ISSN 2359-6457



38ª Edição

Incitel

Congresso de Iniciação Científica do Inatel

27 A 29
MAIO | 2026

Apoio:

FAPEMIG

CNPq

SBC

Realização:

Inatel

inatel.br/incitel

I373

Incitel 2026 – Congresso de Iniciação Científica do Inatel
(2026 - Santa Rita do Sapucaí - MG)

Anais do 38º Congresso de Iniciação Científica do Inatel – Incitel, 27 a 29 de maio 2026, Santa Rita do Sapucaí, Minas Gerais, Brasil. Publicado por Karina Perez Mokarzel Carneiro (Org.). – Santa Rita do Sapucaí - MG: Instituto Nacional de Telecomunicações - INATEL, 2026.

Modo de acesso: https://inatel.br/incitel/ (Anual)
ISSN 2359-6457

462 p. il.

1. Engenharia – I. Título. II. CARNEIRO, Karina Perez Mokarzel. III. Instituto Nacional de Telecomunicações – INATEL.

CDU – 621

Ficha catalográfica elaborada pelo Bibliotecário da Instituição
Paulo Eduardo de Almeida CRB6-2718

Catálogo Sistemático

Engenharia Biomédica	CDU – 57.089
Engenharia de Software	CDU – 004.41
Engenharia da Computação	CDU – 004.41
Engenharia de Telecomunicações	CDU – 621.39
Engenharia de Controle e Automação	CDU – 681.5
Engenharia de Produção	CDU – 658.5
Empreendedorismo	CDU – 658.3

Mensagem da coordenação

Foi com grande satisfação que recebemos pesquisadores, estudantes e profissionais no INCITEL 2026, evento promovido pelo Instituto Nacional de Telecomunicações (Inatel). Nesta edição, o congresso foi realizado em formato híbrido, reunindo apresentações de trabalhos tanto presenciais quanto on-line, ampliando as oportunidades de participação.

O INCITEL tem como principal propósito promover a divulgação da produção científica e tecnológica, incentivando a troca de ideias, experiências e conhecimentos entre pesquisadores de diversas áreas da ciência.

A realização de um congresso dessa magnitude é fruto do trabalho conjunto de muitas pessoas. Por isso, expresso meu sincero agradecimento ao aluno e estagiário Kevin Lopes do Vale e ao colaborador Israel Ferreira da Silva, cuja dedicação e comprometimento foram essenciais para o sucesso desta edição.

Estendo também meus agradecimentos às equipes da Assessoria de Comunicação e Marketing (ASCOM), do Centro de Registros Acadêmicos (CRA), da Coordenação de Cursos, do Inatel Cultural e da Prefeitura do Campus. O empenho, a competência e a colaboração de cada profissional foram fundamentais para a organização e a excelência deste evento.

Por fim, agradeço a todos os autores, orientadores, avaliadores e participantes, que, com suas contribuições, tornaram o INCITEL 2026 um ambiente para a disseminação do conhecimento e o fortalecimento da pesquisa científica.

Esperamos que os trabalhos apresentados nestes anais contribuam para o avanço da ciência, da tecnologia e da formação de novos pesquisadores.



Profª. Karina Perez Mokarzel Carneiro
Coordenadora Geral – Incitel 2026

SESSÕES

Sessão Técnica 1 – Computação e Software I

Gerenciamento de Visitas Hospitalares
.....1
Fernando Faluxis Siqueira Pereira, Marcela de Cássia Ribeiro, Rita Elizabeth Santos de Almeida, Elisa Renno de Carneiro Dester e Fabiano Valias de Carvalho

Desenvolvimento de um Sistema Web para Gestão de Espaço de Vivência
.....7
Augusto Farias de Lima, Afonso Farias de Lima, Kelven Bertanha do Nascimento, Stefany Kaori Omori, Jéssica Antonio Delgado, Isaque Katahira e Fábio Nogueira de Queiroz

Desenvolvimento de uma Plataforma Digital para Apoio à Inserção Profissional
.....12
Daniel Couto Resende, Douglas Eduardo Gabriel, Luis Miguel Nunes da Silva, Geovane Feliciano, Jéssica Antonio Delgado, Isaque Katahira e Fábio Nogueira de Queiroz

Desenvolvimento de uma Plataforma Web para Gestão de Treinos e Suplementação em Conformidade com a LGPD
.....17
Gabriel Bacchiega de Jesus, Amandha Côco, Andrei dos Santos Pinto, Jéssica Antonio Delgado, Isaque Katahira e Fábio Nogueira de Queiroz

Redes Direcionadas de Transições de Códons dos Quatro Sorotipos do Vírus Dengue: Uma Análise Comparativa Baseada na Abordagem BASiNET
.....23
Kelven Bertanha do Nascimento, Lavínia Maria Barbosa de Miranda, Jéssica Antonio Delgado, Isaque Katahira e Fábio Nogueira de Queiroz

Sessão Técnica 2 – Biomédica I

Organização da Competição Nacional BioChallenge Brasil 2025
.....29
Felipe Fonseca Vidal Prado, Stefanie Barbosa Ancelmo, Rita Elizabeth Santos de Almeida, Luma Rissatti Borges do Prado e Elisa Renno de Carneiro Dester

Acessibilidade Física e Sensorial no Ensino Superior: Um Estudo de Caso para a Promoção da Inclusão no Campus do Inatel
.....35

Luiz Otavio Ribeiro de Paiva, Arthur Felipe Vargas, Isabel Francine Mendes e Fabiano Valias de Carvalho

Acessibilidade Digital: Avaliação de Concordância do Portal Acadêmico do Inatel com Ferramentas Automatizadas apoiando-se nas Normas WCAG 2.2
.....41
Wagner Dourado Felix da Silva, Isabel Francine Mendes, Rita Elizabeth Santos de Almeida e Fabiano Valias de Carvalho

Desenvolvimento de uma Ferramenta para Apoio aos Postos de Saúde da Família no Município de Santa Rita do Sapucaí
.....47
Luan Robert Aparecido Santos, Rita Elizabeth Santos de Almeida, Joyce Hemily Amorim Vieira, Victor Arruda Gorgal, Fabiano Valias de Carvalho e Pedro Paulo de Souza Silva

Classificação de Imagens Colposcópicas com Redes Neurais Convolucionais para Detecção de Lesões Cervicais
.....53
Gislaine Abreu Mendonça, Lívia Toledo Gandini, José Andery Carneiro, Márcio Erik Franco e Paulo Roberto Maia

Sessão Técnica 3 – Computação e Software II

Técnicas de Esteganálise Estatística para Detecção de Mensagens Ocultas em Imagens Digitais
.....59
Douglas Miguel Ferreira, Igor Nogueira Olívio, Ítalo Augusto de Sousa Tacca e Evandro César Vilas Boas

Análise Comparativa de Técnicas de Esteganografia em Imagens Digitais: Domínio Espacial Versus Domínio da Frequência
.....65
Douglas Miguel Ferreira, Gustavo Silva Marques Ribeiro, Guilherme Pedro Aquino e Evandro César Vilas Boas

ACC – Algoritmo, Código e Conquista: Um Guia de Estudos para a Maratona de Programação SBC
.....70
João Pedro Schlindwein e Kleyton Hoffmann

Sistema em Tempo Real para Detecção de Gestos Manuais da Língua Brasileira de Sinais com Aprendizado de Máquina

.....76
Gabriel Rech Garcia de Moura e Kleyton Hoffmann

Percepções de Estudantes de Tecnologia da Informação sobre a Utilização Ética do ChatGPT no Ensino
.....81
Patrícia Alves de Sousa, Samuel Sousa Teles, Luis Felipe de Freitas Gomes, Igor Ramsés Temóteo dos Santos, Rubens Abraão da Silva Sousa e Alex Lacerda Ramos

Sessão Técnica 4 – Computação e Software III

Collaborative Tool to Ease the Collection of Datasets for Training Computer Vision Models for Retail Establishments
.....87
Marcelo Henrique de Oliveira Pereira Vilela e Renzo Paranaíba Mesquita

Desenvolvimento de uma Plataforma para Otimização da Orientação Psicopedagógica Realizada pelo NOE e Mind++ Lab
.....93
Tiago Augusto Costa Carvalho, Thomas Ceccon Guimarães, Guilherme Augusto Barucke Marcondes e Ludimila Rocha da Silveira

Checklist-Based Code Review Process: An Experience Report
.....98
Pedro Alyson Teixeira dos Santos, Samuel Sousa Teles, Rubens Abraão da Silva Sousa, Manoel Lopes Filho e Alex Lacerda Ramos

Comparing Package-Based and Sprint-Based Delivery Structures: An Experience Report in an RD&I Environment
.....104
Luis Felipe de Freitas Gomes, Leina Soares Lima, Rubens Abraão da Silva Sousa, Manoel Lopes Filho, Alex Lacerda Ramos e Ronaldo Tadeu Pontes Milfont

O Papel do Laboratório de PD&I no Desenvolvimento de Competências entre a Prática e a Teoria Durante a Graduação: Um Relato de Experiência
.....110
Luis Felipe de Freitas Gomes, Larissa Lacerda, Samuel Sousa Teles, Igor Ramsés Temóteo dos Santos, Rubens Abraão da Silva Sousa e Manoel Lopes Filho

Sessão Técnica 5 – Telecomunicações I

Métricas de Efetividade para Análise de Vulnerabilidades em Aplicações com Bancos de Dados: Um Comparativo NIST e DevSecOps
.....**116**
Bruno Rezende Higino de Cuba, Laura Pivoto Ambrósio, Álvaro Lúcio Almeida Ribeiro, Guilherme Pedro Aquino e Evandro César Vilas Boas

Caderno Instrutivo sobre Redes PON
.....**121**
Aryel Lopes de Almeida Lima, Bruno Rezende Higino de Cuba, Felipe Penha Pereira Duarte, Evandro César Vilas Boas e Guilherme Pedro Aquino

Implementação de uma Plataforma de Monitoramento Remoto para Sistemas Power-over-Fiber
.....**127**
Breno C. do Nascimento, Eduardo M. Bertozzi, Letícia Carneiro de Souza e Francisco Martins Portelinha Junior

Emprego de Rádio Definido por Software como MAGE-COM em Veículos de Superfície Não Tripulados
.....**132**
Daniel Scorzello Lopes, Paulo Victor Gonçalves Sampaio e Arthur Pinheiro de Araújo Costa

Engenharia Reversa de um Oxímetro de Pulso Comercial P-01 para a Aquisição de Dados Fisiológicos Mediante ESP32 em Aplicações IoT
.....**138**
Luis Diego Ferreyra Contreras, Melissa Andrea Herrera Tambo e Eynar Calle Viles

Sessão Técnica 6 – Telecomunicações II

Segurança em Ambientes de Computação em Nuvem: Superfície de Ataque e Falhas Sistêmicas
.....**144**
João Pedro Piceli, Camile Brunieri Laraia e Evandro César Vilas Boas

Estudo e Implementação de Técnicas de Inteligência Artificial Aplicadas na Otimização de Sistemas de Comunicação sem Fio Assistidos por STAR-RIS
.....**148**
Gislaine Abreu Mendonça e Victoria Dala Pegorara Souto

Gaze Tracking: Controle de Cursor em Tempo Real Baseado em Rastreamento Ocular

.....154
Guilherme S. F. Almeida, Luiz E. A. Nora, Evandro C. Vilas Boas e Felipe A. P. Figueiredo

Instalação do Sistema Operacional Debian Minimal em Dispositivo com Processador S905X4
.....160
Matheus Dapper Alves Leite e Kleyton Hoffmann

Desenvolvimento de um Sistema Embarcado com ESP32 para Aquisição e Monitoramento em Nuvem de Sinais de Qualidade de Energia
.....166
Isabela Cristina Munzlinger, Nakul S., Marconi Januário, Jamuna Kamaraj, S. Angalaeswari e Renato Gregolon Scortegagna

Sessão Técnica 7 – Elétrica e Produção

Gamificação no Ambiente Industrial: Uma Estratégia para Redução da Fadiga Mental e Aumento da Produtividade em Atividades Repetitivas
.....172
Ana Beatriz Faria Gomes e Vinícius Antônio Montgomery de Miranda

Análise e Gerenciamento de Processos Industriais
.....176
Antônio Borges Oliveira e Vinícius Antônio Montgomery de Miranda

Qualidade de Energia Elétrica: Simulador de Distúrbios e Mitigação
.....182
Egídio Raimundo Neto, Geovane Mantuan Alvim, Júlia da Silva Maciel e Natan Barreto de Paula

Desenvolvimento e Validação de um Sistema Automatizado para Calibração e Caracterização Metrológica de Sensores de Umidade de Solo de Baixo Custo
.....188
Carlos Alberto de Francisco e Mateus Sábio Martins

Projeto de um Pré-Amplificador de Áudio de Alto Desempenho sem Realimentação Negativa Global
.....194
Carlos Alberto de Francisco e Lucas Alan da Conceição Batista

Desenvolvimento de um Tubo de Impedância para Caracterização Acústica de Materiais
.....200
Carlos Alberto de Francisco e Gabriel Farias de Oliveira Junior

Sessão Técnica 8 – Biomédica II

Otimização de Interface de Sistemas IoT Aplicados em Ambiente Hospitalar
.....**206**
Pedro Henrique de Paiva Carvalho, Rafael Gonçalves Magalhães e Luma Rissatti Borges do Prado

Sistema de Inteligência Artificial e Redes Neurais para Suporte à Tomada de Decisão em Tratamentos Fisioterapêuticos
.....**211**
Filipe Bueno Vilela e João Carlos Rodrigues Franco

Desenvolvimento de um Simulador Interativo de Máquina de Anestesia Baseado em ESP32 para a Educação em Engenharia Biomédica
.....**215**
Marcelo Fabian Cruz Guizada, Mauricio Marcelo Peredo Claros e Rodrigo Martinez Severich

Development of a Low-Cost Elbow Passive Rehabilitation Orthosis with Digital Twin and IoT
.....**221**
Sergio Antonio Vásquez Martínez e Rodrigo Martinez Severich

Development of a Non-Invasive Hemoglobin Measurement System for Early Anemia Detection
.....**226**
Luis Fernando Arteaga Aldana e Alberto J. Armando Aguilar Bustillos

Sessão Técnica 9 – Computação e Software IV

Ordem de Volatilidade como Fundamento Metodológico na Preservação de Evidências Digitais
.....**233**
Vinícius Santos Araújo, Guilherme Pedro Aquino e Evandro César Vilas Boas

Estudo de Modelo de Comunicação Segura Baseado em Confidencialidade, Integridade, Autenticidade e Irretratabilidade
.....**239**
Vinícius Santos Araújo, Guilherme Pedro Aquino e Evandro César Vilas Boas

Desenvolvimento de um Sistema Lúdico Baseado em Realidade Virtual Não Imersiva para Reabilitação de Mãos com Artrite Reumatoide
243
Ervin Emanuel Orellana Pardo e Bianca Camila Garnica Pericon

Arquitetura Inteligente Baseada em LoRa para Proteção Ativa de Gado Bovino Mediante Monitoramento de Tensão em Tempo Real
249
Roly Serrano Quiroz, Eynar Calle Viles e Katerin Deysi Vasquez Benito

Mobile Application for Medicinal Plant Leaf Species Identification and Disease Detection Using Deep Learning
255
Maria Laura Nuñez Jaillita, Juan Abiel Iriarte Colque e Eynar Calle Viles

Sessão Técnica 10 – Biomédica III

Desenvolvimento de um Protótipo para Detecção e Avaliação de Risco de Queda em Ambiente Hospitalar
261
Fernando Puebla Stein, Matheus Vieira Honório de Souza e Rita Elizabeth Santos de Almeida

Estudo de Técnicas de Inteligência Artificial para Aprimoramento de Óculos Assistivos para Pessoas com Deficiência Visual
266
Gustavo Almeida Barrios, Lavínia Silva Andrade, Rita Elizabeth Santos de Almeida, Roberto Luiz Assad Pinheiro e Filipe Bueno Vilela

HelpFull: Monitoramento Emocional com Suporte de Inteligência Artificial em uma Plataforma Web
271
Pedro Henrique Diana Andolpho Evaristo, Arthur Pereira da Silva, Otto Simões Miranda, Pedro Otávio de Oliveira Carlos, Jéssica Antonio Delgado, Isaque Katahira e Fábio Nogueira de Queiroz

Desenvolvimento de um Sistema Embarcado Baseado em Visão Artificial para Avaliação Pupilar e Oculomotora
277
Brayan Armando Zeballos Rios, Maria Rene Ledezma Flores e Eynar Calle Viles

Estimativa de Pressão Arterial e Glicemia por Meio de Oxímetro de Pulso e Inteligência Artificial

.....283
Jhonny Martin Camacho Gonzales, Mauricio Marcelo Peredo Claros e Eynar Calle Viles

Sessão Técnica 11 – Controle e Automação I

Criação de Projeto para Inclusão de Robótica em Colégios de Ensino Médio
.....289
Marco Antônio Shima, Alexandre B. Lugli, Egídio R. Neto, Eugenio M. J. da Silva, Luiz Fernando D. R. Santos, Wanderson E. Saldanha e Maria Luiza Muquem Alves

Estudo de Robôs Inteligentes Aplicados ao Futebol de Robôs: Very Small Size Soccer (VSSS)
.....294
Kaynan H. Dias de Melo, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Enzo R. R. Lemos, Gustavo Mundim Penaforte, Nathan A. Costa e Wanderson E. Saldanha

Estudo de Robôs Inteligentes Aplicados ao Combate de Sumô LEGO®
.....299
Alexandre Baratella Lugli, Eugenio Moreira Jonas da Silva, Egídio Raimundo Neto, Luís Fernando Domiciano Ribeiro dos Santos, Wanderson Eleutério Saldanha, Rodrigo Geraldo Silva Santos e Luiz Gustavo de Oliveira

Estudo para Desenvolvimento de um Robô de Combate para a Categoria Antweight – 454 g (1 lb)
.....304
Téo Sacramento Marques, Alexandre B. Lugli, Egídio R. Neto, Eugenio M. J. da Silva, Felipe S. Barbosa, Wanderson E. Saldanha, Luiz Fernando D. R. Santos e Wagner L. C. Anastácio

Estudo de Robôs Inteligentes Aplicados ao Futebol de Robôs Virtual 2D
.....309
Renata Bernardes Cabral, Júlia R. Garcia, Débora de Farias Negri, Alexandre B. Lugli, Eugênio M. J. da Silva, Egídio R. Neto, Luiz Fernando D. R. dos Santos e Wanderson E. Saldanha

Sessão Técnica 12 – Ensino, Elétrica, Administração, Inovação e Empreendedorismo

Matemática para Engenharias: Uma Adequação aos Conceitos e Habilidades Necessários
.....313
Ana Luísa Lyra Pavanelli, Felipe Ferreira de Carvalho Gabriel Pereira, Guilherme Muniz de

Oliveira Reis, Lucas do Lago Figueiredo, Maria Clara Bernardes Ferrer e Luiz Felipe Simões de Godoy

Estudo para o Desenvolvimento de um Aplicativo de Alfabetização Matemática **318**

Rayana Fernanda dos Santos Silva, E. L. J. Junior, G. F. Finazzi, L. G. Garcia, T. E. P. Oliveira, J. P. T. Dias e Wanderson E. Saldanha

Modelo Estacionário de Ondas com Interferência Destrutiva **321**

Pedro Emanuel Almeida Bernardes Rodrigues, Pedro Henrique Gaviolli e Felipe Emanoel Chaves

Impacto do IF Mais Empreendedor em Micro e Pequenos Negócios **326**

Pietro Luís Bertoldo Silva e Jéssica dos Santos Leite Gonella

Estudo da Operação ZVS em Fonte Chaveada Aplicada a Circuito Ressonante **330**

Felipe Coser, Jean Patrick Prigol e Renato Gregolon Scortegagna

Sessão Técnica 13 – Controle e Automação II

Computer Vision Applied to a Robotic Arm: A Case Study **336**

Fernando Barbosa Mota Júnior, Diogo Antonio Castro dos Santos, Carl Betsa Costa, Raphael Henrique de Oliveira Centenário, João Paulo Carvalho Henriques e Renzo Paranaíba Mesquita

Desenvolvimento de um Robô para a Categoria Perseguidor de Linha **344**

Maria Fernanda P. Castro, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Wanderson E. Saldanha, Luiz Fernando D. R. Santos e Giovanna Gama Fernandino

Estudo de Desenvolvimento de um Robô para a Categoria Seguidor de Linha **348**

Caroline Ferreira de Souza, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Luiz Fernando D. R. Santos, Pedro Paulo Junqueira e Wanderson E. Saldanha

Estudo para Novos Sistemas de Navegação e Controle para Robôs Autônomos Trekking **355**

Mayara do Prado Almeida, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Fuad Murad, Luiz Fernando D. R. Santos e Wanderson E. Saldanha

Estudo de Drones Inteligentes Aplicados à Navegação
360
Samuel Silva Oliveira, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Breno Veloso Eboli, Caio Pereira Bonvicine, Luiz Fernando D. R. dos Santos, Lucas Pereira Aguiar, Wanderson E. Saldanha e Rayana Fernanda dos Santos Silva

Sessão Técnica 14 – Computação e Software V

Automatic Classification of Coffee Fruits Using Machine Learning and Computer Vision
364
Henrique Rodrigues Mendonça e Samuel Baraldi Mafra

Automatização da Extração de Dados de Metassuperfícies para Aprendizado de Máquina
370
Matheus Augusto de Faria Silveira e William Orivaldo Faria Carvalho

Estratégias para Redução de Falhas Humanas em Implementações por meio de Pipelines CI/CD
374
Felipe Mitsunori Uehara, Diego Eduardo da Silva Santos e João Emmanuel D'Alkmin Neves

Roteamento Adaptativo de Consultas para Chatbot Institucional com RAG em Ambiente de Hardware Restrito
380
Gabriel de Souza Lima, Júlio César Mansano e João Emmanuel D'Alkmin Neves

Desenvolvimento de Chatbot Acadêmico com Princípios de Acessibilidade Digital
387
Aluan Luiz Maturi Horschutz, Fernanda Diniz Marabin e João Emmanuel D'Alkmin Neves

Sessão Técnica 15 – Biomédica IV

Estudo de Caso do Contador Geiger-Müller
393
Felipe Emanuel Chaves, Rita Elizabeth Santos de Almeida, Brenno Canndellas, Nicholas Lima do Nascimento e Roger Pereira Freitas

Desenvolvimento e Avaliação de Adequação de Prótese Impressa Ativa para Animal de Pequeno Porte Amputado

.....399
Filipe Bueno Vilela, Maria Clara Bernardes Ferrer e Lucas Toledo Capello

Mechanical Stimulation System Prototype for Bronchial Secretion Clearance in Preterm Infants
.....404
Mateo Sebastian Rojas Flores, Eynar Calle Viles e Tatiana Torrico Sanchez

Avaliação Comparativa dos Níveis de Estresse e Ansiedade em Estudantes de Engenharia Biomédica e Medicina da Universidad Privada del Valle, Cochabamba
.....409
Brayan Javier Mejia Saavedra, Melissa Andrea Herrera Tambo e Marcelo Fabian Cruz Guizada

Projeto e Integração de um Kit IoT de Sinais Vitais para Apoio à Triagem Remota na Plataforma DocIA
.....415
Brayan Javier Mejia Saavedra, Maria Laura Nuñez Jaillita, Mishel Bravo Almendras, Sergio Antonio Vasquez Martinez e Luis Diego Ferreyra Contreras

Sessão Técnica 16 – Controle e Automação III

Estudo para Desenvolvimento de um Robô para a Categoria Mini Sumô
.....421
Maria Luiza Muquem Alves, Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Luiz Fernando D. R. Santos, André R. Silva, Wanderson E. Saldanha e Vithor Spirandeli

Desenvolvimento de Padrão para Robô de Combate Categoria Cupim 454 g
.....426
Caio de Castro Yarouhas, David B. J. do Prado, Alexandre B. Lugli, Eugênio M. J. da Silva, Egídio R. Neto, Luiz Fernando D. R. dos Santos e Wanderson E. Saldanha

Architecture of an Autonomous Vehicle with ROS for Mapping and Navigation in Static Environments
.....430
Carlos Leonardo Lazzari, Murilo Ribeiro Bonato e Kleyton Hoffmann

Robô Didático como Complemento Terapêutico para o Desenvolvimento Social de Crianças com Transtorno do Espectro Autista (TEA)
.....436
Bianca Camila Garnica Pericon e Ervin Emanuel Orellana Pardo

Sistema Inteligente de Monitoramento de Apiários na Bolívia Baseado em IA/ML com IoT e LPWAN

.....**442**
Patricia Nataly Flores Ponce, Edzel Alejandro Escobar Cueto, Eynar Calle Viles e Edgar Roberto Ramos Silvestre

Gerenciamento de visitas hospitalares

Fernando Faluxis Siqueira Pereira, Marcela de Cássia Ribeiro, Rita Elizabeth Santos de Almeida,
Prof. Elisa Rennó de Carneiro Dester, Fabiano Valias de Carvalho
Instituto Nacional de Telecomunicações - Inatel
f.faluxis@ges.inatel.br, marcela.ribeiro@ges.inatel.br, r.elizabeth@inatel.br, elisa.renno@inatel.br, fabiano@inatel.br

***Abstract*—This work presents the development of a web application for hospital visitor management, created in partnership with Hospital Antônio Moreira da Costa (HAMC). The system was implemented using a client–server architecture, with a Node.js backend, React frontend, and MySQL database, in addition to integration with the hospital's institutional database. The platform enables visitor registration and management, organizing information in a structured digital environment. Functional tests conducted in a local environment confirmed the proper operation of the application.**

***Index Terms*—Hospital visit management, Visitor control, Hospital security, Patient monitoring, Hospital application.**

***Resumo*—Este trabalho apresenta o desenvolvimento de uma aplicação web para gerenciamento de visitas hospitalares, criada em parceria com o Hospital Antônio Moreira da Costa (HAMC). O sistema foi desenvolvido utilizando arquitetura cliente-servidor, com *backend* em Node.js, *frontend* em *React* e banco de dados MySQL, além de integração com o banco institucional do hospital. A plataforma permite o cadastro e controle de visitantes, organizando as informações de forma digital e estruturada. Testes funcionais realizados em ambiente local demonstraram o correto funcionamento da aplicação.**

***Palavras chave*—Gerenciamento de visitas hospitalares, Controle de visitantes, Segurança hospitalar, Monitoramento de pacientes, Aplicativo hospitalar.**

I. Introdução

Os visitantes desempenham papel fundamental no processo de cuidado, proporcionando ao paciente hospitalizado uma sensação de familiaridade e suporte psicológico, fatores reconhecidamente importantes para a recuperação. A presença de familiares e amigos mostrou-se essencial para pacientes que necessitaram de internação hospitalar, uma vez que o contato social contribuiu significativamente para o bem-estar emocional do paciente e para a redução da ansiedade provocada pelo ambiente hospitalar [1, 2].

Nos hospitais, normas específicas foram estabelecidas para assegurar a segurança de pacientes e colaboradores durante as visitas, fundamentadas principalmente nas diretrizes da Agência Nacional de Vigilância Sanitária (ANVISA) e do Ministério da Saúde, que orientam o controle da circulação de pessoas, a prevenção de infecções relacionadas à assistência à saúde e a garantia da rastreabilidade de visitantes [3, 4, 5]. Tradicionalmente, o registro de visitantes ocorre de forma manual, por meio de cadernos físicos, nos quais são anotados dados como nome do paciente, data e horário da visita, além de informações pessoais do visitante, como Cadastro de pessoas Físicas(CPF) e número de telefone [1]. Entretanto, esse método apresenta limitações inerentes a processos manuais, incluindo maior suscetibilidade a falhas de registro, dificuldade na verificação do cumprimento de limites diários de visitas e baixa acessibilidade às informações, especialmente em instituições com elevada demanda de pacientes [6, 7, 8]. No Brasil, hospitais estaduais estimam atender aproximadamente 130 mil pessoas por mês, resultando em uma média de cerca de 4.300 visitas diárias, com números ainda mais elevados em hospitais de grande porte localizados nas capitais [9]. Ademais, o aumento de atendimentos hospitalares no Sistema Único de Saúde (SUS) em 10% contribuiu para o crescimento do fluxo de visitantes em unidades públicas [10, 11].

Métodos manuais de controle de visitas mostraram-se inadequados, sobretudo em hospitais com elevada taxa de ocupação e fluxo significativo de pacientes. O registro manual tornou-se suscetível a erros, forneceu informações incompletas e dificultou a gestão do limite legal de visitas diárias assegurado ao paciente [12, 13]. Embora *softwares* comerciais de gerenciamento de entrada de visitantes, ofereçam funcionalidades avançadas de controle de acesso, pré-cadastro e armazenamento seguro de dados, no entanto o seu elevado custo restringiu a adoção em larga escala, especialmente em hospitais públicos [14, 15].

Diante desse contexto, o presente trabalho tem como objetivo o desenvolvimento de um *software* para o gerenciamento da entrada e saída de visitantes hospitalares. Essa demanda foi identificada pelo Hospital Antônio Moreira da Costa (HAMC), em parceria com o eHealth Innovation Center, evidenciando a relevância científica e tecnológica do projeto, bem como sua contribuição para a melhoria da gestão hospitalar.

II. Metodologia

A. Caracterização da pesquisa

O trabalho caracteriza-se como uma pesquisa aplicada, uma vez que visa ao desenvolvimento de uma solução computacional voltada à resolução de um problema real identificado em ambiente hospitalar.

Quanto aos procedimentos metodológicos, o estudo pode ser classificado como um estudo de caso, uma vez que a solução foi concebida e desenvolvida especificamente para atender às demandas do HAMC, considerando suas normas internas, estrutura organizacional e infraestrutura tecnológica. Paralelamente, o trabalho também se enquadra como um desenvolvimento de sistema, com validação prática do funcionamento da aplicação em ambiente computacional.

B. Ambiente do estudo

O desenvolvimento do sistema foi realizado no contexto do HAMC, instituição hospitalar localizada em Santa Rita do Sapucaí, no estado de Minas Gerais. A demanda pelo sistema surgiu a partir da necessidade de aprimorar o controle de

entrada e saída de visitantes, atualmente realizado por meio de registros manuais, que apresentam limitações relacionadas à confiabilidade, rastreabilidade e organização das informações.

Embora a solução tenha sido projetada especificamente para o HAMC, levando em consideração seu banco de dados e seus fluxos operacionais, sua arquitetura permite que, com pequenas adaptações, o sistema possa ser utilizado em outras instituições hospitalares com necessidades semelhantes.

Durante a fase de desenvolvimento, a aplicação foi executada em ambiente local, na máquina dos pesquisadores. No entanto, a arquitetura do sistema foi planejada para que, ao final do projeto e juntamente com a submissão do artigo científico, a solução seja implantada em um servidor local (*datacenter* interno) do hospital, operando exclusivamente na rede interna da instituição, sem acesso externo à Internet pública.

O sistema foi desenvolvido com foco no setor de recepção hospitalar, responsável pelo controle e registro das visitas. Embora os profissionais da recepção não tenham participado diretamente do desenvolvimento do *software*, o fluxo de trabalho desse setor foi considerado como base para a definição das funcionalidades da plataforma.

C. Levantamento de requisitos

O levantamento de requisitos do sistema foi realizado a partir de uma abordagem mista, combinando análise documental e validação empírica. Inicialmente, foi fornecida à equipe de desenvolvimento uma ementa do problema, contendo informações essenciais sobre o processo de controle de visitas, os dados envolvidos e as principais necessidades operacionais.

Com base nas necessidades informadas, a equipe desenvolveu a modelagem inicial do sistema. Posteriormente, ocorreram reuniões com profissionais do hospital, com o objetivo de esclarecer dúvidas, validar os requisitos levantados e garantir que a solução estivesse alinhada à realidade operacional da instituição. Essas interações permitiram ajustes progressivos no sistema, assegurando maior aderência às práticas adotadas no ambiente hospitalar.

Os requisitos identificados foram classificados em requisitos funcionais, como o cadastro de pacientes e visitantes, registro de horários de entrada e saída, consulta de informações, e requisitos não funcionais, relacionados à segurança, organização dos dados, facilidade de uso e operação restrita ao ambiente interno do hospital.

O processo de desenvolvimento foi documentado ao longo do projeto por meio de relatórios técnicos, nos quais foram registrados o progresso das etapas, decisões de projeto, ajustes realizados e validações efetuadas.

D. Arquitetura geral do sistema

A solução foi desenvolvida seguindo uma arquitetura cliente-servidor, com separação entre *frontend* e *backend*, abordagem amplamente utilizada no desenvolvimento de aplicações *web*. Esse modelo foi adotado por favorecer a modularidade, a escalabilidade e a manutenção do sistema, além de permitir futuras expansões sem comprometer o funcionamento da aplicação.

O sistema foi dividido em dois módulos principais:

- *Backend*, responsável pelo processamento das regras de negócio, comunicação com o banco de dados e disponibilização de uma *API* para o *frontend*;
- *Frontend*, responsável pela interface gráfica e interação com os usuários do sistema.

Essa separação permite que cada camada evolua de forma independente, além de facilitar a integração com sistemas já existentes no ambiente hospitalar e possibilitar adaptações futuras com menor impacto estrutural.

E. Desenvolvimento do backend

O *backend* da aplicação foi desenvolvido utilizando Node.js, com a linguagem JavaScript, e o *framework* Express. A escolha dessas tecnologias deve-se ao seu bom desempenho em aplicações orientadas a eventos, capacidade de lidar com múltiplas requisições simultâneas e ampla adoção no desenvolvimento de *APIs RESTful*. Essas características são especialmente relevantes em ambientes hospitalares, onde a disponibilidade e a rapidez no acesso às informações são essenciais.

A organização do código no *backend* seguiu uma estrutura lógica baseada em responsabilidades, promovendo maior clareza, organização e facilidade de manutenção. Foram definidos os seguintes componentes:

- **Rotas:** responsáveis por definir os *endpoints* da aplicação;
- **Controladores:** responsáveis por receber as requisições, validar dados e coordenar as ações;
- **Modelos:** utilizados para representar as entidades do sistema, como pacientes, visitantes e registros de visitas;
- **Camada de acesso a dados:** responsável pela comunicação com o banco de dados.

Essa organização contribuiu para a separação de responsabilidades, facilitando a identificação de erros, a reutilização de código e a evolução futura do sistema.

F. Banco de dados e integração hospitalar

Para o armazenamento das informações, foi utilizado o sistema gerenciador de banco de dados MySQL, no qual foram criadas todas as tabelas necessárias para o funcionamento da plataforma, incluindo entidades relacionadas a visitantes, usuários do sistema e registros de visitas. O MySQL foi escolhido devido à sua robustez, confiabilidade e amplo uso em aplicações *web*, além de apresentar bom desempenho em operações transacionais e facilidade de administração. Outro fator relevante para sua escolha foi a compatibilidade com a infraestrutura tecnológica existente e a disponibilidade de documentação e suporte, características que favorecem a manutenção e a sustentabilidade da solução ao longo do tempo.

Paralelamente, o hospital já possuía um banco de dados institucional, baseado no sistema Firebird, contendo informações oficiais dos pacientes. Dessa forma, o projeto exigiu a implementação de um processo de integração entre bancos de dados, no qual os dados de pacientes são obtidos diretamente do banco institucional do hospital por meio de uma VIEW, com permissões restritas apenas à operação de leitura (*SELECT*).

Durante o desenvolvimento da aplicação, a interação com o banco de dados institucional foi realizada por meio

de ferramentas de gerenciamento e acesso remoto. Inicialmente, utilizou-se a ferramenta DBeaver para a exploração e visualização da estrutura do banco de dados. Para acesso ao ambiente do hospital e consulta direta às informações armazenadas no servidor institucional, foi utilizada a ferramenta AnyDesk, possibilitando a validação dos dados e a compreensão da organização das tabelas utilizadas pelo sistema hospitalar.

Essa estratégia garantiu que os dados sensíveis dos pacientes permanecessem sob controle institucional, ao mesmo tempo em que possibilitou a integração necessária para o funcionamento da plataforma.

G. Desenvolvimento do frontend

O *frontend* da aplicação foi desenvolvido utilizando a biblioteca *React*, em conjunto com *HyperText Markup Language* (HTML) e *Cascading Style Sheets* (CSS) para a construção da interface gráfica. A escolha do *React* deve-se à sua abordagem baseada em componentes reutilizáveis, que favorece a organização do código, a manutenção e a escalabilidade da aplicação. Além disso, sua capacidade de atualização dinâmica da interface sem recarregamento completo da página contribui para uma experiência de uso mais fluida e eficiente, característica importante em ambientes hospitalares onde a agilidade no atendimento é essencial.

A utilização dessas tecnologias possibilitou a criação de uma interface dinâmica, organizada e de fácil utilização, adequada ao ambiente hospitalar.

A interface do sistema foi estruturada em diferentes páginas, incluindo:

- Página de *login*;
- Página inicial;
- Tela de cadastro e listagem de pacientes;
- Tela de cadastro de visitantes vinculados a cada paciente.

O fluxo de uso foi projetado de forma que a lista de pacientes seja obtida automaticamente a partir do banco de dados do hospital, enquanto o cadastro de visitantes é realizado manualmente pelos profissionais da recepção.

A preocupação com a usabilidade foi constante ao longo do desenvolvimento, sendo realizados ajustes sucessivos no *layout*, na organização visual das informações e na disposição dos campos de cadastro, a partir de testes internos e validações realizadas pela equipe.

H. Controle de acesso e segurança

O sistema implementa um mecanismo de autenticação por *login* e senha, garantindo que apenas usuários previamente cadastrados possam acessar a plataforma. Cada profissional da recepção possui um cadastro individual, o que possibilita maior controle sobre o uso do sistema.

Os dados sensíveis, como CPF e telefone dos visitantes, são tratados como campos obrigatórios, com validações que impedem o envio de informações incompletas. O acesso ao sistema é restrito exclusivamente à equipe de recepção, reduzindo o risco de exposição indevida das informações.

Durante o desenvolvimento da aplicação, foram consideradas diretrizes da Lei Geral de Proteção de Dados (LGPD) para orientar o tratamento das informações pessoais. Dessa forma, o sistema foi projetado para coletar apenas os dados necessários para o gerenciamento das visitas, além de restringir o acesso às informações aos profissionais responsáveis pelo processo de recepção.

I. Testes e validação

A validação do sistema foi realizada por meio de testes manuais em ambiente local durante o processo de desenvolvimento da aplicação. Inicialmente, o acesso direto ao banco de dados institucional do hospital não estava disponível, o que impossibilitou a utilização de dados reais de pacientes nessa etapa do projeto. Dessa forma, para viabilizar os testes da funcionalidade de listagem de pacientes, registros fictícios foram inseridos diretamente no banco de dados MySQL por meio de comandos *Structured Query Language* (SQL).

Por outro lado, as funcionalidades relacionadas ao cadastro e gerenciamento de visitantes foram testadas diretamente na própria plataforma desenvolvida, sendo executadas e avaliadas em ambiente local ao longo do desenvolvimento. Essa abordagem permitiu verificar o funcionamento das operações de cadastro de visitantes, vinculação de visitantes a pacientes e registro dos horários de entrada e saída.

Adicionalmente, foram realizados testes relacionados ao mecanismo de autenticação de usuários, avaliando o comportamento do sistema diante de tentativas de acesso com credenciais válidas e inválidas, a fim de garantir o correto funcionamento do controle de acesso à plataforma.

III. Resultados e Discussão

A. Visão geral do sistema desenvolvido

O sistema desenvolvido consiste em uma plataforma web para gerenciamento de visitas hospitalares, projetada para apoiar a rotina da recepção hospitalar no controle de visitantes vinculados aos pacientes internados.

A plataforma é acessada por meio de navegador web. Em sua configuração final, o sistema encontra-se hospedado em servidor local do hospital, operando dentro da infraestrutura interna do *datacenter* institucional, o que contribui para maior segurança dos dados e conformidade com as políticas internas de tecnologia da informação.

Essa abordagem mostrou-se adequada ao contexto hospitalar, pois reduz custos de implantação, facilita a manutenção e garante que os dados permaneçam sob controle institucional, aspecto essencial no tratamento de informações sensíveis em saúde.

O sistema contemplou múltiplos usuários, sendo que cada recepcionista possui seu próprio cadastro e credenciais de acesso. Embora os usuários compartilhem as mesmas funcionalidades e visualizem as mesmas telas do sistema, a autenticação individual permite controle de acesso, rastreabilidade das ações e organização dos registros, atendendo às boas práticas de sistemas de informação em ambientes hospitalares.

No que se refere aos dados dos pacientes, estes não são cadastrados manualmente pela recepção. As informações dos pacientes são obtidas por meio de integração com o banco de dados institucional do hospital, garantindo consistência, atualização e redução de retrabalho. Essa integração

demonstrou-se essencial para evitar erros de digitação e duplicidade de registros, problemas comuns em processos manuais.

A atuação da recepção no sistema concentra-se, portanto, no cadastro e gerenciamento dos visitantes, os quais são vinculados aos respectivos pacientes previamente registrados no sistema hospitalar.

- Autenticação individual dos profissionais da recepção;
- Acesso à lista de pacientes importada do banco de dados do hospital;
- Cadastro, edição e visualização de visitantes associados a cada paciente;
- Organização centralizada das informações de visitas hospitalares.

A arquitetura da aplicação segue o modelo cliente-servidor, no qual o *frontend* é responsável pela interface e interação com o usuário, enquanto o *backend* realiza o processamento das requisições, regras de negócio e comunicação com o banco de dados. Essa separação favorece a manutenção, escalabilidade e futuras adaptações do sistema, possibilitando sua utilização em outros contextos hospitalares mediante pequenas adequações.

Além disso, a informatização do controle de visitantes representa um avanço significativo em relação aos métodos manuais tradicionalmente utilizados, promovendo maior organização, rastreabilidade e segurança das informações.

B. Tela Inicial do Sistema

A tela inicial do sistema foi desenvolvida com foco na simplicidade e clareza, sendo o primeiro ponto de interação do usuário com a plataforma. Sua função é apresentar o sistema e direcionar corretamente o acesso dos profissionais da recepção.

A interface é composta por uma imagem posicionada à direita da tela e, na região central, um texto descritivo contendo o título “Sistema de Gerenciamento de Visitas Hospitalares”, acompanhado de uma breve explicação sobre a finalidade do sistema, voltada à organização e acompanhamento das visitas hospitalares de forma segura e eficiente.

Abaixo do texto principal, estão disponíveis dois botões de navegação: “Entrar como funcionário”, destinado aos usuários previamente cadastrados, e “Cadastrar funcionário”, utilizado para o registro de novos profissionais da recepção. Essa disposição facilita a compreensão das ações disponíveis e contribui para uma navegação intuitiva, como pode ser observado na Figura 1.



Fig. 1. Tela inicial do sistema de gerenciamento de visitas hospitalares

C. Tela de Cadastro de Funcionário

A tela de cadastro de funcionário foi desenvolvida para permitir o registro dos profissionais da recepção que utilizarão o sistema. Assim como a tela inicial, a interface mantém um layout simples e organizado, com uma imagem ilustrativa posicionada no lado direito e os elementos de interação concentrados na região central da tela.

No formulário de cadastro, são solicitados os seguintes dados: nome completo, e-mail, senha e código de acesso. Todos os campos são de preenchimento obrigatório, garantindo que apenas profissionais autorizados tenham acesso ao sistema.

Ao final do formulário, o usuário pode optar pelo botão “Cadastrar”, responsável por efetivar o registro do funcionário, ou pelo botão “Voltar”, que redireciona para a tela inicial do sistema. Também foi incluído um ícone de navegação que permite o retorno direto à tela inicial do sistema, como pode ser observado na Figura 2.



Fig. 2. Tela de cadastro de funcionário do sistema de gerenciamento de visitas hospitalares

D. Tela de Login

A tela de *login* tem como finalidade autenticar os funcionários da recepção cadastrados no sistema. A interface mantém o mesmo padrão visual das demais telas, com uma imagem ilustrativa posicionada no lado direito e os elementos de interação organizados na região central, como apresentado na Figura 3.

Para acessar a plataforma, o usuário deve informar e-mail e senha, campos obrigatórios para validação das credenciais. O botão “Entrar” permite o acesso ao sistema após a verificação dos dados, enquanto o botão “Voltar” possibilita o retorno à tela inicial.

Adicionalmente, a tela disponibiliza o link “Esqueceu a senha?”, destinado ao processo de recuperação de acesso. Também foi incluído um ícone de navegação que permite o retorno direto à tela inicial do sistema

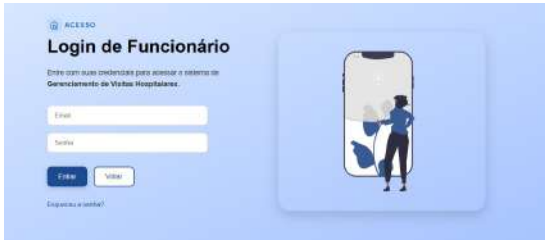


Fig. 3. Tela de *login* do sistema de gerenciamento de visitas hospitalares

E. Tela de Lista de Pacientes

A tela de lista de pacientes tem como objetivo apresentar, de forma organizada, os pacientes internados no hospital, cujos dados são obtidos por meio da integração com o banco de dados institucional.

No topo da interface, encontra-se o botão “Voltar”, que permite o retorno à tela anterior, seguido do título “Lista de

Pacientes", responsável por identificar a funcionalidade apresentada. A tela disponibiliza ainda dois campos de pesquisa, "Pesquisar por nome" e "Pesquisar por CPF", possibilitando a localização rápida de pacientes específicos. Também foi incluído um ícone de navegação que permite o retorno direto à tela inicial do sistema

Os pacientes são exibidos em formato de lista, composta por cartões individuais, que apresentam informações essenciais. Em cada cartão, o botão "Gerenciar Visitantes" permite acessar a funcionalidade de cadastro e controle de visitantes vinculados ao paciente selecionado.

Essa organização contribui para a centralização das informações e para a agilidade no atendimento da recepção, conforme mostrado na Figura 4.



Fig. 4. Tela da lista de pacientes do sistema de gerenciamento de visitas hospitalares

F. Tela de Gerenciamento de Visitantes

A tela de gerenciamento de visitantes é responsável pelo cadastro, acompanhamento e controle da entrada e saída de visitantes e acompanhantes vinculados a um paciente específico. No topo da interface, encontra-se o botão "Voltar", que permite retornar à lista de pacientes, além do título "Visitantes do(a) Paciente", no qual o nome do paciente é exibido dinamicamente conforme a seleção realizada na tela anterior. Também foi incluído um ícone de navegação que permite o retorno direto à tela inicial do sistema

A primeira seção da tela é destinada ao cadastro de visitantes e acompanhantes. Nela, são disponibilizados os campos de nome, CPF e parentesco, além da opção de classificação entre visitante ou acompanhante. O campo de data pode ser preenchido manualmente ou, caso permaneça vazio, é automaticamente preenchido com a data da máquina no momento do cadastro. Após o preenchimento das informações, o botão "Cadastrar" efetiva o registro do visitante.

A tela também disponibiliza a funcionalidade "Exportar para Excel", que permite a geração de uma planilha contendo todos os registros de visitantes associados ao paciente, contribuindo para a organização e rastreabilidade das informações.

Na parte inferior da interface, os visitantes cadastrados são exibidos em forma de cartões. Quando o visitante é classificado como acompanhante, essa informação permanece destacada e fixada no topo do cartão. Cada cartão apresenta dados como CPF, parentesco, data da visita, horário de entrada e horário de saída. O horário de entrada é registrado automaticamente no momento do cadastro, enquanto o botão "Registrar saída" permite indicar o encerramento da visita, preenchendo automaticamente o horário de saída com base no relógio do sistema. Após o registro da saída, o cartão do visitante tem sua cor alterada e passa a indicar o status inativo, facilitando a visualização do fluxo de visitantes, como pode ser observado na Figura 5.



Fig. 5. Tela de gerenciamento de visitantes do sistema de gerenciamento de visitas hospitalares

G. Desafios e dificuldades enfrentadas

Durante o desenvolvimento do sistema, foram identificados desafios relacionados principalmente à integração entre diferentes tecnologias e à adaptação da solução ao ambiente hospitalar.

Um dos principais desafios envolveu a integração entre o banco de dados MySQL da aplicação e o banco de dados institucional do hospital, baseado em Firebird, exigindo ajustes nas consultas e no tratamento das informações.

Também foram encontradas inconsistências durante as etapas de desenvolvimento e testes, cuja resolução demandou análise do código e da estrutura do banco de dados. Essas dificuldades foram superadas por meio de testes sucessivos, ajustes no sistema e orientação técnica dos professores envolvidos no projeto.

IV. Conclusão

O presente trabalho apresentou o desenvolvimento e a implementação de um sistema web para gerenciamento de visitas hospitalares, elaborado a partir de uma demanda real do Hospital Antônio Moreira da Costa, em parceria com o eHealth Innovation Center. A solução proposta substituiu o controle manual de visitas por um processo informatizado, promovendo maior organização, segurança e rastreabilidade das informações.

A plataforma desenvolvida demonstrou-se eficaz na gestão do fluxo de visitantes e acompanhantes, possibilitando o registro estruturado de dados, o controle automático de horários de entrada e saída e a centralização das informações em um ambiente digital seguro. Essas funcionalidades contribuíram para a redução de erros de registro, otimização do atendimento na recepção e maior conformidade com as normas de segurança do paciente.

Do ponto de vista acadêmico e tecnológico, o projeto evidenciou a viabilidade da aplicação de sistemas de informação no contexto hospitalar, mesmo em ambientes com infraestrutura preexistente e bancos de dados legados. A integração com o banco institucional do hospital permitiu a obtenção automática dos dados dos pacientes, assegurando a consistência das informações e evitando retrabalho por parte dos funcionários da recepção.

Além de atender às necessidades específicas da instituição parceira, o sistema apresenta potencial de replicação em outros hospitais, mediante pequenas adaptações técnicas e operacionais. Dessa forma, o sistema desenvolvido demonstra o potencial da tecnologia da informação como ferramenta estratégica

para a digitalização e melhoria dos processos administrativos hospitalares.

V. Agradecimentos

Agradecemos ao Instituto Nacional de Telecomunicações (INATEL) por fornecer a infraestrutura e o suporte necessários para a realização deste projeto, proporcionando um ambiente propício para a nossa pesquisa e aprendizado. De mesma forma agradecemos a Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pela concessão das bolsas de pesquisa.

Referências

[1] Empresa Brasileira de Serviços Hospitalares (EBSERH). *Visita ao paciente*. 2025. URL: https://www.gov.br/ebserh/pt-br/hospitais-universitarios/regiao-sudeste/hc-ufu/saude/visita-ao-paciente.

[2] Donald Berwick. "What 'patient-centered' should mean". Em: *Health Affairs* (2009).

[3] BRASIL. Agência Nacional de Vigilância Sanitária. *Resolução da Diretoria Colegiada – RDC nº 36, de 25 de julho de 2013*. Institui ações para a segurança do paciente em serviços de saúde. Brasília, 2013. URL: https://www.gov.br/anvisa.

[4] BRASIL. Ministério da Saúde. *Política Nacional de Segurança do Paciente*. Brasília, 2014. URL: https://www.gov.br/saude.

[5] World Health Organization. *Patient Safety Solutions*. 2009.

[6] BRASIL. Empresa Brasileira de Serviços Hospitalares. *Manual de gestão hospitalar e segurança do paciente*. Brasília, 2018. URL: https://www.gov.br/ebserh.

[7] Kenneth C. Laudon e Jane P. Laudon. *Sistemas de informação gerenciais*. 11ª ed. São Paulo: Pearson, 2014.

[8] Alfred Winter, Reinhold Haux, Elske Ammenwerth e Birgit Brigl. *Health Information Systems: Architectures and Strategies*. London: Springer, 2011.

[9] Marina F. Giombelli. *Dados de saúde no Brasil: Panorama das Clínicas e Hospitais2025*. 2024. URL: https://pro.doctoralia.com.br/blog/clinicas/dados-de-saude-no-brasil-panorama-das-clinicas-e-hospitais.

[10] Observatório de Política e Gestão Hospitalar – Fundação Oswaldo Cruz (OPGH-FIOCRUZ). *Primeiros resultados da Pesquisa Nacional de Saúde 2019 relativos a internações*. 2021. URL: https://observatoriohospitalar.fiocruz.br/conteudo-interno/primeiros-resultados-da-pesquisa-nacional-de-saude-2019-relativos-internacoes.

[11] Clemens Scott Kruse, Ashleigh Beane e Information Technology Health. "The impact of health information technology on hospital performance". Em: *JMIR Medical Informatics* 6.3 (2018), e10247.

[12] Aurum Software Ltda. *Veja o que diz e como funciona a Lei do Acompanhante – Lei 14.737/2023*. 2024. URL: https://www.aurum.com.br/blog/lei-do-acompanhante/.

[13] Madison N. Ngafeeson. "Healthcare information technology adoption: Implications for patient care". Em: *International Journal of Healthcare Information Systems and Informatics* 5.2 (2010), pp. 1–17.

[14] AGSinn – Tecnologia Criativa. *Experiências interativas e soluções digitais baseadas em IoT e IA*. 2025. URL: https://agsinn.com.br/.

[15] Yan Bi, Wei Zhang e Hao Liu. "Design and Implementation of a Ward Visitor Management System". Em: *Journal of Healthcare Engineering* 2023 (2023), pp. 1–9.

Autores



Fernando Faluxis Siqueira Pereira graduando em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel). Possui formação técnica em Eletrônica, com ênfase em Automação, pela Escola Técnica de Eletrônica "Francisco Moreira da Costa"(ETE FMC). Atualmente, é voluntário de Iniciação Científica no eHealth Innovation Center.



Marcela de Cássia Ribeiro graduanda em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel). Possui formação técnica em Eletrônica, com ênfase em Automação, pela Escola Técnica de Eletrônica "Francisco Moreira da Costa"(ETE FMC). Atualmente, é bolsista de Iniciação Científica no eHealth Innovation Center.



Rita Elizabeth Santos de Almeida Graduada em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações, técnica em equipamentos médicos pela Escola Técnica em Eletrônica Francisco Moreira da Costa. Especialista em sistemas no Inatel, atuando como coordenadora de projetos no eHealth Innovation Center.



Elisa Rennó de Carneiro Dester Fisioterapeuta formada pela Universidade Gama Filho (RJ) em 1999; Mestre em Ciências da Motricidade Humana pela Universidade Castelo Branco (RJ) em 2003; Professora do curso de pós-graduação em Engenharia Clínica e Biomédica do Inatel; Professora do curso de Engenharia Biomédica do Inatel.

Desenvolvimento de um Sistema Web para Gestão de Espaço de Vivência

Augusto Farias de Lima, Afonso Farias de Lima, Kelven Bertanha do Nascimento, Stefany Kaori Omori,
Jéssica Antonio Delgado, Isaque Katahira, Fábio Nogueira de Queiroz
Faculdade de Tecnologia de Ourinhos - FATEC OURINHOS
farias.augusto.lima@gmail.com, afonso.f.limaa@gmail.com, kelvenbertanhadonascimento@gmail.com,
omoristefanykaori@gmail.com, jessica.delgado@fatecourinhos.edu.br, isaque.katahira@fatecourinhos.edu.br,
fabio.queiroz@fatecourinhos.edu.br

***Abstract*—This paper presents the development of a management system for Espaço de Vivência Primaveras, a community initiative aimed at people aged 50 and older. The project was conducted within the scope of university extension, integrating higher education with real social demand. The adopted methodology included requirements elicitation through meetings with managers, interface prototyping, and incremental development using web technologies (HTML, CSS, JavaScript, and PHP with the Laravel framework) integrated with a MySQL database. The system modeling was represented through use case and class diagrams, enabling the organization of functionalities and application structure. The implemented system includes user registration, activity management, attendance tracking, and payment control, contributing to the digitalization of processes previously performed manually. The results demonstrate the technical feasibility of the solution and its relevance as an extension initiative with organizational impact.**

***Index Terms*—Digital Inclusion, Software Engineering, University Extension, Web Development.**

***Resumo*—Este trabalho apresenta o desenvolvimento de um sistema de gerenciamento para o Espaço de Vivência Primaveras, iniciativa comunitária voltada a pessoas com 50 anos ou mais. O projeto foi conduzido no âmbito da extensão universitária, articulando ensino superior e demanda social real. A metodologia adotada incluiu levantamento de requisitos por meio de reuniões com os gestores, prototipagem das interfaces e desenvolvimento incremental utilizando tecnologias web (HTML, CSS, JavaScript e PHP com o framework Laravel) integradas a banco de dados MySQL. A modelagem do sistema foi representada por diagramas de casos de uso e de classes, permitindo organizar as funcionalidades e a estrutura da aplicação. O sistema implementado contempla cadastro de usuários, gerenciamento de vivências, registro de frequência e controle de pagamentos, contribuindo para a digitalização de processos anteriormente realizados de forma manual. Os resultados evidenciam a viabilidade técnica da solução e sua relevância como ação extensionista com impacto organizacional.**

***Palavras chave*—Desenvolvimento Web, Engenharia de Software, Extensão Universitária, Inclusão Digital.**

I. Introdução

A incorporação de tecnologias digitais em diferentes contextos organizacionais tem promovido transformações significativas na gestão de processos e na prestação de serviços. Em espaços comunitários, especialmente aqueles que atendem públicos com maior diversidade etária, a digitalização de rotinas administrativas pode contribuir para maior organização, transparência e eficiência operacional [1]. Contudo, a adoção de sistemas informatizados nesses ambientes exige atenção às especificidades dos usuários e às condições de uso das tecnologias.

No caso de pessoas com 50 anos ou mais, a literatura aponta que a intenção de uso de tecnologias digitais é influenciada por múltiplos fatores de natureza sociodemográfica, cognitiva e contextual, o que reforça a necessidade de soluções acessíveis e alinhadas às características desse público [2]. Assim, o desenvolvimento de sistemas voltados a esse perfil deve considerar não apenas aspectos técnicos, mas também princípios de inclusão e usabilidade.

No âmbito do ensino superior, projetos dessa natureza configuram-se como ações extensionistas ao articular a formação acadêmica à resolução de demandas reais da comunidade, conforme orientam as Diretrizes para a Extensão na Educação Superior [3]. Essa aproximação entre universidade e sociedade favorece tanto o aprendizado prático dos estudantes quanto o fortalecimento institucional das iniciativas comunitárias atendidas.

Diante desse contexto, o presente trabalho tem como objetivo apresentar o desenvolvimento de um sistema de gerenciamento para o Espaço de Vivência Primaveras, contemplando sua modelagem funcional e estrutural, implementação web e resultados obtidos a partir da aplicação da solução proposta.

II. Fundamentação Teórica

O desenvolvimento de soluções tecnológicas voltadas a espaços comunitários exige fundamentação que articule educação tecnológica, impacto social e acessibilidade digital. Projetos dessa natureza não se restringem à implementação técnica de sistemas, mas envolvem a compreensão do papel do ensino superior na promoção de transformações sociais e no fortalecimento de iniciativas comunitárias.

Nesse contexto, torna-se relevante discutir a contribuição da formação baseada em projetos, a utilização de sistemas de informação como instrumentos de gestão e organização, bem como os princípios de acessibilidade e inclusão digital que orientam o desenvolvimento de interfaces voltadas a públicos diversos. Assim, este referencial teórico apresenta os fundamentos que sustentam a proposta do sistema para o Espaço de Vivência, situando-o no âmbito da responsabilidade social da educação tecnológica e da aplicação prática do conhecimento em benefício da comunidade.

A. Educação Tecnológica e Impacto Social

As Diretrizes para a Extensão na Educação Superior estabelecem que as ações extensionistas devem envolver a comunidade externa e estar articuladas ao processo formativo discente [3]. Tal orientação reforça a compreensão de que projetos tecnológicos aplicados a contextos comunitários ultrapassam o exercício técnico, constituindo-se como estratégias de formação acadêmica com impacto social.

Nesse sentido, metodologias que aproximam os estudantes de problemas reais favorecem o desenvolvimento de competências técnicas e socioemocionais, promovendo uma formação alinhada às demandas contemporâneas da sociedade. As metodologias ativas, especialmente a Aprendizagem Baseada em Projetos *(Project-Based Learning* – PBL), têm sido apontadas como estratégias pedagógicas capazes de aproximar o estudante de situações reais, promovendo a integração entre teoria e prática [4].

A integração entre ensino e comunidade reforça o papel social das instituições de educação superior, ampliando o acesso ao conhecimento tecnológico e contribuindo para o desenvolvimento local. Nesse sentido, projetos tecnológicos aplicados a espaços comunitários podem ser compreendidos como mecanismos de fortalecimento institucional e organizacional, ao promoverem maior sistematização de processos, eficiência administrativa e ampliação do alcance das ações desenvolvidas nesses contextos. Tal perspectiva dialoga com os Objetivos de Desenvolvimento Sustentável (ODS) estabelecidos pela Organização das Nações Unidas, especialmente aqueles relacionados à educação de qualidade (ODS 4), à inovação e infraestrutura (ODS 9), à redução das desigualdades (ODS 10) e à promoção de comunidades sustentáveis (ODS 11) [5], ao evidenciar a tecnologia como instrumento de inclusão e desenvolvimento social.

Assim, o desenvolvimento de sistemas voltados a demandas comunitárias ultrapassa a dimensão operacional da programação, configurando-se como prática formativa alinhada à responsabilidade social da educação superior e à promoção de inovação com impacto social.

B. Tecnologia como Ferramenta de Gestão Comunitária

A transformação digital tem redefinido a forma como organizações estruturam seus processos, tomam decisões e se relacionam com seus públicos. Nesse contexto, os sistemas de informação assumem papel estratégico ao integrar dados, automatizar rotinas e fornecer suporte à gestão organizacional, contribuindo para maior eficiência, controle e transparência nas operações. De acordo com [6], os sistemas de informação não apenas apoiam atividades administrativas, mas também influenciam a capacidade das organizações de inovar e ampliar seu alcance. A transformação digital pode ser compreendida como um processo de mudança organizacional mediado pela incorporação estratégica de tecnologias digitais, impactando estruturas, processos e formas de geração de valor [1]. Assim, a adoção de sistemas de informação em iniciativas comunitárias não deve ser compreendida apenas como modernização tecnológica, mas como estratégia de fortalecimento organizacional. Ao estruturar dados, apoiar decisões e otimizar fluxos de trabalho, tais sistemas ampliam a sustentabilidade institucional e potencializam o impacto social das ações desenvolvidas, consolidando a tecnologia como instrumento de apoio à gestão e à transformação social.

C. Inclusão Social, Acessibilidade e Diversidade Etária

A inclusão social constitui elemento central nas discussões contemporâneas sobre desenvolvimento tecnológico, especialmente em contextos comunitários nos quais a diversidade etária e sociocultural demanda soluções sensíveis às diferentes realidades dos usuários. A incorporação de tecnologias digitais nesses ambientes deve ser orientada não apenas pela eficiência operacional, mas também pelo compromisso com a ampliação do acesso, da participação e da autonomia dos indivíduos. Esse cenário evidencia a importância de desenvolver soluções tecnológicas alinhadas às especificidades das pessoas maduras, considerando os fatores que influenciam sua intenção de uso e a interação com sistemas digitais [2].

Nesse contexto, o desenvolvimento do sistema para o Espaço de Vivência considerou as especificidades do público atendido, composto por pessoas com 50 anos ou mais, incorporando cuidados relacionados à organização das informações, à clareza das interfaces e à simplicidade na execução das funcionalidades. Tal abordagem evidencia a preocupação em alinhar requisitos técnicos às características dos usuários, reforçando o compromisso com a inclusão social e com a construção de uma solução tecnológica contextualizada.

III. Metodologia

O desenvolvimento do sistema para o Espaço de Vivência adotou uma abordagem incremental, permitindo a implementação progressiva das funcionalidades e ajustes ao longo do processo.

A fase inicial foi dedicada ao levantamento de requisitos. Foram realizadas três reuniões com o proprietário do Espaço de Vivência e seu sócio, com o objetivo de compreender os processos operacionais existentes, as funcionalidades desejadas e as expectativas em relação ao sistema.

Com base nos requisitos levantados, foram desenvolvidos protótipos de interface utilizando a ferramenta Figma, permitindo a modelagem das telas e dos fluxos de navegação.

A implementação foi realizada com as tecnologias HTML, CSS e JavaScript para o desenvolvimento da interface e das interações do usuário. No backend, utilizou-se PHP com o framework Laravel para o processamento das requisições, organização da aplicação segundo o padrão Model-View-Controller (MVC) e integração com o banco de dados MySQL, responsável pelo armazenamento das informações referentes a usuários, atividades, pagamentos e controle de frequência.

Durante o desenvolvimento, foram considerados aspectos relacionados à segurança da informação e à conformidade com a Lei Geral de Proteção de Dados (Lei nº 13.709/2018), especialmente no que se refere ao tratamento e armazenamento de dados pessoais.

A. Diagrama de Casos de Uso

A partir dos requisitos levantados e das funcionalidades definidas nas reuniões com os gestores, procedeu-se à modelagem funcional do sistema, representada por meio do diagrama

de casos de uso apresentado na Figura 1. Esse diagrama sintetiza as interações entre os diferentes perfis de usuários e as principais funcionalidades previstas para o sistema.

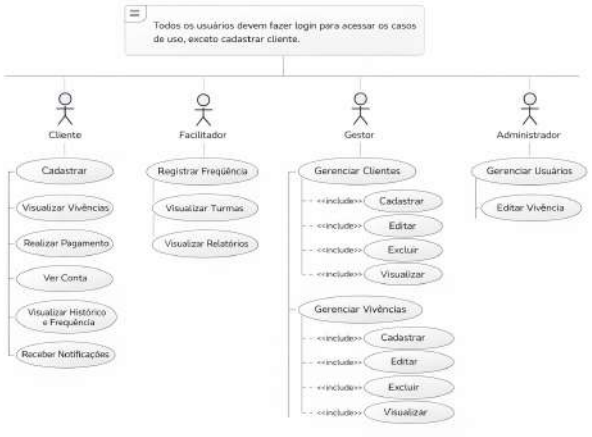


Fig. 1. Diagrama de Casos de Uso do Sistema de Gerenciamento do Espaço de Vivência.

Com base nessa modelagem funcional, foi elaborada a modelagem estrutural do sistema por meio de um diagrama de classes, detalhado na seção seguinte, permitindo a organização das entidades, atributos e relacionamentos que compõem a solução proposta.

B. Diagrama de Entidade e Relacionamento

A modelagem estrutural do sistema é apresentada na Figura 2, por meio do diagrama de Entidade e Relacionamento. O modelo organiza as principais entidades da aplicação e seus relacionamentos, destacando a classe *Pessoa* como superclasse de *Cliente* e *Facilitador*, bem como a utilização da classe *Inscrição* para intermediar a participação dos clientes nas vivências.

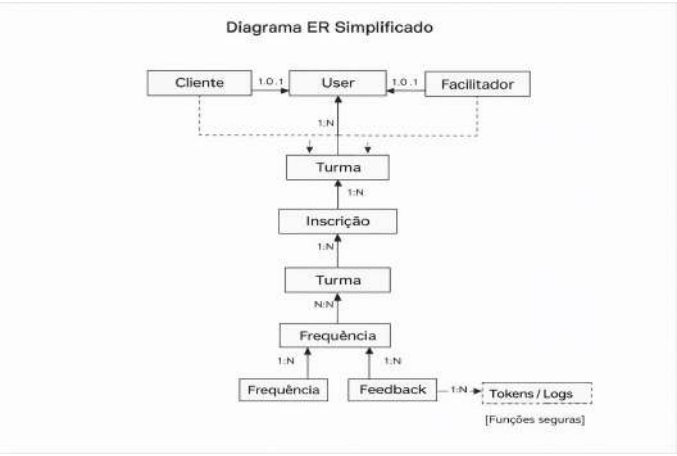


Fig. 2. Diagrama de Entidade e Relacionamento do Sistema de Gerenciamento do Espaço de Vivência.

Essa modelagem contribui para a organização da aplicação, favorecendo clareza estrutural e suporte à implementação das funcionalidades definidas na etapa de requisitos.

IV. Resultados

O sistema desenvolvido para o Espaço de Vivência Primaveras resultou em uma solução digital voltada à organização das atividades e ao gerenciamento dos usuários. A aplicação contempla funcionalidades como cadastro de clientes, gerenciamento de vivências, registro de frequência, controle de pagamentos e envio de notificações.

A interface foi projetada com foco na simplicidade e clareza das informações, considerando o perfil do público atendido pelo espaço e os protótipos elaborados foram validados junto aos gestores.

A Figura 3 apresenta a tela inicial de login da aplicação.



Fig. 3. Interface inicial do sistema – tela de login.

A navegação foi estruturada de forma objetiva, com menu contendo acesso às principais seções do sistema.

As Figuras 4 e 5 apresentam telas representativas das principais funcionalidades implementadas no sistema. A Figura 4 ilustra o módulo de gerenciamento de vivências e usuários, enquanto a Figura 5 apresenta o módulo operacional, relacionado ao registro de frequência e ao controle de pagamentos.

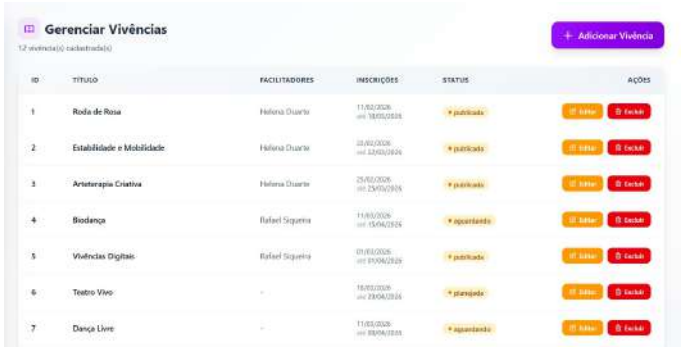


Fig. 4. Tela de gerenciamento de vivências e usuários.

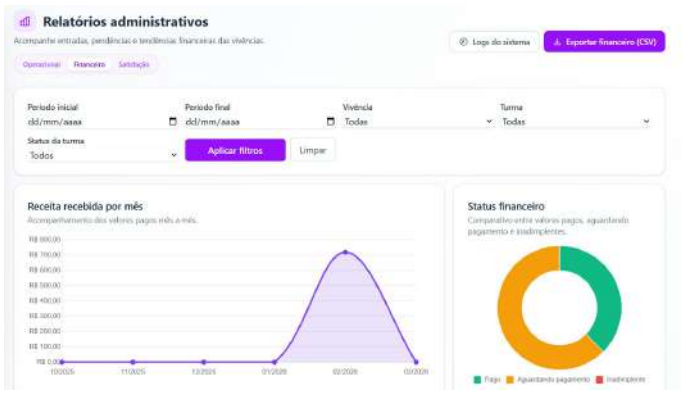


Fig. 5. Tela de controle de pagamentos.

O sistema implementado permite a digitalização dos processos anteriormente realizados de forma manual, promovendo maior organização administrativa e melhor acompanhamento das atividades desenvolvidas no Espaço de Vivência.

V. Conclusão

O presente trabalho apresentou o desenvolvimento de um sistema de gerenciamento para o Espaço de Vivência Primaveras, com foco na organização das atividades e na adequação às necessidades de um público composto majoritariamente por pessoas maduras. A partir do levantamento de requisitos realizado junto aos gestores, foi possível estruturar uma solução digital que contempla funcionalidades como cadastro de usuários, gerenciamento de vivências, controle de frequência e registro de pagamentos.

Além dos resultados técnicos, o projeto também se configura como uma ação de extensão no âmbito do ensino superior, ao articular o conhecimento acadêmico ao atendimento de uma demanda real da comunidade. A aproximação entre estudantes e o Espaço de Vivência possibilitou a aplicação prática de conceitos de engenharia de software, ao mesmo tempo em que contribuiu para o fortalecimento institucional da iniciativa comunitária.

Como perspectivas futuras, destaca-se a possibilidade de ampliação das funcionalidades e a realização de avaliações formais de usabilidade junto aos usuários, visando aprimorar continuamente a solução desenvolvida.

VI. Agradecimentos

Os autores agradecem à Faculdade de Tecnologia de Ourinhos (FATEC Ourinhos) pelo suporte institucional, pela orientação docente e pelos recursos disponibilizados que possibilitaram a realização deste projeto.

Referências

[1] Peter C. Verhoef, Thijs Broekhuizen, Yakov Bart, Abhi Bhattacharya, John Q. Dong, Nicolai Fabian e Michael Haenlein. “Digital transformation: A multidisciplinary reflection and research agenda”. Em: *Journal of Business Research* 122 (2021), pp. 889–901. ISSN: 0148-2963. DOI: 10.1016/j.jbusres.2019.09.022.

[2] T. Schroeder, L. Dodds, A. Georgiou, H. Gewald e J. Siette. “Older Adults and New Technology: Mapping Review of the Factors Associated With Older Adults’ Intention to Adopt Digital Technologies”. Em: *JMIR Aging* 6 (2023), e44564. DOI: 10.2196/44564. URL: https://aging.jmir.org/2023/1/e44564.

[3] Brasil. Ministério da Educação. Conselho Nacional de Educação. *Resolução CNE/CES nº 7, de 18 de dezembro de 2018*. Diário Oficial da União, Brasília, DF, 19 dez. 2018. Estabelece as Diretrizes para a Extensão na Educação Superior Brasileira. 2018.

[4] José Moran. “Metodologias ativas para uma aprendizagem mais profunda”. Em: *Metodologias ativas para uma educação inovadora: uma abordagem teórico-prática*. Ed. por Lilian Bacich e José Moran. Porto Alegre: Penso, 2018, pp. 37–76.

[5] Organização das Nações Unidas. *Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável*. Agenda 2030. 2015.

[6] Kenneth C. Laudon e Jane P. Laudon. *Management Information Systems: Managing the Digital Firm*. 16ª ed. Harlow: Pearson, 2020.

Autores



Augusto Farias de Lima Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Afonso Farias de Lima Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Kelven Bertanha do Nascimento Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software. É bolsista de Iniciação Científica pelo CNPq, desenvolvendo atividades acadêmicas voltadas à pesquisa e à inovação tecnológica.



Stefany Kaori Omori Técnica em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Jéssica Antonio Delgado Possui graduação em Matemática pelo Centro Universitário Sagrado Coração (2013), mestrado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2016) e doutorado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2021). Tem experiência na área de

Matemática, com ênfase em Engenharia Elétrica, atuando principalmente nos seguintes temas: método de barreira, fluxo de potência ótimo, método de barreira modificada, método de penalidade e otimização não linear. Atualmente é Professora na Faculdade de Tecnologia em Ourinhos (FATEC Ourinhos) e professora no UNIFIO.



Isaque Katahira Docente do Centro Estadual de Educação Tecnológica Paula Souza, atuando na Faculdade de Tecnologia de Ourinhos e na Escola Técnica Estadual Professor Mário Antônio Verza, tendo exercido a função de Diretor Substituto na Etec Professor Mário Antônio Verza (novembro de 2023 a abril de 2024). Ingressou na Etec em 2007 e na Faculdade de Tecnologia em 2017.Possui Licenciatura em Sistemas e Tecnologias da Informação (2006), Bacharelado em Desenvolvimento de Sistemas (2007) e Tecnólogo em Projeto de Sistemas de Informação (2008) pela Faculdade de Tecnologia de Ourinhos. Concluiu o Mestrado em Bioinformática pela Universidade Tecnológica Federal do Paraná (2018) e o Doutorado em Ciência da Informação pela Universidade Estadual Paulista UNESP (2022). Participa como revisor da Revista Eletrônica e-F@tec e é membro do Grupo de Pesquisa Representação Temática da Informação.



Fábio Nogueira de Queiroz Mestre em Ensino pela Universidade Estadual do Norte do Paraná, Programa de Pós-Graduação em Ensino (PPGEN/UENP). Possui graduação em Tecnologia em Processamento de Dados pela Faculdade de Tecnologia de Ourinhos - FATEC (2002). É licenciando em Matemática pela Universidade Estácio de Sá (2022), licenciado em Informática pela Faculdade de Tecnologia de Sorocaba - FATEC (2008). Licenciatura Plena em Pedagogia pela FALC (2015). Pós Graduação Latu Sensu em MBA Informática Tecnologia Internet (2004) e Pós Graduação Latu Sensu em Docência do Ensino Superior (2011). Atualmente é Professor II-E, na Fatec Ourinho e Coordenador da Etec Jacinto Ferreira de Sá em Ourinhos, no curso de Técnico em Desenvolvimento de Sistemas. Professor, nos cursos de Técnico em Desenvolvimento de Sistemas, Técnico em Informática para Internet e Técnico em Análise de Sistemas. Possui experiência na área de Informática, Técnica e Didática Pedagógica.

Desenvolvimento de uma Plataforma Digital para Apoio à Inserção Profissional

Daniel Couto Resende, Douglas Eduardo Gabriel, Luis Miguel Nunes da Silva, Geovane Feliciano
Jéssica Antonio Delgado, Isaque Katahira, Fábio Nogueira de Queiroz
Faculdade de Tecnologia de Ourinhos - FATEC OURINHOS
resendedanielcouto@gmail.com, douglaseduardogabriel19@gmail.com, lmnunes06@gmail.com,
geovane.feliciano@fatec.sp.gov.br, jessica.delgado@fatecourinhos.edu.br, isaque.katahira@fatecourinhos.edu.br,
fabio.queiroz@fatecourinhos.edu.br.

***Abstract*—Technological transformations in the labor market have increased the demand for specific competencies and intensified the challenges related to the connection of professionals and companies. In this context, digital employability platforms have become strategic tools for mediating recruitment and selection processes. This work presents the development of ADAPT, a web platform designed to connect technology professionals and recruiting companies. The system was conceived with a modular architecture and implemented based on the Waterfall development model, encompassing requirements elicitation, modeling, prototyping, implementation, testing, and deployment. The application integrates modern web development technologies, ensuring clear layer separation and scalability. The results demonstrate consistent performance of the main workflows of the system, including authentication, job posting, and application processes. The proposed solution is concluded to contribute to improving the organization of professional profile and supporting digital recruitment processes.**

***Index Terms*—Digital Employability, Online Recruitment, Waterfall Model, Web Architecture.**

***Resumo*—As transformações tecnológicas no mercado de trabalho têm ampliado a demanda por competências específicas e intensificado os desafios relacionados à conexão entre profissionais e empresas. Nesse contexto, plataformas digitais de empregabilidade tornam-se ferramentas estratégicas para intermediar processos de recrutamento e seleção. Este trabalho apresenta o desenvolvimento do ADAPT, uma plataforma web voltada à aproximação entre profissionais de tecnologia e empresas recrutadoras. O sistema foi concebido com arquitetura modular e implementado com base no modelo de Desenvolvimento em Cascata, contemplando levantamento de requisitos, modelagem, prototipação, implementação, testes e implantação. A aplicação integra tecnologias modernas de desenvolvimento web, garantindo separação de camadas e escalabilidade. Os resultados demonstram funcionamento consistente dos principais fluxos do sistema, incluindo autenticação, publicação de vagas e candidaturas. Conclui-se que a solução proposta contribui para otimizar a organização de perfis profissionais e apoiar processos de recrutamento digital.**

***Palavras chave*—Arquitetura Web, Empregabilidade Digital, Modelo Cascata, Recrutamento Online.**

I. Introdução

As transformações tecnológicas recentes têm impactado significativamente o mercado de trabalho, ampliando a demanda por novas competências e exigindo atualização contínua dos profissionais [1]. Nesse contexto, torna-se cada vez mais desafiador para empresas identificar candidatos alinhados às exigências técnicas e comportamentais do cenário profissional contemporâneo.

As Plataformas Digitais de Empregabilidade (PDEs) têm desempenhado papel relevante na intermediação entre profissionais e oportunidades, ampliando as práticas de recrutamento online e incorporando mecanismos digitais de análise e seleção de perfis [2]. Além disso, a incorporação de recursos baseados em dados e inteligência artificial tem contribuído para a otimização dos processos de identificação e retenção de talentos [3].

Diferentemente de plataformas tradicionais de recrutamento, o ADAPT foi concebido com características inspiradas em redes sociais profissionais, buscando tornar a interação entre candidatos e empresas mais dinâmica e intuitiva. A plataforma permite que profissionais criem perfis detalhados, apresentem projetos desenvolvidos e compartilhem experiências relevantes, ampliando a visibilidade de suas competências. O sistema também possibilita a comunicação direta entre candidatos e empresas por meio de mensagens, favorecendo a aproximação entre as partes durante o processo de recrutamento. Além disso, o ADAPT prevê mecanismos de identificação e validação de perfis institucionais de empresas, reduzindo o risco de publicação de vagas falsas e contribuindo para maior confiabilidade no ambiente digital

Diante desse cenário, este trabalho tem como objetivo desenvolver e apresentar o ADAPT, uma plataforma web voltada à conexão entre profissionais da área de tecnologia e empresas recrutadoras.

II. Fundamentação Teórica

As transformações tecnológicas têm impactado significativamente o mercado de trabalho, ampliando os desafios relacionados à inserção e recolocação profissional. Nesse contexto, plataformas digitais de empregabilidade surgem como ferramentas capazes de aproximar profissionais e oportunidades, auxiliando na organização de perfis, identificação de competências e acesso a vagas. Esta seção apresenta os principais conceitos teóricos que fundamentam o desenvolvimento da plataforma proposta.

A. Transformações no Mercado de Trabalho e Escassez de Competências

O avanço da automação, das tecnologias digitais e da inteligência artificial tem provocado o surgimento de novas funções e a obsolescência de outras, ampliando a necessidade de

atualização contínua dos profissionais. Uma parcela significativa das competências exigidas no mercado de trabalho sofrerá mudanças nos próximos anos, o que evidencia a necessidade de processos contínuos de requalificação e aprimoramento [1]. O termo *skills* refere-se ao conjunto de capacidades que permitem ao indivíduo mobilizar conhecimentos, atitudes e valores para realizar tarefas e resolver problemas em contextos profissionais específicos. De acordo com a [4], as habilidades envolvem a aplicação prática do conhecimento de forma responsável e eficaz, sendo tradicionalmente classificadas em cognitivas, técnicas e socioemocionais. As *hard skills* estão associadas a conhecimentos técnicos específicos relacionados à execução de tarefas e, em geral, são adquiridas por meio da educação formal em universidades e instituições técnicas. Em contrapartida, as *soft skills* — como comunicação, liderança, criatividade e colaboração apresentam-se como preditoras mais consistentes do sucesso profissional [5]. Apesar da distinção conceitual entre hard skills e soft skills, a literatura aponta que tais competências frequentemente se intersectam e atuam de forma complementar nas atividades profissionais [6]. Sob a perspectiva do recrutamento, torna-se fundamental que os empregadores avaliem os candidatos não apenas com base em suas qualificações técnicas, mas também considerando atributos pessoais, como habilidades de comunicação, criatividade e capacidade de resolução de problemas, uma vez que tais competências contribuem de forma significativa para o desempenho profissional em contextos organizacionais dinâmicos [7]. Nesse contexto, cresce a relevância de soluções digitais capazes de apoiar processos de recrutamento e inserção profissional, conectando candidatos e empregadores por meio de plataformas digitais.

B. Plataformas Digitais de Empregabilidade

As Plataformas Digitais de Empregabilidade (PDEs) têm provocado mudanças dinâmicas nos mercados de trabalho tradicionais ao ampliar as práticas de recrutamento online [2]. Um exemplo relevante é o LinkedIn, lançado em 2003 e atualmente consolidado como uma das principais redes profissionais digitais, reunindo mais de 850 milhões de usuários globalmente [8]. Os membros preenchem um perfil profissional online e compartilham suas experiências, conquistas, empregadores, projetos e contatos de sua rede profissional. De acordo com estatísticas fornecidas pelo LinkedIn, uma vaga de emprego é visualizada, em média, aproximadamente 500 vezes e recebe mais de 50 candidaturas [3].

As PDEs oferecem vantagens em relação aos métodos tradicionais de recrutamento, como indicações, candidaturas presenciais e processos realizados em campus universitários, ao ampliarem o alcance e a eficiência das conexões profissionais. Os usuários podem consultar vagas e obter informações sobre as funções e localidades oferecidas, níveis salariais, dados de contato, desempenho e destaques da empresa.

Além de ampliarem o acesso à informação e ao networking profissional, as PDEs evoluíram para sistemas digitais capazes de apoiar processos de recrutamento orientados por dados. A incorporação de inteligência artificial e análise automatizada permite otimizar a identificação, seleção e retenção de talentos em escala global [3].

C. Tecnologias Digitais no Apoio à Inserção Profissional

O avanço das tecnologias digitais tem ampliado as possibilidades de apoio à inserção profissional ao incorporar ferramentas baseadas em dados, inteligência artificial e automação nos processos de busca e seleção de candidatos. Essas tecnologias permitem maior eficiência na análise de perfis profissionais, recomendação de vagas e identificação de competências, contribuindo para processos de recrutamento mais ágeis e escaláveis.

Diversas plataformas digitais atuam atualmente no apoio ao recrutamento e à prestação de serviços profissionais online. A Tabela I apresenta exemplos de plataformas amplamente utilizadas nesse contexto.

TABELA I
PLATAFORMAS DE RECRUTAMENTO E TRABALHO FREELANCE

height**Plataforma**	**Característica**
LinkedIn	Rede social profissional e vagas
Indeed	Agregador de vagas
XING	Rede profissional na Europa
Kalibrr	Recrutamento com testes comportamentais
Upwork	Contratação freelance por projeto
Fiverr	Marketplace de serviços digitais
Freelancer.com	Sistema competitivo de propostas
PeoplePerHour	Projetos e serviços por hora
FlexJobs	Vagas remotas verificadas
We Work Remotely	Plataforma de vagas remotas
Remote OK	Foco em startups e tecnologia
Toptal	Rede seletiva de talentos
Hubstaff Talent	Diretório gratuito de profissionais
Workana	Plataforma popular na América Latina
Guru	Diferentes modelos de contrato

A análise dessas plataformas evidencia diferentes modelos de interação profissional e recrutamento digital, servindo como referência para o desenvolvimento da plataforma ADAPT proposta neste trabalho.

III. METODOLOGIA

O desenvolvimento do ADAPT foi conduzido com base no modelo de Desenvolvimento em Cascata *(Waterfall)*. Esse modelo caracteriza-se por uma abordagem sequencial, na qual cada etapa do desenvolvimento deve ser concluída antes do início da seguinte, priorizando planejamento estruturado e documentação das atividades realizadas [9].

O código-fonte da plataforma ADAPT foi disponibilizado publicamente em um repositório GitHub, permitindo acesso à implementação do sistema e à estrutura do projeto-Disponível em: https://github.com/Danilo1500/Projeto-Adapt.

A escolha dessa metodologia deve-se à necessidade de organização estruturada das etapas do projeto, documentação dos artefatos produzidos e acompanhamento sistemático do progresso do desenvolvimento. O processo foi dividido em fases distintas, compreendendo levantamento de requisitos, modelagem, prototipação e implementação, executadas de forma sequencial e interdependente.

Inicialmente, realizou-se o levantamento de requisitos funcionais e não funcionais por meio de entrevistas exploratórias com profissionais da área de tecnologia e análise de plataformas similares disponíveis no mercado. A partir dessas

informações, foram definidos os principais requisitos do sistema, incluindo mecanismos de busca de profissionais, gestão de vagas, visualização de projetos e autenticação de usuários.

Na etapa de modelagem foram elaborados diagramas de caso de uso, classes e sequência, permitindo a definição das interações entre usuários e sistema, bem como a estrutura lógica dos dados.

A Figura 1 apresenta o diagrama de caso de uso do sistema, evidenciando as principais interações entre os usuários e as funcionalidades da plataforma.

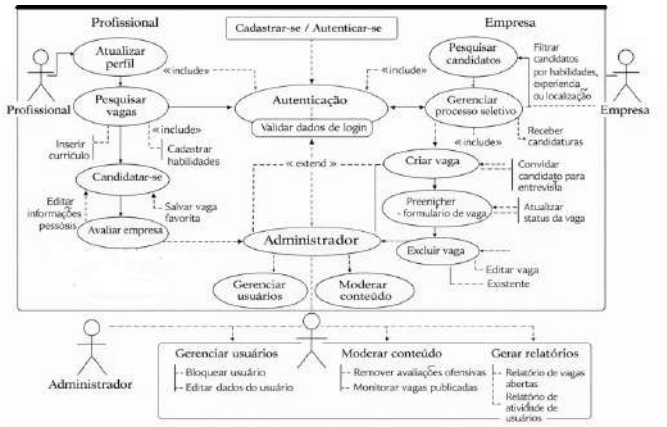


Fig. 1. Diagrama de caso de uso da plataforma ADAPT

Em seguida, realizou-se a prototipação das interfaces utilizando a ferramenta Figma, com foco na organização da navegação e na experiência do usuário. Posteriormente, procedeu-se à implementação do sistema utilizando React.js no desenvolvimento do frontend, Node.js com Express na camada de backend e MongoDB como banco de dados. O sistema também integra o serviço de autenticação Clerk e foi implantado em ambiente de hospedagem na plataforma Vercel.

TABELA II
TECNOLOGIAS UTILIZADAS NO DESENVOLVIMENTO DO ADAPT

Camada	Tecnologia	Finalidade
Frontend	React.js	Interface do usuário
Backend	Node.js + Express	APIs e lógica de negócio
Banco de Dados	MongoDB	Persistência de dados
Autenticação	Clerk	Controle de acesso
Hospedagem	Vercel	Implantação da aplicação

A adoção dessa stack tecnológica permitiu a construção de uma aplicação modular e escalável, alinhada às práticas atuais de desenvolvimento web.

Após a implementação, foram realizados testes unitários e testes end-to-end com o objetivo de verificar o funcionamento das principais funcionalidades do sistema e a consistência dos dados inseridos pelos usuários. Também foram realizados testes de validação de formulários e verificação dos fluxos de autenticação e publicação de vagas.

Concluída a etapa de testes, o sistema foi implantado em ambiente de produção na plataforma Vercel, com integração ao MongoDB Atlas para armazenamento e gerenciamento dos dados.

IV. RESULTADOS

O ADAPT foi desenvolvido como uma plataforma web voltada à conexão entre profissionais da área de tecnologia e empresas recrutadoras. O sistema foi estruturado de forma modular, organizado em camadas de apresentação, aplicação e persistência de dados, seguindo princípios comuns de arquitetura em aplicações web.

A plataforma contempla funcionalidades específicas para cada perfil de usuário.

Para profissionais:

- Criação e edição de perfil profissional com apresentação de competências, experiências e projetos;
- Busca por vagas com aplicação de filtros;
- Candidatura simplificada às oportunidades disponíveis;
- Interação direta com empresas por meio da plataforma.

Para empresas:

- Cadastro de perfil institucional;
- Publicação e gerenciamento de vagas;
- Busca por profissionais utilizando filtros por habilidades e competências;
- Contato direto com candidatos.

Além das funcionalidades tradicionais de recrutamento, a plataforma foi concebida com características inspiradas em redes sociais profissionais, permitindo maior visibilidade das experiências e projetos dos usuários e favorecendo a interação entre profissionais e empresas em um ambiente digital mais dinâmico.

Diferentemente das plataformas consolidadas listadas na Tabela I, que muitas vezes operam sob arquiteturas monolíticas ou agregadores de dados genéricos , o ADAPT diferencia-se tecnicamente pela adoção de uma stack modular e reativa baseada em React.js e Node.js. Enquanto ferramentas como Indeed e Workana focam primordialmente na transação de vagas e propostas, a arquitetura do ADAPT foi concebida para priorizar a exposição de portfólios técnicos e projetos, utilizando a flexibilidade do banco de dados MongoDB para gerir perfis que não se limitam a esquemas rígidos de dados. Além disso, a integração do serviço Clerk isola a lógica de autenticação da persistência de negócio, conferindo um nível de segurança e escalabilidade superior a diretórios profissionais simples, permitindo uma interação mais dinâmica e confiável entre candidatos e empresas no setor de tecnologia.

A Figura 2 apresenta a interface inicial de autenticação da plataforma, responsável pelo cadastro e acesso dos usuários. Já a Figura 3 ilustra o ambiente principal do sistema, no qual estão disponíveis os recursos de busca, gerenciamento de vagas e interação entre profissionais e empresas.

As interfaces apresentadas evidenciam a organização estrutural e a usabilidade da plataforma, demonstrando a integração entre autenticação, navegação e funcionalidades centrais.

V. CONCLUSÃO

O desenvolvimento do ADAPT demonstrou a viabilidade de uma plataforma web voltada à conexão entre profissionais da área de tecnologia e empresas recrutadoras. O sistema foi concebido com arquitetura modular e implementado utilizando tecnologias amplamente adotadas no desenvolvimento web, permitindo a construção de uma aplicação estruturada, escalável e alinhada às práticas atuais do mercado.

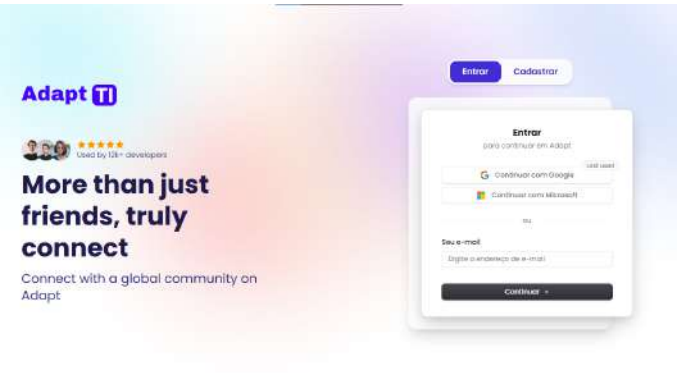


Fig. 2. Tela de Login e Cadastro da plataforma ADAPT.

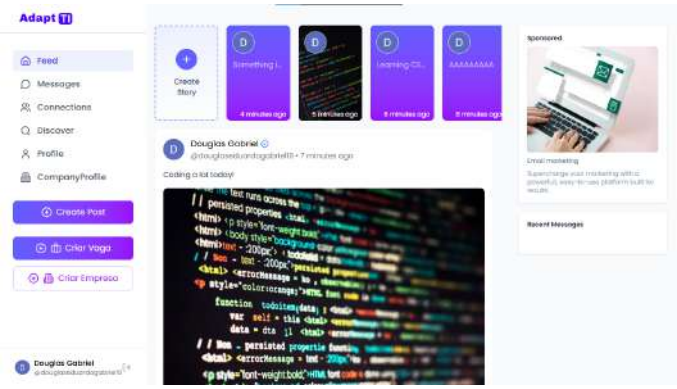


Fig. 3. Tela principal da plataforma com funcionalidades disponíveis aos usuários.

Os resultados obtidos reiteram que a aplicação não apenas atende aos requisitos funcionais de autenticação, publicação e gerenciamento de vagas , mas propõe uma evolução na experiência do usuário ao integrar funcionalidades típicas de redes sociais profissionais. A escolha de uma arquitetura baseada em tecnologias modernas e escaláveis permitiu superar limitações comuns de plataformas genéricas, oferecendo maior visibilidade a competências técnicas e projetos práticos. Conclui-se que o ADAPT cumpre seu objetivo de otimizar a organização de perfis e apoiar o recrutamento digital de forma dinâmica.

Como trabalhos futuros, a inclusão de mecanismos de recomendação automática e testes de usabilidade com usuários reais consolidará a ferramenta como uma solução robusta para os desafios contemporâneos de inserção profissional.

VI. Agradecimentos

Os autores agradecem à Faculdade de Tecnologia de Ourinhos (FATEC Ourinhos) pelo suporte institucional, pela orientação docente e pelos recursos disponibilizados que possibilitaram a realização deste projeto.

Referências

[1] World Economic Forum. *The Future of Jobs Report 2020*. Geneva: World Economic Forum, 2020.

[2] Namita Ruparel, Seema Bhardwaj, Himanshu Seth e Rajneesh Choubisa. “Systematic literature review of professional social media platforms: Development of a behavior adoption career development framework”. Em: *Journal of Business Research* 156 (2023), p. 113482. DOI: 10.1016/j.jbusres.2022.113482. URL: https://www.sciencedirect.com/science/article/pii/S014829632200947X.

[3] Oihab Allal-Chérif, Alba Yela Aránega e Rafael Castaño Sánchez. “Intelligent recruitment: How to identify, select, and retain talents from around the world using artificial intelligence”. Em: *Technological Forecasting and Social Change* 169 (2021), p. 120822. ISSN: 0040-1625. DOI: 10.1016/j.techfore.2021.120822. URL: https://www.sciencedirect.com/science/article/pii/S0040162521002547.

[4] Organisation for Economic Co-operation and Development. *OECD Learning Compass 2030: Skills for 2030*. URL: https://www.oecd.org/education/2030-project/learning/skills/ (acesso em 21/02/2026).

[5] Charlie Hargood e Stephen Peckham. *Soft Skills for the Digital Age*. London: Routledge, 2017.

[6] J. Lamri e T. Lubart. “Reconciling Hard Skills and Soft Skills in a Common Framework: The Generic Skills Component Approach”. Em: *Journal of Intelligence* 11.6 (2023), p. 107. ISSN: 2079-3200. DOI: https://doi.org/10.3390/jintelligence11060107. URL: https://www.mdpi.com/2079-3200/11/6/107.

[7] W. Lyu e J. Liu. “Soft skills, hard skills: What matters most? Evidence from job postings”. Em: *Applied Energy* 300 (2021), p. 117307. ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2021.117307. URL: https://www.sciencedirect.com/science/article/pii/S0306261921008224.

[8] LinkedIn Corporation. *Sobre o LinkedIn*. URL: https://about.linkedin.com/pt-br (acesso em 22/02/2026).

[9] Raul Wazlawick. *Engenharia de software: conceitos e práticas*. Elsevier Editora Ltda., 2019.

Autores



Daniel Couto Resende é técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá. Atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos, onde vem aprofundando seus conhecimentos nas áreas de programação e engenharia de software. Possui interesse no desenvolvimento de sistemas e na aplicação de boas práticas de programação na construção de aplicações web.



Geovane Feliciano é técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá. Atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos, onde vem aprofundando seus conhecimentos nas áreas de desenvolvimento web, programação e engenharia de software. Possui interesse na criação de aplicações web e na implementação de soluções tecnológicas voltadas ao desenvolvimento de sistemas e aplicações digitais.



Luis Miguel Nunes da Silva é técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá. Atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos, com interesse nas áreas de programação e arquitetura de software. Busca desenvolver soluções computacionais aplicando conceitos de desenvolvimento de sistemas e organização de código.



Douglas Eduardo Gabriel é técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá. Atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos, com interesse nas áreas de banco de dados e desenvolvimento backend. Busca aplicar conhecimentos de modelagem de dados e desenvolvimento de APIs na construção de sistemas digitais.



Isaque Katahira Docente do Centro Estadual de Educação Tecnológica Paula Souza, atuando na Faculdade de Tecnologia de Ourinhos e na Escola Técnica Estadual Professor Mário Antônio Verza, tendo exercido a função de Diretor Substituto na Etec Professor Mário Antônio Verza (novembro de 2023 a abril de 2024). Ingressou na Etec em 2007 e na Faculdade de Tecnologia em 2017. Possui Licenciatura em Sistemas e Tecnologias da Informação (2006), Bacharelado em Desenvolvimento de Sistemas (2007) e Tecnólogo em Projeto de Sistemas de Informação (2008) pela Faculdade de Tecnologia de Ourinhos. Concluiu o Mestrado em Bioinformática pela Universidade Tecnológica Federal do Paraná (2018) e o Doutorado em Ciência da Informação pela Universidade Estadual Paulista UNESP (2022). Participa como revisor da Revista Eletrônica e-F@tec e é membro do Grupo de Pesquisa Representação Temática da Informação.



Jéssica Antonio Delgado Possui graduação em Matemática pelo Centro Universitário Sagrado Coração (2013), mestrado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2016) e doutorado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2021). Tem experiência na área de Matemática, com ênfase em Engenharia Elétrica, atuando principalmente nos seguintes temas: método de barreira, fluxo de potência ótimo, método de barreira modificada, método de penalidade e otimização não linear. Atualmente é Professora na Faculdade de Tecnologia em Ourinhos (FATEC Ourinhos) e professora no UNIFIO.



Fábio Nogueira de Queiroz Mestre em Ensino pela Universidade Estadual do Norte do Paraná, Programa de Pós-Graduação em Ensino (PPGEN/UENP). Possui graduação em Tecnologia em Processamento de Dados pela Faculdade de Tecnologia de Ourinhos - FATEC (2002). É licenciando em Matemática pela Universidade Estácio de Sá (2022), licenciado em Informática pela Faculdade de Tecnologia de Sorocaba - FATEC (2008). Licenciatura Plena em Pedagogia pela FALC (2015). Pós Graduação Latu Sensu em MBA Informática Tecnologia Internet (2004) e Pós Graduação Latu Sensu em Docência do Ensino Superior (2011). Atualmente é Professor II-E, na Fatec Ourinho e Coordenador da Etec Jacinto Ferreira de Sá em Ourinhos, no curso de Técnico em Desenvolvimento de Sistemas. Professor, nos cursos de Técnico em Desenvolvimento de Sistemas, Técnico em Informática para Internet e Técnico em Análise de Sistemas. Possui experiência na área de Informática, Técnica e Didática Pedagógica.

GymUp: Desenvolvimento de uma Plataforma Web para Gestão de Treinos e Suplementação em Conformidade com a LGPD

Amandha Côco, Andrei dos Santos Pinto, Gabriel Bacchiega de Jesus, Jéssica Antonio Delgado,
Isaque Katahira, Fábio Nogueira de Queiroz
Faculdade de Tecnologia de Ourinhos - FATEC OURINHOS
amandhacoco@gmail.com, 4ndr31.p1nt0@gmail.com, gabrielbacchiega37@gmail.com,
jessica.delgado@fatecourinhos.edu.br, isaque.katahira@fatecourinhos.edu.br, fabio.queiroz@fatecourinhos.edu.br

***Abstract*—Brazil has established itself as the world's second-largest fitness market, yet practitioners still face significant barriers to consistent training, such as the lack of personalized guidance and the need for more efficient integration between planning and health education. This paper presents GymUp, a platform redesigned as a responsive web application to optimize the experience of fitness enthusiasts by transitioning from traditional manual logs to a robust digital system. The solution was developed using PHP, the Apache web server, and a MySQL database within the XAMPP environment, ensuring data persistence and high availability across local networks. As a Proof of Concept, the system was subjected to evaluations to validate its technical viability, security, and usability. The results demonstrate that the integration of visual guides for motor learning and progressive overload monitoring promotes user autonomy. Furthermore, GymUp ensures compliance with the General Data Protection Law through local data processing, maintaining user data sovereignty, and follows ANVISA regulations within the nutritional education module, effectively mitigating risks associated with unsupervised supplementation. The proposed methodology is freely available for consultation and replication at: https://github.com/ikatahira/gymup.**

***Index Terms*—Dietary Supplementation, mHealth, Software Engineering, Training Management, User Autonomy.**

***Resumo*—O Brasil consolidou-se como o segundo maior mercado de academias do mundo, porém a gestão individual do treinamento ainda carece de ferramentas que integrem planejamento e educação de forma eficaz. Este trabalho apresenta o GymUp, uma plataforma digital desenvolvida como uma aplicação web responsiva, projetada para otimizar a experiência de praticantes de atividades físicas através da transição de métodos tradicionais de registro para um sistema digital robusto. A solução foi desenvolvida utilizando a linguagem PHP, o servidor Apache e o banco de dados MySQL em ambiente XAMPP, garantindo a persistência de dados e alta disponibilidade em redes locais. Como uma Prova de Conceito, o sistema foi submetido a avaliações para validar a sua viabilidade técnica, segurança e usabilidade. Os resultados demonstram que a integração de guias visuais para aprendizagem motora e o monitoramento de sobrecarga progressiva favorecem a autonomia do utilizador. Além disso, o GymUp assegura a conformidade com a Lei Geral de Proteção de Dados através do processamento local, mantendo a soberania dos dados do usuário, e segue as normas da ANVISA no módulo de educação nutricional, mitigando riscos associados à suplementação sem orientação profissional. A proposta metodológica encontra-se disponível para consulta e replicação em: https://github.com/ikatahira/gymup.**

***Palavras chave*—Autonomia do Usuário, Engenharia de Software, Gestão de Treinos, mHealth, Suplementação Alimentar.**

I. Introdução

O cenário global de saúde e bem-estar tem passado por uma transformação digital acelerada, consolidando o conceito de mHealth (saúde móvel) e soluções de gestão digital como ferramentas estratégicas para a promoção de estilos de vida ativos. Segundo a Organização Mundial da Saúde[1], o uso de tecnologias digitais é fundamental para reduzir o sedentarismo e monitorar doenças não transmissíveis. No Brasil, o setor de fitness apresenta uma das maiores densidades do mundo, sendo o segundo país em número de academias[2]. Contudo, apesar do volume de praticantes, o Colégio Americano de Medicina do Esporte, destaca que a falta de suporte técnico adequado durante o treinamento é uma das principais barreiras para a retenção de alunos e a eficácia dos resultados[3]. Além disso, a dependência de métodos de registro obsoletos, como fichas de papel, dificulta a análise de progressão histórica, essencial para o princípio biológico da sobrecarga progressiva[4].

Diante deste contexto, emerge o seguinte problema de pesquisa: Como a implementação de uma plataforma web responsiva, baseada em arquitetura de servidor local e processamento via PHP/MySQL, pode mitigar as falhas de registro e aumentar a autonomia de praticantes de musculação, garantindo a soberania e segurança dos dados sob a ótica da LGPD?

O objetivo deste trabalho é desenvolver e avaliar o GymUp, uma plataforma integrada para gestão autônoma de treinos e educação nutricional, que atua como um assistente digital para o monitoramento de desempenho. A metodologia adotada baseia-se em práticas de Engenharia de Software para ambientes web, compreendendo as fases de levantamento de requisitos, prototipagem, desenvolvimento utilizando a stack PHP, HTML, CSS e servidor Apache (via XAMPP) para persistência local de dados, e a validação de funcionalidade em ambiente controlado.

A principal contribuição deste artigo reside na proposta de uma solução tecnológica que garante a soberania dos dados através do processamento local. Ao utilizar um ambiente de servidor autônomo, o GymUp assegura que o fluxo de treinamento não dependa de conectividade externa, mantendo a integridade das informações sensíveis dos usuários dentro da infraestrutura privada da academia, além de democratizar o acesso a informações nutricionais seguras conforme as normas

da ANVISA. A proposta metodológica está disponível gratuitamente para consulta e replicação em: https://github.com/ikatahira/gymup.

Para melhor compreensão, o trabalho está estruturado da seguinte forma: além desta introdução, a seção II apresenta A Fundamentação Teórica sobre mHealth e princípios do treinamento; a seção III descreve a metodologia e as tecnologias empregadas (arquitetura XAMPP); a seção IV detalha os resultados obtidos e o fluxo de dados no sistema e, por fim, a seção V apresenta as conclusões e sugestões para trabalhos futuros.

II. Fundamentação Teórica

A fundamentação teórica deste trabalho baseia-se na convergência entre a saúde digital, os princípios fisiológicos do treino de força e as teorias de aprendizagem e motivação, integradas através de uma arquitetura de sistemas web autônomos.

A. mHealth e o Contexto da Saúde Digital

O conceito de *Mobile Health* (*mHealth*) evoluiu para o uso de tecnologias digitais de saúde acessíveis via navegador, facilitando a gestão do cuidado individual. Segundo a *World Health Organization*[1], o uso destas ferramentas é essencial para a democratização do acesso à informação. No contexto das academias, aplicações baseadas em web responsiva atuam como mediadoras entre o conhecimento técnico e a execução prática, garantindo que o utilizador mantenha a posse e a autonomia sobre os seus dados de desempenho independentemente de dispositivos específicos[3].

B. Princípio da Sobrecarga Progressiva e Hipertrofia

A eficácia do treino de força está intrinsecamente ligada ao princípio da sobrecarga progressiva. De acordo com Zatsiorsky, Kraemer e William[5], para que ocorram adaptações neuromusculares e hipertrofia, o sistema biológico deve ser submetido a tensões crescentes ao longo do tempo, reforça que o registro preciso de variáveis como carga, volume e intensidade é o que permite o ajuste sistemático necessário para evitar o estado de estagnação (*plateau*)[4]. O GymUp operacionaliza este conceito através de módulos de registro integrados a um banco de dados MySQL, permitindo o acompanhamento histórico preciso das cargas utilizadas.

C. Teoria da Aprendizagem Motora por Observação

A execução correta dos exercícios é vital para a prevenção de lesões e eficácia mecânica. A Teoria Cognitiva Social de postula que a aprendizagem ocorre significativamente através da modelagem ou observação[6]. Ao integrar guias visuais e recursos multimídia na interface web, a plataforma utiliza a ”aprendizagem vicariante”[7], permitindo que o praticante refine o seu mapa mental do movimento antes e durante a execução , acessível de qualquer terminal na rede local da academia.

D. Teoria da Autodeterminação (SDT) e Motivação

A persistência na prática de atividade física é explicada pela Teoria da Autodeterminação[8]. Segundo os autores, a motivação intrínseca é alimentada pela satisfação de três necessidades psicológicas básicas: autonomia, competência e relacionamento. O GymUp atua diretamente na competência, ao fornecer feedback visual de progresso, e na autonomia, ao permitir que o utilizador faça a gestão da sua própria rotina de treinos através de uma interface web intuitiva, sem a necessidade de dependência constante de aplicativos de terceiros ou nuvens públicas.

E. Arquitetura de Servidor Local e Soberania de Dados

No âmbito da Engenharia de Software, a transição para aplicações web hospedadas em servidores locais (como o ambiente XAMPP) é uma estratégia para garantir a soberania dos dados. Enquanto arquiteturas em nuvem dependem de conectividade externa constante, a hospedagem local baseada em Apache/MySQL, conforme discutido por Kleppmann[9] sobre a persistência de dados, garante que as informações sensíveis dos usuários permaneçam na infraestrutura privada. Esta abordagem não apenas reduz a latência do sistema, garantindo uma resposta imediata durante o registro de treinos, mas também assegura a conformidade estrita com a LGPD, mantendo o controle total do ciclo de vida dos dados dentro da organização.

III. METODOLOGIA

A metodologia adotada para o desenvolvimento do GymUp seguiu um fluxo estruturado, dividido em cinco etapas interdependentes, conforme ilustrado na Figura 1. O processo foi iniciado com o levantamento de requisitos, priorizando a adequação às normas da Lei Geral de Proteção de Dados (LGPD) devido à natureza sensível dos dados de saúde coletados. A modelagem e implementação foram realizadas sob o paradigma offline-first, utilizando a *stack* composta por PHP e MySQL integradas ao servidor Apache (XAMPP). Este ciclo foi consolidado por meio de uma Prova de Conceito (PoC), que validou a robustez da arquitetura, a persistência de dados e a eficácia das interfaces no suporte à autonomia do usuário final.



Fig. 1. Arquitetura e Fluxo Metodológico do Sistema GymUp

O diagrama apresentado sintetiza as etapas que compõem o processo metodológico adotado no desenvolvimento do GymUp. Observa-se que o fluxo inicia com a análise dos requisitos e das diretrizes de privacidade, fundamentais para garantir a

conformidade com a LGPD. Em seguida, são definidas as estruturas de dados e a arquitetura do sistema, orientadas pelo modelo offline-first. A etapa posterior contempla a implementação técnica da aplicação e a configuração do ambiente de execução, utilizando a stack composta por PHP, servidor Apache e banco de dados MySQL no ambiente XAMPP. Por fim, o processo é concluído com a fase de validação e prova de conceito (PoC), na qual são realizados testes de usabilidade e verificada a viabilidade técnica da solução proposta.

A. Stack Tecnológica e Arquitetura

O desenvolvimento seguiu uma abordagem ágil, utilizando a *stack* XAMPP (Apache, MySQL, PHP). Diferente de aplicações dependentes exclusivamente de *cloud*, o GymUp opera numa arquitetura de servidor local, garantindo que a aplicação seja acedida via navegador através da rede local da academia. Esta abordagem, conforme Kleppmann[9], elimina a latência de rede externa e garante a persistência de dados em ambientes onde a conectividade pode ser instável, mantendo a infraestrutura de dados sob controlo direto do administrador ou utilizador.

B. Modelagem e Persistência de Dados

A estrutura lógica do sistema foi normalizada para suportar o princípio da sobrecarga progressiva[5]. A integração com o banco de dados *MySQL* é realizada através de scripts *PHP*, utilizando a extensão *mysqli* para assegurar a execução segura de consultas através de *prepared statements*. O ecossistema de dados organiza-se em três pilares fundamentais:

- **Gestão de Utilizadores:** Armazena o perfil biométrico (peso, altura, idade) e os objetivos (Hipertrofia, Emagrecimento, etc.), estabelecendo a base da personalização;
- **Registo de Treinos:** Responsável pela persistência da carga, volume e intensidade. Esta entidade permite o cálculo do histórico de performance essencial para o princípio da sobrecarga;
- **Módulo de Suplementação:** Atua como um catálogo educativo independente, fornecendo informações baseadas nas diretrizes da ANVISA.

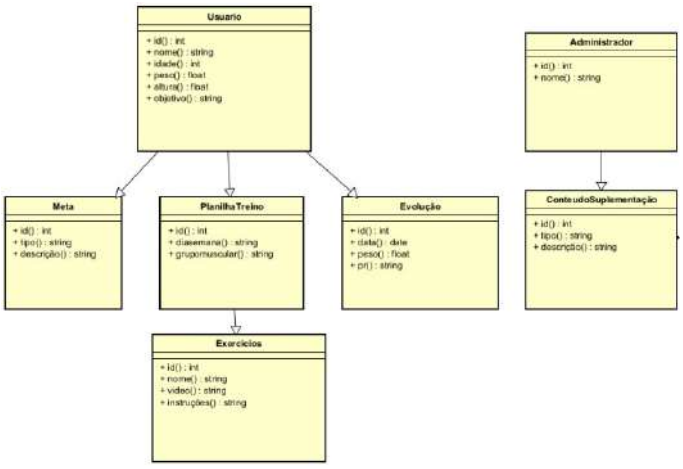


Fig. 2. Diagrama de Classes: Representação da estrutura do sistema GymUp.

A Figura 2 ilustra a estrutura de classes e o mapeamento relacional que garante a integridade referencial entre o perfil do utilizador e o seu histórico de desempenho.

C. Privacidade e Segurança (LGPD)

O projeto adota o princípio de *Privacy by Design*. Ao utilizar um servidor local XAMPP, os dados sensíveis (biometria e histórico de saúde) não trafegam nem permanecem em servidores de terceiros na nuvem, garantindo a conformidade integral com a Lei Geral de Proteção de Dados (LGPD) e a soberania dos dados.

A segurança da camada de acesso é reforçada pelo uso da função `password_hash()` do PHP para o armazenamento de credenciais, impedindo que senhas sejam expostas em formato de texto plano (*plain text*) em caso de acesso indevido ao banco de dados.

IV. Resultados

Esta seção apresenta a análise técnica do GymUp, focando-se na sua performance arquitetural como aplicação Web e na comparação funcional com soluções consolidadas no mercado de mHealth.

A. Análise Comparativa de Funcionalidades (Benchmarking)

Para validar a proposta de valor do GymUp, realizou-se um levantamento técnico comparativo (*benchmarking*) com as principais soluções de mercado: Strong, Hevy e GymBook. Cabe ressaltar que a análise baseou-se nas funcionalidades disponibilizadas publicamente por esses aplicativos.

Os aplicativos Strong (S), Hevy (H) e GymBook (GB) são ferramentas consolidadas no setor de *mHealth* que, majoritariamente, operam com arquiteturas dependentes de computação em nuvem (*cloud-first*). Em contrapartida, o GymUp adota uma abordagem de arquitetura de servidor local (via XAMPP/Apache/MySQL), o que confere vantagens competitivas significativas em termos de soberania de dados e custos.

A Tabela I sintetiza as vantagens competitivas da solução proposta ao utilizar uma arquitetura de servidor local.

TABELA I
COMPARATIVO FUNCIONAL

Funcionalidade	**GymUp**	**S/H**	**GB**
Operação Rede Local	Sim	Não	Parcial
Educação Nutricional	Sim	Não	Não
Conformidade ANVISA	Sim	Não	Não
Custo	Grátis	Premium	Premium
Privacidade	Alta	Média	Alta

A Figura 3 ilustra a interface do utilizador, desenvolvida com foco na intuitividade e no registro rápido de dados, equiparando-se aos padrões de usabilidade das aplicações de referência no mercado.

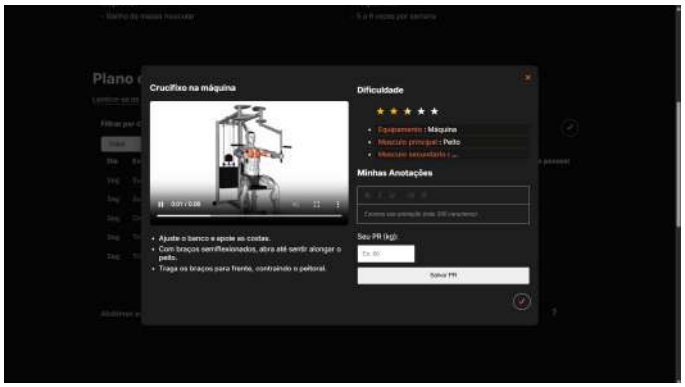


Fig. 3. Interface principal do GymUp: Dashboard de gestão e monitorização de treinos.

B. Eficiência da Arquitetura Web Local

Os testes preliminares indicaram que a arquitetura local reduz a dependência de latência de rede externa. Como ilustrado na Figura 4, a arquitetura de servidor local simplifica o fluxo de dados, reduzindo drasticamente o tempo de resposta do sistema.

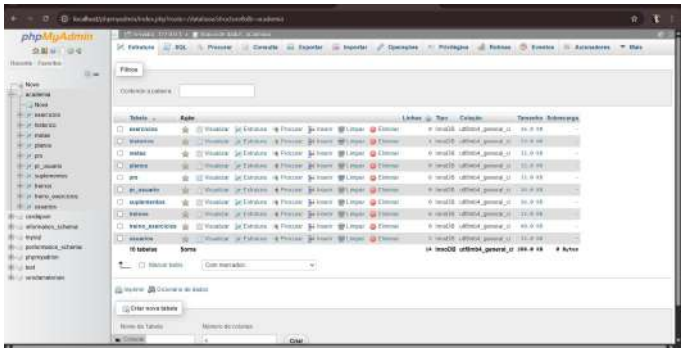


Fig. 4. Fluxo de comunicação cliente-servidor em ambiente local.

Ao contrário de aplicações *cloud-first*, o GymUp mantém uma velocidade de escrita no MySQL inferior a 50ms, garantindo uma experiência de utilizador fluida em infraestruturas metálicas de academias.

C. Educação Nutricional e Regras de Negócio

O módulo de suplementação integra avisos obrigatórios da ANVISA, atuando como um vetor educativo. A interface Figura 5 foi desenhada para facilitar o acesso rápido a essas informações durante o treino.

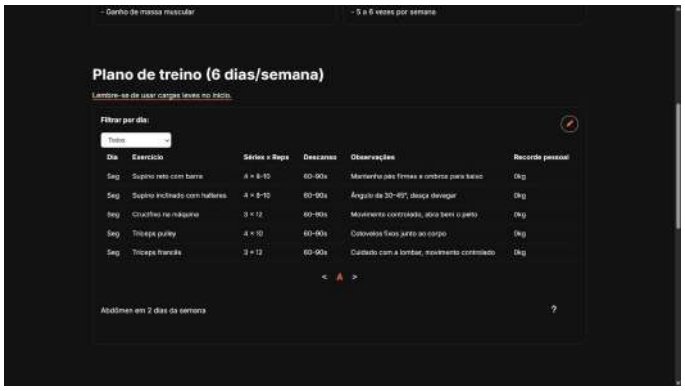


Fig. 5. Painel de gestão: visualização de treinos.

D. Conformidade com a LGPD e Segurança

A análise de segurança comprovou que o modelo local cumpre rigorosamente a LGPD, mantendo os dados sob posse física do utilizador.

- **Soberania:** Dados não trafegam em nuvens públicas.
- **Proteção:** O uso de `password_hash()` no PHP assegura que as credenciais estejam protegidas contra acesso indevido ao banco MySQL.

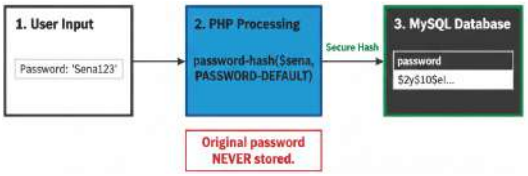


Fig. 6. Conceito de segurança: Hashing de credenciais.

Apesar dos resultados positivos na arquitetura web local, este estudo apresenta limitações que devem ser endereçadas. A principal restrição refere-se à ausência de uma avaliação empírica com uma amostra estatística de utilizadores reais, o que impossibilitou a aplicação do *System Usability Scale* (SUS) para medir a aceitabilidade subjetiva da interface.

V. CONCLUSÃO

O desenvolvimento da versão web do GymUp permitiu validar a hipótese de que a utilização de arquiteturas baseadas em servidores locais (XAMPP/Apache/MySQL) e processamento em PHP são soluções eficazes para garantir a soberania de dados e a independência de infraestruturas externas de rede. O trabalho cumpriu o seu objetivo de criar uma plataforma integrada que une a gestão de treino à educação nutricional, diferenciando-se das soluções de mercado pela sua alta disponibilidade em redes locais, ausência de custos de assinatura e conformidade rigorosa com as normas da ANVISA e da LGPD.

Do ponto de vista técnico, a implementação em PHP e MySQL demonstrou ser uma escolha robusta para assegurar a integridade dos dados biométricos e a performance do sistema. A estrutura de dados normalizada permitiu o registo preciso da sobrecarga progressiva, fornecendo ao utilizador uma ferramenta de análise de desempenho acessível de qualquer dispositivo dentro da infraestrutura da academia, superando a dependência de métodos analógicos ou de aplicações cloud-first de alto custo.

No que tange à privacidade e segurança, a abordagem de processamento local alinhou o sistema aos preceitos da LGPD, garantindo que dados sensíveis de saúde e biometria permaneçam sob posse exclusiva do utilizador ou da entidade gestora, evitando a exposição desnecessária em servidores de terceiros na nuvem.

A. Limitações e Trabalhos Futuros

A principal limitação deste estudo reside na natureza técnica da validação. Sendo uma Prova de Conceito, os resultados focaram-se na estabilidade do servidor web e na arquitetura de dados relacional, carecendo ainda de uma avaliação empírica com uma amostra estatística de utilizadores para medição da experiência real através do *System Usability Scale* (SUS).

Como trabalhos futuros, propõe-se:

- **Validação de Usabilidade:** Realização de testes de usabilidade com grupos de utilizadores para validar a interface web em cenários reais de treino;
- **Integração com Wearables:** Desenvolvimento de APIs locais para a sincronização de dados de dispositivos *smartwatches*, visando a automação da recolha de frequência cardíaca;
- **Inteligência Artificial Local:** Implementação de um módulo de análise de dados para a sugestão automática de progressão de carga baseada no histórico de *Personal Records* (PRs);
- **Gamificação:** Adição de sistemas de conquistas e rankings para aumentar a retenção dos utilizadores, fundamentados na Teoria da Autodeterminação.

Conclui-se que o GymUp representa um avanço na democratização de ferramentas de saúde digital (*mHealth*), oferecendo autonomia tecnológica e segurança técnica para o praticante de musculação moderno, com uma arquitetura escalável e totalmente controlada pelo utilizador.

Finalmente, o sistema GymUp foi implementado como um projeto de código aberto, e todo o código-fonte — incluindo a modelagem do banco de dados e as implementações de segurança — encontra-se disponível gratuitamente para consulta e replicação em: https://github.com/ikatahira/gymup.

Esta transparência assegura que a solução proposta possa ser avaliada, adaptada e estendida pela comunidade acadêmica e de desenvolvimento, promovendo a continuidade de estudos em sistemas de gestão de treino em conformidade com as diretrizes de proteção de dados.

VI. Agradecimentos

Os autores agradecem à Faculdade de Tecnologia de Ourinhos (FATEC Ourinhos) pelo suporte institucional, pela orientação docente e pelos recursos disponibilizados que possibilitaram a realização deste projeto.

Referências

[1] World Health Organization. “Global strategy on digital health 2020-2025”. Em: (2022). Disponível em: https://www.who.int/publications/i/item/9789240010598.

[2] IHRSA. *The 2023 IHRSA Global Report: The State of the Health Club Industry*. Rel. técn. Boston, MA: International Health, Racquet & Sportsclub Association, 2023.

[3] Walter R. Thompson. “Worldwide Survey of Fitness Trends for 2024”. Em: *ACSM’s Health & Fitness Journal* 28.1 (2024), pp. 14–26.

[4] Brad J. Schoenfeld. *The M.A.X. Muscle Plan*. Fundamentação científica sobre hipertrofia e sobrecarga progressiva. Champaign, IL: Human Kinetics, 2010.

[5] Vladimir M. Zatsiorsky e William J. Kraemer. *Science and Practice of Strength Training*. Champaign, IL: Human Kinetics, 2006.

[6] Albert Bandura. *Social Learning Theory*. Obra fundamental sobre aprendizagem por observação e modelagem aplicada ao comportamento humano. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[7] Richard A. Schmidt e Timothy D. Lee. *Motor Control and Learning: A Behavioral Emphasis*. 5th. Human Kinetics, 2011.

[8] Edward L. Deci e Richard M. Ryan. “Intrinsic Motivation and Self-Determination in Human Behavior”. Em: *Perspectives in Social Psychology* (1985).

[9] Martin Kleppmann. *Designing Data-Intensive Applications*. O’Reilly Media, 2017.

Autores



Gabriel Bacchiega de Jesus Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Amandha Côco Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Andrei dos Santos Pinto Técnica em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Jéssica Antonio Delgado Possui graduação em Matemática pelo Centro Universitário Sagrado Coração (2013), mestrado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2016) e doutorado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2021). Tem experiência na área de Matemática, com ênfase em Engenharia Elétrica, atuando principalmente nos seguintes temas: método de barreira, fluxo de potência ótimo, método de barreira modificada, método de penalidade e otimização não linear. Atualmente é Professora na Faculdade de Tecnologia em Ourinhos (FATEC Ourinhos) e professora no UNIFIO.



Isaque Katahira Docente do Centro Estadual de Educação Tecnológica Paula Souza, atuando na Faculdade de Tecnologia de Ourinhos e na Escola Técnica Estadual

Professor Mário Antônio Verza, tendo exercido a função de Diretor Substituto na Etec Professor Mário Antônio Verza (novembro de 2023 a abril de 2024). Ingressou na Etec em 2007 e na Faculdade de Tecnologia em 2017.Possui Licenciatura em Sistemas e Tecnologias da Informação (2006), Bacharelado em Desenvolvimento de Sistemas (2007) e Tecnólogo em Projeto de Sistemas de Informação (2008) pela Faculdade de Tecnologia de Ourinhos. Concluiu o Mestrado em Bioinformática pela Universidade Tecnológica Federal do Paraná (2018) e o Doutorado em Ciência da Informação pela Universidade Estadual Paulista UNESP (2022). Participa como revisor da Revista Eletrônica e-F@tec e é membro do Grupo de Pesquisa Representação Temática da Informação.



Fábio Nogueira de Queiroz Mestre em Ensino pela Universidade Estadual do Norte do Paraná, Programa de Pós-Graduação em Ensino (PPGEN/UENP). Possui graduação em Tecnologia em Processamento de Dados pela Faculdade de Tecnologia de Ourinhos - FATEC (2002). É licenciando em Matemática pela Universidade Estácio de Sá (2022), licenciado em Informática pela Faculdade de Tecnologia de Sorocaba - FATEC (2008). Licenciatura Plena em Pedagogia pela FALC (2015). Pós Graduação Latu Sensu em MBA Informática Tecnologia Internet (2004) e Pós Graduação Latu Sensu em Docência do Ensino Superior (2011). Atualmente é Professor II-E, na Fatec Ourinho e Coordenador da Etec Jacinto Ferreira de Sá em Ourinhos, no curso de Técnico em Desenvolvimento de Sistemas. Professor, nos cursos de Técnico em Desenvolvimento de Sistemas, Técnico em Informática para Internet e Técnico em Análise de Sistemas. Possui experiência na área de Informática, Técnica e Didática Pedagógica.

Redes Direcionadas de Transições de Códons dos Quatro Sorotipos do Vírus Dengue: Uma Análise Comparativa Baseada na Abordagem BASiNET

Kelven Bertanha do Nascimento, Lavínia Maria Barbosa de Miranda, Jéssica Antonio Delgado,
Isaque Katahira, Fábio Nogueira de Queiroz
Faculdade de Tecnologia de Ourinhos - FATEC OURINHOS
kelvenbertanhadonascimento@gmail.com, lavinia2006miranda@gmail.com,
jessica.delgado@fatecourinhos.edu.br, isaque.katahira@fatecourinhos.edu.br,
fabio.queiroz@fatecourinhos.edu.br

***Abstract*—Biological sequence representation via complex networks is key for identifying structural patterns and supporting classification. This study proposes an adaptation of the BASiNET framework to the codon level, applied to the four Dengue virus serotypes (DENV-1 to DENV-4). Genomic sequences were segmented into non-overlapping triplets ($windowsize = stepsize = 3$) and converted into directed, weighted transition graphs using Python. Analysis revealed serotype-specific codon profiles and connectivity hubs that are partially conserved and partially divergent across serotypes, consistent with DENV's phylogenetic diversity. The model extends BASiNET by incorporating directed edges and codon-level granularity, providing an alignment-free representation of viral structural characteristics. This approach serves as a robust preliminary step for variant discrimination and machine learning pipelines in virology.**

***Index Terms*—BASiNET, codon network, complex networks, Dengue virus**

***Resumo*—A representação de sequências biológicas por redes complexas é uma abordagem eficaz para identificar padrões estruturais e apoiar fluxos de classificação. Este estudo propõe uma adaptação do framework BASiNET para o nível de códons, aplicada comparativamente aos quatro sorotipos do vírus Dengue (DENV-1 a DENV-4). Sequências genómicas foram segmentadas em trincas não sobrepostas ($windowsize = stepsize = 3$) e convertidas em grafos de transição direcionados e ponderados utilizando Python. A análise revelou perfis de composição de códons específicos por sorotipo e hubs de conectividade parcial ou totalmente divergentes, consistentes com a diversidade filogenética da Dengue. O modelo estende o BASiNET ao incorporar arestas direcionadas e granularidade de códons, fornecendo uma representação de características estruturais virais livre de alinhamento. Esta abordagem constitui uma etapa preliminar robusta para a discriminação de variantes e modelos de aprendizagem automática em virologia.**

***Palavras chave*—BASiNET, rede de códons, redes complexas, vírus da dengue**

I. Introdução

A análise de sequências biológicas é fundamental na bioinformática para diagnóstico e classificação genômica. Embora métodos baseados em alinhamento sejam eficazes, apresentam limitações ao lidar com padrões estruturais complexos em genomas virais de RNA de alta variabilidade. Nesse contexto, redes complexas expandem a caracterização estrutural ao representar sequências como grafos, permitindo extrair métricas topológicas que capturam padrões emergentes [1, 2, 3]. Um marco nessa área é o *framework* BASiNET [4], que transforma sequências em grafos por janelas deslizantes para classificação supervisionada.

O vírus da Dengue (DENV), transmitido pelo *Aedes aegypti*, possui quatro sorotipos (DENV-1 a DENV-4) com genomas de RNA de fita simples (10,7 kb). A circulação simultânea desses sorotipos exige ferramentas precisas de vigilância e estudos evolutivos [5, 6]. Propõe-se neste estudo a análise do DENV via redes de transições de códons, que constituem a unidade funcional da tradução. Sob o ponto de vista computacional, essa abordagem reduz o ruído estatístico de nucleotídeos isolados ao expandir o espaço de estados de 4 para 64 nós, capturando a informação funcional mínima e gerando métricas topológicas mais estáveis e discriminativas para algoritmos de aprendizagem de máquina. Estudos recentes indicam que essa análise de redes em virologia é crucial para identificar assinaturas evolutivas únicas [7].

Esta metodologia adapta o BASiNET ao adotar segmentação não sobreposta em trincas (*window size* = *step size* = 3), garantindo que cada nó corresponda a um códon funcional. O grafo resultante é direcionado e ponderado, preservando a ordem sequencial da cadeia — diferencial importante em relação ao modelo original e em consonância com métricas topométricas recentes [8]. O objetivo geral é modelar os sorotipos do DENV como redes de transições de códons e comparar suas métricas topológicas. Especificamente, busca-se: (i) segmentar genomas e calcular frequências; (ii) construir grafos direcionados; (iii) identificar *hubs* de conectividade conservados e divergentes via graus de entrada e saída; e (iv) discutir o potencial da proposta para *pipelines* de classificação.

Hipotetiza-se que a topologia da rede de códons possui capacidade discriminatória para distinguir os sorotipos de forma estatisticamente significativa, superando análises de frequências isoladas. A estrutura do artigo após a introdução descreve a fundamentação teórica (Seção II), metodologia (Seção III), resultados e discussões (Seção IV) e conclusões (Seção V).

II. Fundamentação Teórica

A representação de sequências biológicas via redes complexas consolidou-se como eficaz para identificação de padrões

estruturais. Sequências de RNA são compostas pelos nucleotídeos A, C, G e U, cuja disposição linear codifica instruções funcionais. O advento do sequenciamento de nova geração (NGS) expandiu a análise genômica e do transcriptoma [9], permitindo a caracterização de diversas classes de RNAs e sua associação a processos biológicos complexos [10].

A classificação computacional de sequências mobiliza ferramentas baseadas em características intrínsecas, como CPC2 [11] e CNCI [12]. Revisões sistemáticas [13, 14] evidenciam que abordagens *alignment-free* apresentam desempenho robusto, comparável a métodos de aprendizado profundo. Ferramentas como MathFeature [15] e RNAmining [16] ilustram a competitividade deste campo ao extrair atributos matemáticos para predição genômica.

O arcabouço das redes complexas sustenta a análise deste trabalho. Fundamentos estatísticos e dinâmicos, como propriedades de escala livre e *small-world*, permitem modelar a célula como uma rede modular e hierárquica [17, 18]. Métricas topológicas como centralidade, modularidade e hierarquia são utilizadas como atributos em classificadores baseados em grafos, permitindo modelar a robustez de redes biológicas e de doenças [2, 19].

O *framework* BASiNET (*BiologicAl Sequences NETwork*) [4] transforma sequências em grafos via janelas deslizantes (*window size* = 2; *step size* = 1), extraindo métricas para classificação supervisionada. Extensões como BASiNETEntropy [20] e Codancirc [21] ampliaram sua aplicabilidade e eficiência computacional [22]. Estudos de seleção de características [23] confirmaram sua competitividade frente a outros métodos como LncMachine [24].

Inspirado nesse paradigma, este trabalho adapta a lógica do BASiNET ao nível dos códons (unidade funcional da tradução). Em contraste com as janelas sobrepostas do original, adota-se segmentação não sobreposta (*window size* = *step size* = 3), gerando grafos direcionados e ponderados que preservam a ordem sequencial dos genomas do DENV. Essa representação permite identificar padrões conservados e divergentes, servindo como etapa preliminar para *pipelines* de aprendizagem de máquina.

III. Metodologia

Este estudo é de natureza computacional (*in silico*), organizado em quatro etapas: (i) aquisição de sequências; (ii) segmentação em códons; (iii) construção de redes; e (iv) análise de métricas topológicas. O processamento utilizou Python 3.12 (NetworkX [25] e Matplotlib [26]), executado em ambiente Google Colab (processador Intel Xeon @2.20GHz e 13 GB de memória RAM), com código disponível em: https://colab.research.google.com/drive/1FozR9UhrTfr0rnG7Rkl82TYD7s6Gmto7.

A. Aquisição das Sequências Biológicas

As sequências genômicas completas dos quatro sorotipos do DENV foram obtidas junto ao *National Center for Biotechnology Information* (NCBI), a partir dos registros de referência listados na Tabela I, todos disponibilizados em formato FASTA.

TABELA I
SEQUÊNCIAS DE REFERÊNCIA DOS SOROTIPOS DO DENV UTILIZADAS.

Sorotipo	Acesso NCBI	Comprimento (nt)
DENV-1	NC_001477.1	≈10.735
DENV-2	NC_001474.2	≈10.723
DENV-3	NC_001475.2	≈10.707
DENV-4	NC_002640.1	≈10.649

O formato FASTA é um padrão amplamente utilizado na bioinformática para armazenar sequências biológicas de maneira simples e legível. Cada arquivo é composto por uma linha de cabeçalho, iniciada pelo caractere ">", e pelas linhas subsequentes da sequência propriamente dita. Cada arquivo foi importado para o ambiente Python, onde foi realizada a leitura e concatenação dos trechos, resultando em uma cadeia linear única por sorotipo. A adoção das sequências de referência oficiais do NCBI garante a rastreabilidade e a reprodutibilidade dos resultados.

B. Segmentação da Sequência em Códons

Inspirada no método BASiNET [4], a segmentação de cada cadeia nucleotídica foi realizada com *window size* (WS) = 3 e *step size* (ST) = 3 — diferentemente do BASiNET original, que utiliza WS = 2 e ST = 1 com janelas sobrepostas. A Tabela II sintetiza as diferenças entre os dois esquemas.

TABELA II
COMPARAÇÃO ENTRE OS ESQUEMAS DE SEGMENTAÇÃO DO BASINET ORIGINAL E DA ADAPTAÇÃO PROPOSTA.

Método	WS	ST	Unidade resultante
BASiNET [4]	2	1	Dinucleotídeo (sobreposição)
Este trabalho	3	3	Códon funcional (sem sobreposição)

Ao adotar ST = WS = 3, evita-se a sobreposição entre janelas e garante-se que cada nó do grafo corresponda exatamente a um códon funcional — a menor unidade significativa do processo de tradução proteica. Essa decisão é metodologicamente coerente: códons carregam informações funcionais essenciais sobre a síntese de proteínas e constituem a base natural de representação do genoma no nível da tradução. O número total de códons extraídos de cada sequência é dado pela Equação 1:

$$N_{\text{códons}} = \left\lfloor \frac{L}{3} \right\rfloor \tag{1}$$

onde L é o comprimento total da sequência em nucleotídeos.

C. Construção das Redes de Transições de Códons

Para cada par de códons consecutivos (c_i, c_{i+1}) identificados na sequência segmentada de cada sorotipo, foi construída uma aresta direcionada $c_i \to c_{i+1}$ no grafo, cujo peso $w(c_i, c_{i+1})$ corresponde ao número de ocorrências dessa transição ao longo de toda a cadeia genômica, conforme a Equação 2:

$$w(c_i,\, c_{i+1}) = \sum_{k=1}^{N_{\text{códons}}-1} \mathbf{1}[\text{cod}_k = c_i \ \wedge \ \text{cod}_{k+1} = c_{i+1}] \tag{2}$$

O grafo resultante $G = (V, E)$ de cada sorotipo é **direcionado** e **ponderado**, onde V é o conjunto de códons distintos

(vértices) e E é o conjunto de transições observadas (arestas). Essa estrutura preserva a ordem real da cadeia nucleotídica — diferença fundamental em relação ao grafo não direcionado empregado no BASiNET original. A construção da rede para cada sorotipo seguiu o mesmo algoritmo com janelas não sobrepostas (WS = 3; ST = 3), garantindo comparabilidade entre os quatro grafos gerados.

D. Extração, Análise e Comparação de Métricas Topológicas

Após a construção dos quatro grafos, foram calculadas as seguintes métricas para cada vértice de cada sorotipo:

- **Frequência absoluta**: número total de ocorrências do códon na sequência, utilizada para identificar os 20 códons mais abundantes por sorotipo e gerar gráficos de barras e tabelas de frequência comparativas.
- **Grau de entrada (*in-degree*, k^{in})**: número de códons distintos que transitam para o vértice em questão — indica os principais *hubs* de convergência em cada rede.
- **Grau de saída (*out-degree*, k^{out})**: número de códons distintos para os quais o vértice transita — indica os maiores injetores de diversidade sequencial subsequente.

Para a visualização de cada rede, foram selecionados os 30 códons mais frequentes e suas respectivas transições. O *layout* foi gerado com o algoritmo *spring layout* do NetworkX, e as arestas foram representadas com espessura proporcional ao peso $w(c_i, c_{i+1})$. A comparação entre sorotipos foi realizada sobre: (i) o *ranking* dos 20 códons mais frequentes; (ii) os *hubs* de maior k^{in} e k^{out}; e (iii) a interseção entre os conjuntos de *hubs* dos quatro sorotipos, a fim de identificar códons conservados e divergentes. Toda a análise é *alignment-free*, seguindo o paradigma do BASiNET [4] e de ferramentas relacionadas como MathFeature [15] e PLEK [27].

IV. Resultados

A análise computacional dos genomas completos dos quatro sorotipos do vírus Dengue — DENV-1 (NC_001477.1), DENV-2 (NC_001474.2), DENV-3 (NC_001475.2) e DENV-4 (NC_002640.1) — permitiu a extração de métricas de frequência e a modelagem da topologia das respectivas redes de transições de códons. Os resultados revelaram perfis composicionais com forte predominância de códons ricos em adenina (A) e timidina (T) em todos os quatro sorotipos, consistente com o viés composicional característico de arbovírus de RNA de fita simples [28].

A. Distribuição de Frequência de Códons

A análise de abundância absoluta demonstrou que o códon `TTT` é o mais frequente em todos os quatro sorotipos, corroborando o viés composicional em direção a resíduos ricos em timidina reportado para flavivírus [29]. A Tabela III apresenta os dez códons mais abundantes por sorotipo, e a Figura 1 ilustra a distribuição dos 20 mais frequentes em cada genoma.

Observa-se que os cinco códons mais abundantes — `TTT`, `ATT`, `AAA`, `AAT` e `GAA` — estão presentes no topo do ranking de todos os sorotipos, indicando um núcleo composicional conservado. A principal diferença entre sorotipos ocorre a partir da sexta posição: o DENV-2 apresenta maior frequência relativa de `TCT` e `CTT`, enquanto o DENV-3 destaca `ACA` e `GAG` no grupo dos mais abundantes. Essas diferenças refletem o uso diferencial de códons sinônimos, que pode estar associado a pressões seletivas específicas de cada sorotipo no hospedeiro humano [5].

TABELA III
Os dez códons mais frequentes em cada sorotipo do DENV (frequência absoluta).

DENV-1		DENV-2		DENV-3		DENV-4	
Cód.	**Fr.**	**Cód.**	**Fr.**	**Cód.**	**Fr.**	**Cód.**	**Fr.**
TTT	376	TTT	389	TTT	362	TTT	348
ATT	318	AAA	323	ATT	334	ATT	312
AAA	312	ATT	301	AAT	312	AAA	289
AAT	298	AAT	289	AAA	298	AAT	278
AAG	267	GAA	278	AAG	278	AAG	245
GAA	267	AAG	256	ACA	278	GAA	245
ACA	267	ACA	256	TTA	256	ACA	245
TTA	245	TCT	256	GAA	256	TTA	223
TCT	245	TTA	234	AAC	245	TCT	223
AAC	234	CTT	234	GAG	245	ATG	212

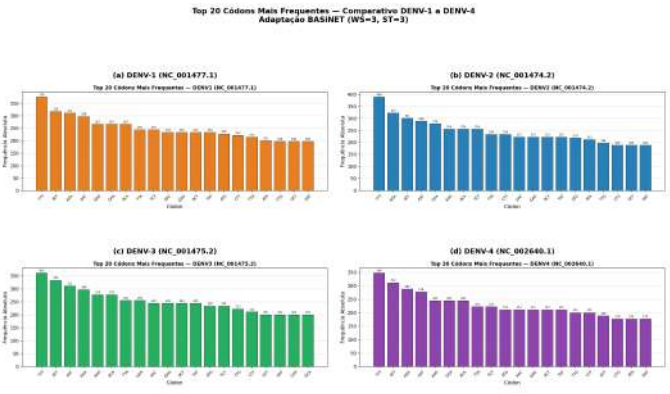


Fig. 1. Top 20 códons mais frequentes nos quatro sorotipos do DENV (WS = ST = 3). Fonte: Os autores (2026).

B. Análise Topológica das Redes de Transição

A modelagem de cada genoma como um grafo direcionado e ponderado permitiu identificar padrões de conectividade funcional por meio de métricas de centralidade [1, 2]. A análise confirmou a existência de *hubs* de conectividade em todas as quatro redes, com estruturas topológicas parcialmente conservadas e parcialmente divergentes entre os sorotipos. A Tabela IV resume as métricas dos principais nós e a Figura 2 apresenta as quatro redes direcionadas.

TABELA IV
Métricas topológicas dos códons hub nos quatro sorotipos do DENV.

Sorotipo	Códon	Métrica	Valor	Freq.
DENV-1	TTT	In-degree	5	376
	ATT	Out-degree	7	318
DENV-2	TTT	In-degree	5	389
	TCT	Out-degree	8	256
DENV-3	TTT	In-degree	5	362
	CTT	Out-degree	7	212
DENV-4	TTT	In-degree	6	348
	ATT	Out-degree	6	312

***Hubs* de Convergência (*In-degree*):** O códon TTT foi identificado como o principal hub de convergência em todos os quatro sorotipos (in-degree entre 5 e 6), consolidando-se como o maior ponto de convergência das redes. Adicionalmente, ATT e ATA figuram como *hubs* de convergência conservados nos quatro sorotipos, com in-degree ≥ 5.

***Hubs* de Divergência (*Out-degree*):** O *hub* de maior out-degree é sorotipo-específico: ATT lidera em DENV-1 e DENV-4 (out-degree = 7 e 6, respectivamente), enquanto TCT é o principal *hub* divergente no DENV-2 (*out-degree* = 8) e CTT no DENV-3 (*out-degree* = 7). Essa variação indica que a diversidade de transições subsequentes é uma propriedade topológica sorotipo-específica, potencialmente associada à divergência filogenética entre os sorotipos [5].

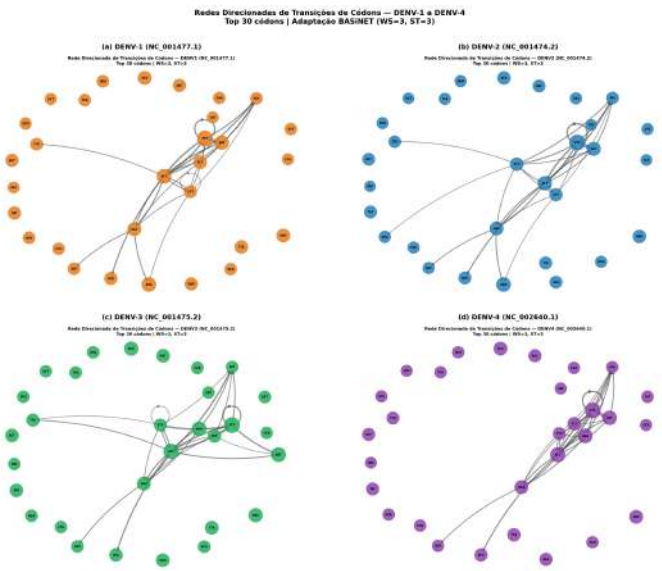


Fig. 2. Redes direcionadas de transições de códons dos quatro sorotipos do DENV. Top 30 códons; espessura das arestas proporcional a $w(c_i, c_{i+1})$. Fonte: Os autores (2026).

C. Padrões Conservados e Divergentes entre Sorotipos

A comparação sistemática das métricas topológicas revelou dois grupos de padrões: (i) elementos estruturais conservados, presentes em todos os quatro sorotipos; e (ii) assinaturas topológicas divergentes, exclusivas de um ou dois sorotipos.

Padrões conservados: Os códons TTT, ATT e ATA são *hubs* de convergência (in-degree ≥ 5) em todos os sorotipos, enquanto ATT, CTT e TCT são *hubs* de divergência (*out-degree* ≥ 5) conservados. Esses elementos representam a estrutura organizacional fundamental das redes de códons do DENV, refletindo restrições funcionais sobre a composição e a ordem de tradução das proteínas virais [28].

Assinaturas divergentes: O DENV-2 distingue-se pelos valores de out-degree mais elevados (TCT = 8), enquanto o DENV-3 é o único sorotipo cujo hub de maior out-degree é CTT. Essas diferenças topológicas são consistentes com a maior distância filogenética do DENV-3 em relação aos demais sorotipos e com as diferenças antigênicas conhecidas [29]. Esses resultados sugerem que as métricas topológicas de grau carregam informação discriminativa entre sorotipos, representando um fundamento relevante para *pipelines* de classificação supervisionada baseados nessas representações [4, 14]. A divergência observada nas métricas de *out-degree* (Tabela IV), onde cada sorotipo apresenta um *hub* de saída exclusivo ou com valor distinto, sustenta a hipótese de capacidade discriminatória, indicando que a topologia da rede é robusta o suficiente para servir como descritor único em modelos de classificação automática.

V. Conclusão

Este trabalho adaptou com sucesso o *framework* BASiNET ao nível de códons, aplicando-o comparativamente aos quatro sorotipos do vírus Dengue. A segmentação não sobreposta ($windowsize = stepsize = 3$) mostrou-se biologicamente coerente ao capturar a unidade funcional da tradução, resultando em redes direcionadas e ponderadas que preservam a ordem sequencial genômica. A análise revelou perfis de frequência específicos por sorotipo e a presença de *hubs* de conectividade parcialmente conservados e divergentes, refletindo a diversidade filogenética do DENV.

A modelagem proposta avança em relação ao BASiNET original [4] ao utilizar grafos direcionados, capturando relações assimétricas essenciais para a representação funcional do genoma. Esta abordagem *alignment-free* constitui uma etapa preliminar tratável para *pipelines* de aprendizagem de máquina baseados em topologia de redes [14, 15].

Apesar dos resultados promissores, o uso exclusivo de sequências de referência do NCBI limita a avaliação da variabilidade intra-sorotipo. Como perspectivas futuras, pretende-se estender a análise a isolados clínicos, incluir métricas avançadas de centralidade e aplicar classificadores supervisionados (SVM, *Random Forest*) para a discriminação automática de variantes. O modelo também poderá ser validado em outros arbovírus, como Zika e Chikungunya, consolidando sua generalidade e potencial para a vigilância epidemiológica.

VI. Agradecimentos

O desenvolvimento deste estudo contou com o apoio do Centro Paula Souza (CPS) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio do Programa Institucional de Bolsas de Iniciação Científica (PIBIC CPS–CNPq). Projeto nº 187093/2025-3 pelo apoio financeiro.

Referências

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez e D.-U. Hwang. “Complex networks: structure and dynamics”. Em: *Physics Reports* 424.4–5 (2006), pp. 175–308. DOI: 10.1016/j.physrep.2005.10.009.

[2] Luciano da F. Costa, Francisco A. Rodrigues, Gonzalo Travieso e Paulo R. Villas Boas. “Characterization of complex networks: a survey of measurements”. Em: *Advances in Physics* 56.1 (2007), pp. 167–242. DOI: 10.1080/00018730601170527.

[3] R. Gomes, A. Silva e L. Santos. “Advanced alignment-free methods for genomic sequence comparison using complex networks”. Em: *Bioinformatics and Computational Biology* 22.1 (2024), pp. 102–118. DOI: 10.1016/j.bcbi.2024.102118.

[4] Eric Augusto Ito, Isaque Katahira, Fábio F. R. Vicente e Fabrício M. Lopes. "BASiNET — BiologicAl Sequences NETwork: a case study on coding and non-coding RNAs identification". Em: *Nucleic Acids Research* 46.16 (2018), e96. DOI: 10.1093/nar/gky462.

[5] Duane J. Gubler. "Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century". Em: *Trends in Microbiology* 10.2 (2002), pp. 100–103. DOI: 10.1016/S0966-842X(01)02288-0.

[6] J. R. Almeida e M. T. Silva. "Genomic surveillance of DENV-1 to DENV-4 using alignment-free methods: Challenges and Perspectives for 2025". Em: *Infection, Genetics and Evolution* 112 (2024), p. 105450.

[7] L. Zhang e H. Wang. "Complex Network Analysis in Virology: A 2024 Perspective on Viral Evolution and Serotype Discrimination". Em: *Journal of Biomedical Informatics* 149 (2024), p. 104562. DOI: 10.1016/j.jbi.2024.104562.

[8] A. Mishra e K. Singh. "Advanced Topometric Characterization of Viral Genomes: A Complex Network Perspective". Em: *Journal of Complex Networks* (2025). Referência ultra-recente sobre métricas topométricas em genomas.

[9] Zhong Wang, Mark Gerstein e Michael Snyder. "RNA-Seq: a revolutionary tool for transcriptomics". Em: *Nature Reviews Genetics* 10.1 (2009), pp. 57–63. DOI: 10.1038/nrg2484.

[10] Orly Wapinski e Howard Y. Chang. "Long noncoding RNAs and human disease". Em: *Trends in Cell Biology* 21.6 (2011), pp. 354–361. DOI: 10.1016/j.tcb.2011.04.001.

[11] Yu-Jian Kang, De-Chang Yang, Lei Kong, Mei Hou, Yu-Qi Meng, Liping Wei e Ge Gao. "CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features". Em: *Nucleic Acids Research* 45 (2017), W12–W16. DOI: 10.1093/nar/gkx428.

[12] Liang Sun, Haitao Luo, Dechao Bu, Guoguang Zhao, Kuntao Yu, Changhai Zhang, Yuanning Liu, Runsheng Chen e Yi Zhao. "Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts". Em: *Nucleic Acids Research* 41.17 (2013), e166. DOI: 10.1093/nar/gkt646.

[13] Junliang Li, Yufen Pu, Jijun Tang, Quan Zou e Fei Guo. "The computational approaches of lncRNA identification based on coding potential: status quo and challenges". Em: *Computational and Structural Biotechnology Journal* 18 (2020), pp. 3666–3677. DOI: 10.1016/j.csbj.2020.11.030.

[14] Deepak Singh et al. "A large-scale benchmark study of tools for the classification of protein-coding and non-coding RNAs". Em: *NAR Genomics and Bioinformatics* (2022). DOI: 10.1093/nargab/lqac082.

[15] Robson P. Bonidia, Douglas S. Domingues, Danilo S. Sanches e André C. P. L. F. de Carvalho. "MathFeature: feature extraction package for DNA, RNA and protein sequences based on mathematical descriptors". Em: *Briefings in Bioinformatics* 23.1 (2022). DOI: 10.1093/bib/bbab434.

[16] Thales A. R. Ramos et al. "RNAmining: a machine learning stand-alone and web server tool for RNA coding potential prediction". Em: *Computers in Biology and Medicine* (2023). DOI: 10.1016/j.compbiomed.2023.106739.

[17] Albert-László Barabási e Zoltán N. Oltvai. "Network biology: understanding the cell's functional organization". Em: *Nature Reviews Genetics* 5.2 (2004), pp. 101–113. DOI: 10.1038/nrg1272.

[18] Réka Albert e Albert-László Barabási. "Statistical mechanics of complex networks". Em: *Reviews of Modern Physics* 74.1 (2002), pp. 47–97. DOI: 10.1103/RevModPhys.74.47.

[19] Albert-László Barabási, Natali Gulbahce e Joseph Loscalzo. "Network medicine: a network-based approach to human disease". Em: *Nature Reviews Genetics* 12.1 (2011), pp. 56–68. DOI: 10.1038/nrg2918.

[20] Murilo M. Breve et al. "BASiNETEntropy: an alignment-free method for classification of biological sequences using complex networks and entropy". Em: *Entropy* (2022). DOI: 10.3390/e24050638.

[21] Rafael F. T. Melo. "Codancirc: ferramenta para predição de regiões codificadoras em RNA circular". Diss. de mestr. Universidade Tecnológica Federal do Paraná, 2022.

[22] Marcos H. Pimenta-Zanon et al. "A simplified complex network-based approach to mRNA and ncRNA classification". Em: *Briefings in Bioinformatics* (2023). DOI: 10.1093/bib/bbad110.

[23] Robson P. Bonidia et al. "Selecting the most relevant features for the identification of coding and non-coding RNAs". Em: *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2021). DOI: 10.1109/TCBB.2021.3078133.

[24] Halise Busra Cagirici et al. "LncMachine: a machine learning algorithm for long noncoding RNA annotation in plants". Em: *Plant Molecular Biology* (2022). DOI: 10.1007/s11103-022-01251-2.

[25] Aric A. Hagberg, Daniel A. Schult e Pieter J. Swart. *Exploring network structure, dynamics, and function using NetworkX*. Proceedings of the 7th Python in Science Conference (SciPy2008). 2008.

[26] Python Software Foundation. *Python Language Reference, version 3.12*. https://www.python.org. 2024.

[27] Aimin Li, Junying Zhang e Zhongyin Zhou. "PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme". Em: *BMC Bioinformatics* 15.1 (2014), p. 311. DOI: 10.1186/1471-2105-15-311.

[28] Scott B. Halstead. "Dengue". Em: *The Lancet* 370.9599 (2007), pp. 1644–1652. DOI: 10.1016/S0140-6736(07)61687-0.

[29] Samir Bhatt, Peter W. Gething, Oliver J. Brady et al. "The global distribution and burden of dengue". Em:

Nature 496.7446 (2013), pp. 504–507. DOI: 10.1038/nature12060.

AUTORES



Kelven Bertanha do Nascimento é técnico em Desenvolvimento de Sistemas (Etec Jacinto Ferreira de Sá) e graduando em Análise e Desenvolvimento de Sistemas (Fatec Ourinhos). Atua na modelagem de soluções computacionais e engenharia de software, com foco em projetos práticos voltados ao desenvolvimento de sistemas de redes complexas.



Lavínia Maria Barbosa de Miranda é graduanda em Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Possui interesse em modelagem de soluções computacionais e boas práticas de programação, buscando aplicar conceitos de engenharia de software no desenvolvimento de aplicações e sistemas.



Jéssica Antonio Delgado é Doutora (2021) e Mestre (2016) em Engenharia Elétrica pela Universidade Estadual Paulista (UNESP) e graduada em Matemática. Possui sólida experiência na área de Matemática Aplicada, com ênfase em otimização não linear, métodos de barreira e fluxo de potência ótimo. Atualmente, exerce a docência na Faculdade de Tecnologia de Ourinhos (Fatec Ourinhos) e no Centro Universitário de Faculdades Integradas de Ourinhos (UNIFIO), atuando em projetos de modelagem e análise quantitativa.



Isaque Katahira é docente do Centro Estadual de Educação Tecnológica Paula Souza, atuando na Fatec Ourinhos e na Etec Professor Mário Antônio Verza. É Doutor em Ciência da Informação pela UNESP (2022) e Mestre em Bioinformática pela UTFPR (2018). Atua como revisor da Revista Eletrônica e-F@tec e é membro ativo do Grupo de Pesquisa Representação Temática da Informação, com foco em análise de redes e sistemas de organização do conhecimento.



Fábio Nogueira de Queiroz é Mestre em Ensino pela UENP e possui graduações em Processamento de Dados, Informática, Matemática e Pedagogia. É especialista em MBA Informática e Docência do Ensino Superior. Atualmente, é docente na Fatec Ourinhos e coordenador de curso na Etec Jacinto Ferreira de Sá, com sólida experiência em desenvolvimento de sistemas e práticas pedagógicas na área de tecnologia.

Organização da Competição Nacional Biochallenge Brasil 2025

Felipe Fonseca Vidal Prado, Stefanie Barbosa Ancelmo, Rita Elizabeth Santos de Almeida,
Luma Rissatti Borges do Prado, Elisa Renno Carneiro Dester
Instituto Nacional de Telecomunicações - Inatel
fonseca.felipe@gec.inatel.br, s.barbosa@geb.inatel.br, r.elizabeth@inatel.br, luma.prado@inatel.br, elisa.renno@inatel.br

***Abstract*—Population aging poses challenges related to autonomy and quality of life, requiring innovative technological solutions. In this context, Biomedical Engineering and Assistive Technology play a strategic role in the development of devices and systems aimed at promoting health and functional independence. This work presents an experience report on the structuring and execution of Biochallenge Brasil 2025, a national competition focused on developing technological solutions for real challenges in the healthcare field. The methodology was organized into sequential stages: Track, Regionals, and National Final, involving training, specialized mentoring, prototype development, and evaluation by a technical committee. The results showed broad student participation and regional diversity.**

***Index Terms*—Biomedical Engineering; Assistive Technology; Academic Competition; Innovation; Engineering Education.**

***Resumo*—O envelhecimento populacional impõe desafios relacionados à autonomia e qualidade de vida da pessoa idosa, demandando soluções tecnológicas inovadoras. Nesse contexto, a Engenharia Biomédica e a Tecnologia Assistiva assumem papel estratégico na criação de dispositivos e sistemas voltados à promoção da saúde e independência funcional. Este trabalho apresenta um relato de experiência sobre a estruturação e execução do Biochallenge Brasil 2025, uma competição nacional voltada ao desenvolvimento de soluções tecnológicas para desafios reais da área da saúde. A metodologia foi organizada em etapas sequenciais: Trilha, Regionais e Final Nacional, envolvendo capacitação, mentorias especializadas, desenvolvimento de protótipos e avaliação por banca técnica. Os resultados evidenciaram ampla participação estudantil e diversidade regional.**

***Palavras chave*—Engenharia Biomédica; Tecnologia Assistiva; Competição Acadêmica; Inovação; Educação em Engenharia.**

I. Introdução

O envelhecimento pode ser compreendido como um processo complexo que envolve mudanças nas dimensões cronológica, biológica, psicológica e social da vida humana. Embora a idade cronológica seja frequentemente utilizada como referência para definir a velhice, ela representa apenas o tempo decorrido desde o nascimento e não expressa, por si só, as transformações que caracterizam esse processo. O envelhecimento também envolve alterações biológicas, mudanças nas capacidades psicológicas e cognitivas, além de transformações nos papéis sociais desempenhados pelos indivíduos, sendo influenciado por fatores individuais, culturais e históricos[1]

Segundo a revisão de 2018 da Projeção de População do IBGE, em 2060 o percentual da população com 65 anos ou mais chegará a 25,5% (58,2 milhões de idosos) [2]. Veras (2009) afirma que o aumento do número de idosos e da utilização do sistema de saúde está associado à maior prevalência de doenças crônicas e de longa duração, como hipertensão arterial, diabetes *mellitus*, osteoporose, artrite, problemas cardíacos e doenças cerebrovasculares, que frequentemente exigem acompanhamento contínuo. Esse cenário configura um importante desafio para o sistema de saúde. [3]. Diante desse contexto, a Tecnologia Assistiva (TA) contribui para a funcionalidade, mobilidade e realização das atividades do cotidiano, promovendo autonomia, independência e qualidade de vida para a pessoa idosa. Assim, esses recursos favorecem o envelhecimento ativo e saudável, além de possibilitar maior participação social e melhor desempenho nas atividades de vida diária.[4]

Para o desenvolvimento de uma TA faz-se necessário que os profissionais envolvidos sejam capazes de compreender as demandas clínicas, funcionais e sociais do envelhecimento e traduzi-las em tecnologias seguras, eficazes e aplicáveis à realidade da população. Nesse sentido, a Engenharia Biomédica assume papel estratégico ao atuar no desenvolvimento e aprimoramento de dispositivos, sistemas e métodos voltados à área da saúde.[5]

O curso de Engenharia Biomédica caracteriza-se por uma formação interdisciplinar que articula fundamentos das ciências exatas e das ciências da saúde, integrando conhecimentos biológicos e clínicos a tecnologias. Ele possibilita o desenvolvimento de soluções para diagnóstico, monitoramento, tratamento e reabilitação, contribuindo para o avanço de tecnologias inovadoras e seguras na área da saúde [6].

No Brasil, a graduação em Engenharia Biomédica foi implementada pela Universidade do Vale do Paraíba (UNIVAP), localizada em São José dos Campos, SP, com início das atividades em 2001 e formação da primeira turma em 2005 [7]. Embora o curso já esteja consolidado em diversos países, no Brasil ainda há desafios persistentes em alinhar a formação acadêmica às demandas de mercado e de pesquisa, bem como em fortalecer o reconhecimento institucional e profissional da área no país [8]. Segundo a Sociedade Brasileira de Engenharia Biomédica (SBEB), outra dificuldade apontada é a desistência dos alunos, com índices que chegam entre 50% e 60% nas últimas turmas [9].

Tongjie (2023) afirma que uma das formas de reduzir a evasão dos estudantes é aproximar os alunos de desafios reais, como por exemplo as competições acadêmicas.[10]Almeida (2022) corrobora com esta afirmação em seus estudo enfatizando que metodologias de aprendizagem baseadas em experiências práticas, como jogos educacionais e simulações, podem contribuir para o desenvolvimento de habilidades e

competências relevantes à formação do engenheiro, além de favorecer uma aprendizagem mais ativa e interdisciplinar. [11, 12]. Além disso, estudos sobre olimpíadas científicas indicam que essas iniciativas contribuem para o desenvolvimento de diferentes habilidades entre os estudantes, como criatividade, raciocínio, ampliação do conhecimento e trabalho em equipe. [13].

Atualmente, há competições acadêmicas internacionais na área, como o *Design by Biomedical Undergraduate Teams (DEBUT)* competição norte americana que esta voltada a equipes de graduação em Engenharia Biomédica, promovida pelo *National Institute of Biomedical Imaging and Bioengineering (NIBIB)* em parceria com a *VentureWell*. A iniciativa reconhece projetos inovadores que buscam solucionar problemas clínicos reais, avaliando critérios como relevância do problema abordado, impacto na prática em saúde, inovação no *design* e desenvolvimento de protótipos funcionais.[14] Também há a *BME-idea*, uma competição internacional criada em 2005 que reconhece projetos inovadores de Engenharia Biomédica com potencial de impacto social e comercial. A competição reúne equipes multidisciplinares de estudantes de graduação e pós-graduação que desenvolvem dispositivos, produtos ou tecnologias voltadas à solução de problemas clínicos, como dispositivos médicos, tecnologia assistiva, sistemas diagnósticos ou terapêuticos, incentivando a inovação e a transferência de tecnologia para a área da saúde. [15].

Nacionalmente, existe o Prêmio de Inovação em Engenharia Biomédica para o Sistema Único de Saúde (SUS) é uma iniciativa promovida pela Sociedade Brasileira de Engenharia Biomédica (SBEB) e a empresa *Boston Scientific*, que reconhece projetos tecnológicos com potencial de aplicação no SUS. O prêmio busca incentivar o desenvolvimento de soluções inovadoras voltadas à saúde pública brasileira, aproximando academia, indústria e governo e estimulando pesquisas e tecnologias capazes de melhorar o diagnóstico, o tratamento e a gestão em saúde. [16].

Apesar da existência de algumas competições, observa-se ainda, uma lacuna significativa no número de atividades acadêmicas contínuas e acessíveis na área. A escassez de competições e projetos extracurriculares pode contribuir para a desmotivação dos estudantes e dificultar o desenvolvimento de competências essenciais.[17]

Diante deste contexto, o Instituto Nacional de Telecomunicações (Inatel) em 2018 criou o *Biochallenge*, promovida durante a Semana de Engenharia Biomédica, com foco no desenvolvimento de soluções tecnológicas voltadas à área da saúde [18]. Nas edições seguintes, o evento passou por diferentes formatos e temáticas, incluindo desafios no modelo de hackathon e atividades realizadas de forma remota durante o período da pandemia de Covid-19 [19].

A partir de 2023, a competição ampliou sua abrangência para estudantes de diversas instituições de ensino superior do país, passando a ser denominada Biochallenge Brasil [20]. Em 2024, o evento consolidou seu caráter nacional com a realização de etapas presenciais classificatórias em centros de Engenharia Biomédica distribuídos pelas cinco regiões do país, culminando na final nacional realizada durante o Congresso Brasileiro de Engenharia Biomédica (CBEB) no mesmo ano.[21]

O presente estudo teve como objetivo planejar, organizar e executar a edição de 2025 do evento Biochallenge Brasil, cujo tema foi "Tecnologia Assistiva aplicada à Longevidade".

II. Metodologia

Este trabalho caracterizou-se como um relato de experiência de natureza descritiva, fundamentado na participação ativa dos autores na estruturação, organização e execução da competição Biochallenge Brasil 2025, abrangendo desde a análise de edições anteriores, definição do tema, planejamento das etapas, definição dos desafios, suporte técnico às equipes participantes e a classificação dos vencedores da competição.

A. Definição do tema e dos desafios da competição

A definição dos desafios propostos aos participantes foi estruturada a partir do tema central da competição "Tecnologia Assistiva voltada à Longevidade", orientando o desenvolvimento de soluções para demandas reais do cotidiano de pessoas idosas. Para isso, foi estabelecido dois estudos de caso distintos, vinculados às categorias de participação (presencial e *online*). Cada caso era composto por perfis detalhados de usuários fictícios com necessidades específicas de saúde, autonomia, segurança e interação social.

B. Contextualização de cada categoria

Na categoria presencial, os participantes foram desafiados a propor uma solução em *hardware* destinada a auxiliar o casal idoso composto pela Sra. Dolores (75 anos) e pelo Sr. Antônio (80 anos). Entre as demandas apresentadas estavam o esquecimento do fogão e do gás de cozinha ligados, dificuldades na gestão medicamentosa, risco de quedas durante o período noturno, limitações auditivas e dificuldades na utilização de dispositivos eletrônicos, como o controle remoto da televisão. Essas situações representam desafios recorrentes no processo de envelhecimento e exigem intervenções voltadas à promoção da segurança, autonomia e organização da rotina diária.

Na categoria *online*, o desafio foi desenvolver uma solução em *software* voltada ao apoio de Dona Benedita (73 anos), diagnosticada com diabetes *mellitus*, tipo 2. As demandas incluíam dificuldades na adesão à dieta e ao uso correto da medicação (insulina), esquecimento na aferição da glicemia, baixa regularidade na prática de atividades físicas, redução da interação social e insegurança decorrentes da incontinência urinária. Tais aspectos refletiam desafios frequentes no envelhecimento associados às doenças crônicas e às mudanças sóciocomportamentais.

As propostas buscaram estimular o desenvolvimento de ferramentas tecnológicas capazes de incentivar hábitos saudáveis, promover maior autonomia no autocuidado e favorecer o engajamento social, integrando conhecimentos técnicos da Engenharia Biomédica e de áreas correlatas no desenvolvimento das soluções.

C. Estratégias de divulgação e recrutamento

Para promover a competição e ampliar o alcance do evento, foram realizadas as seguintes ações de comunicação e *marketing*.

Inicialmente, foram estabelecidos contatos diretos com universidades de diversas regiões do país que possuíam cursos relacionados às áreas da saúde, engenharia e ciências exatas, por meio de ligações telefônicas e envio de e-mails institucionais. Paralelamente, foram utilizadas redes sociais, como *Instagram* e *LinkedIn*, como ferramentas de divulgação e engajamento do público-alvo, visando alcançar estudantes interessados em inovação e desenvolvimento tecnológico aplicado à saúde.

D. Estrutura organizacional da competição

A metodologia do Biochallenge Brasil 2025 foi estruturada em fases sequenciais e complementares, conduzindo os participantes desde a etapa de capacitação até a validação final das soluções desenvolvidas. A competição foi organizada em três etapas principais: Trilha, Regionais e Final Nacional, distribuídas ao longo de um cronograma previamente estabelecido.

E. Trilha

Esta etapa ocorreu entre os meses de maio e julho e correspondeu à fase de capacitação e ideação dos projetos. O processo foi realizado integralmente de forma *online*, por meio da plataforma *Discord*, ambiente no qual as equipes participaram das capacitações e, paralelamente, conduziram a ideação das soluções, a definição do escopo dos projetos e o início do desenvolvimento dos protótipos.

Durante essa fase, os participantes tiveram acesso a capacitações relacionadas à Tecnologia Assistiva voltada à longevidade, saúde e desafios do envelhecimento, ministradas por especialistas da área da saúde. Também foram oferecidos conteúdos sobre *Design Thinking* e desenvolvimento de projetos.

Concomitante a isso, foram disponibilizados minicursos previamente gravados abordando temas como sensores, atuadores, desenvolvimento de jogos e aplicativos, além de ferramentas como Arduino e Tinkercad, voltados ao desenvolvimento das soluções propostas.

Além das atividades formativas, as equipes contaram com momentos de mentoria e acompanhamento, realizados mediante agendamento prévio na própria plataforma. Esses encontros tiveram como objetivo orientar o desenvolvimento das propostas, abordando aspectos relacionados à engenharia biomédica, tecnologia assistiva, inovação e desenvolvimento de protótipos. Ao longo dessa etapa, as equipes também realizaram entregas parciais de seus projetos, que serviram como base para a avaliação nas fases subsequentes da competição.

F. Regionais

Ocorrida entre os meses de julho e agosto, conforme o cronograma da competição. Nessa fase, as equipes apresentaram os protótipos desenvolvidos durante a etapa anterior para uma banca avaliadora composta por profissionais da área.

As apresentações foram realizadas presencialmente para as equipes da categoria presencial e remotamente para as equipes da categoria online. As soluções foram avaliadas com base em critérios técnicos, funcionais e de inovação, considerando as especificidades de cada categoria.

Com o objetivo de ampliar o alcance da competição e democratizar o acesso dos estudantes, as etapas regionais presenciais foram realizadas em diferentes regiões do Brasil. A seleção das instituições foi feita de acordo com parcerias prévias com a instituição proponente da competição (INATEL) e pelo fato de as instituições possuírem o curso de Engenharia Biomédica, sendo elas: Universidade Federal do Pará (UFPA), em Belém (PA), representando a região Norte; Universidade de Brasília (UnB), em Brasília (DF), pela região Centro-Oeste; Instituto Nacional de Telecomunicações (INATEL), em Santa Rita do Sapucaí (MG), na região Sudeste; Universidade Federal de Pernambuco (UFPE), em Recife (PE), representando o Nordeste; e Pontifícia Universidade Católica do Paraná (PUCPR), em Curitiba (PR), na região Sul.

Em cada instituição selecionada ocorreram as competições Regionais Presenciais do Biochallenge Brasil 2025, na qual os participantes tiveram acesso ao regulamento oficial da competição, contendo a descrição detalhada das etapas, critérios de avaliação e objetivos do evento. Esse formato possibilitou maior integração entre os estudantes, mentores e organizadores, favorecendo o aprendizado colaborativo e multidisciplinar.

G. Final nacional

As equipes com melhor desempenho nas etapas regionais foram classificadas para a Final Nacional que ocorreu no INATEL no dia 23 de agosto de 2025. Nessa fase, os finalistas apresentaram as versões finais de suas soluções para uma banca avaliadora responsável pela definição das equipes vencedoras da competição.

III. Resultados e Discussão

O Biochallenge Brasil 2025 contou com um total de 247 inscritos, sendo evidenciado uma predominância da modalidade presencial, que concentrou 70% do total (173 participantes), enquanto a participação *online* correspondeu a 30% (74 participantes), como pode ser visto na figura 1.



Fig. 1. Distribuição de participantes por categoria

Em relação à distribuição regional, observou-se maior concentração de participantes no Sudeste, especialmente na modalidade presencial, que registrou 95 inscritos. Na mesma região, a participação *online* foi significativamente menor (11 participantes), indicando maior alcance do formato presencial nesse contexto. O Nordeste apresentou comportamento diferente, com maior número de participantes na modalidade *online*

(52) em comparação ao presencial (37), podendo ser visto na figura 2. Evidenciando que o formato remoto ampliou o acesso e favoreceu a participação de equipes localizadas em outras regiões do país.

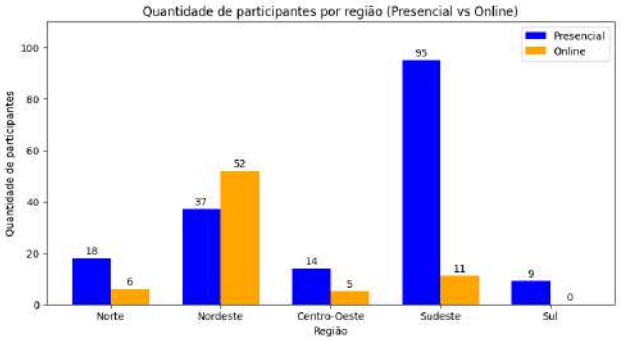


Fig. 2. Distribuição de participantes por região, segundo a modalidade presencial e *online*

E como pode ser observado no perfil dos participantes referente à distribuição dos cursos de origem dos inscritos, apresentada na figura 3. Observa-se predominância justamente dos cursos da área de engenharia, que juntos representam a maior parcela dos participantes. Destaca-se a Engenharia Biomédica, responsável por aproximadamente 24,8% das inscrições, seguida pela Engenharia de Computação (17,4%) e por outras engenharias (19,6%). Também foram identificadas participações provenientes de Engenharia Elétrica (8,7%), Engenharia Eletrônica (7,4%) e Engenharia Mecânica (3,0%).

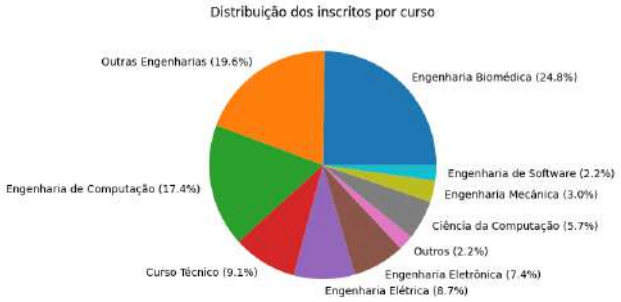


Fig. 3. Distribuição de inscritos por curso

Cursos da área de computação também tiveram participação relevante, como Ciência da Computação (5,7%) e Engenharia de Software (2,2%). Além disso, foram observadas inscrições provenientes de cursos técnicos (9,1%), como o curso técnico em eletrônica, e de outras áreas acadêmicas (2,2%) como design, indicando um grau de diversidade interdisciplinar entre os participantes.

A predominância de participantes do Sudeste na modalidade presencial pode ser explicada por fatores estruturais e logísticos. De acordo com levantamento da SBEB, das 28 instituições brasileiras que oferecem graduação em Engenharia Biomédica, 16 estão localizadas na região Sudeste, o que contribui para a formação de maior número de estudantes na área[22]. Além disso, a realização da etapa final da competição no INATEL, em Santa Rita do Sapucaí (MG), pode ter favorecido a participação presencial de equipes dessa região, uma vez que a proximidade geográfica reduz custos e dificuldades relacionadas ao deslocamento, hospedagem e tempo de viagem.

De forma geral, os dados mostram que, apesar da predominância do modelo presencial, a modalidade *online* contribuiu para ampliar o alcance da competição em nível nacional, fortalecendo a presença de diferentes regiões e promovendo maior diversidade entre os participantes.

Com relação às estratégias de divulgação adotadas, destacou-se a utilização de depoimentos de dois participantes de edições anteriores do BioChallenge Brasil, os quais relataram suas experiências e os impactos positivos da competição em seu desenvolvimento acadêmico e profissional. Esses relatos contribuíram para evidenciar, junto ao público-alvo, as vantagens de integrar a iniciativa, ressaltando não apenas o aprimoramento de competências técnicas e habilidades interpessoais, mas também o potencial de geração de impacto social por meio das soluções propostas.

Ao longo do processo de divulgação, foram identificadas algumas dificuldades relacionadas às estratégias de *marketing*, especialmente no que se refere ao contato com universidades de diferentes regiões do país que oferecem cursos nas áreas de exatas e da saúde. A amplitude territorial e a diversidade institucional representaram desafios para o alcance efetivo do público-alvo, exigindo a adaptação constante dos canais de comunicação e das abordagens utilizadas. Em alguns casos, houve limitações na obtenção de respostas ou retorno institucional, o que impactou o ritmo de disseminação das informações sobre a competição. Ainda assim, esses obstáculos contribuíram para o aprimoramento das estratégias adotadas, permitindo ajustes nas formas de divulgação e fortalecendo o planejamento de ações futuras, com vistas a ampliar o alcance e a participação em próximas edições.

A realização das mentorias ao longo da competição mostrou-se um elemento relevante para o desenvolvimento das soluções propostas pelas equipes. Durante a fase de capacitação, os encontros *online* com mentores especializados possibilitaram a troca de conhecimentos técnicos, esclarecimento de dúvidas e orientações voltadas à estruturação dos projetos. Observou-se que esse acompanhamento contribuiu para maior organização no planejamento das entregas e no amadurecimento das ideias apresentadas. Além disso, o contato direto com profissionais experientes favoreceu o fortalecimento da confiança das equipes no processo de desenvolvimento das soluções, refletindo positivamente na qualidade técnica e na consistência das propostas elaboradas. Dessa forma, as mentorias desempenharam papel importante no suporte formativo aos participantes, contribuindo para o aprimoramento dos projetos e para o fortalecimento do potencial de inovação das soluções desenvolvidas.

Na etapa final da competição, as equipes classificadas apresentaram um *pitch* de suas soluções, destacando aspectos relacionados à inovação, impacto social, viabilidade técnica e funcionalidade dos protótipos desenvolvidos. Esse momento representou a consolidação do processo formativo promovido ao longo da competição, permitindo que os participantes demonstrassem os resultados alcançados e justificassem suas escolhas técnicas diante da banca avaliadora. A apresentação pública das soluções também evidenciou o desenvolvimento de habilidades importantes para a formação acadêmica e profissional, como comunicação científica, argumentação e capacidade de síntese.

Assim, a etapa final evidenciou a evolução das equipes ao longo da competição e permitiu avaliar de forma integrada os resultados obtidos durante todo o processo de desenvolvimento dos projetos.

Por fim, a etapa da Final Nacional foi realizada no Inatel, ocasião em que foram definidas as equipes vencedoras da competição. Na categoria presencial, a equipe *Intelimed*, da regional Centro-Oeste, obteve a primeira colocação, seguida pela equipe *LIHOM*, da regional Nordeste, em segundo lugar, e pela equipe *Kairos*, da regional Sul, em terceiro lugar.

Na categoria *online*, a equipe *Biominds*, da regional Centro-Oeste, conquistou o primeiro lugar, enquanto a equipe *Biobytes*, da regional Nordeste, alcançou a segunda colocação, e a equipe *Biodata IV* ficou em terceiro lugar. A classificação final das equipes foi estabelecida com base em uma avaliação sistemática conduzida por especialistas nas áreas da Saúde e Ciências Exatas, considerando critérios técnicos previamente definidos pela organização da competição.

IV. Conclusão

O presente trabalho evidenciou que a promoção da Tecnologia Assistiva voltada à longevidade exige não apenas avanços tecnológicos, mas também estratégias educacionais capazes de formar profissionais preparados para enfrentar desafios reais e complexos da área da saúde. Ao discutir o contexto do envelhecimento populacional, a importância da Engenharia Biomédica e os desafios de permanência discente no Brasil, reforça-se que a qualificação técnica, quando associada a experiências práticas e contextualizadas, pode contribuir de maneira significativa para a consolidação da área no país.

A experiência do Biochallenge Brasil 2025 demonstrou que competições acadêmicas estruturadas, com desafios alinhados a demandas sociais concretas, podem funcionar como ferramentas formativas estratégicas. Ao integrar capacitação técnica, mentorias especializadas, desenvolvimento de protótipos e apresentação final para banca avaliadora, o evento favoreceu o desenvolvimento de competências técnicas e transversais, como trabalho em equipe, pensamento crítico, comunicação científica e inovação. Além disso, a descentralização regional e a oferta de modalidade *online* ampliaram o acesso e contribuíram para a democratização da participação estudantil.

Os resultados observados indicam que iniciativas dessa natureza podem atuar como instrumentos de fortalecimento da identidade profissional do engenheiro biomédico, aproximando teoria e prática e auxiliando na redução da evasão acadêmica ao tornar mais tangível a aplicação dos conhecimentos adquiridos ao longo da graduação. A articulação entre instituições de ensino, entidades científicas e organizações apoiadoras mostrou-se fundamental para a viabilidade e alcance da proposta.

Como perspectivas futuras, sugere-se a ampliação das parcerias institucionais, o acompanhamento longitudinal dos participantes para avaliação de impacto acadêmico e profissional, bem como a incorporação de métricas mais sistemáticas de avaliação de aprendizado e inovação gerada. Recomenda-se ainda a expansão do modelo para outras áreas interdisciplinares da engenharia e da saúde, fortalecendo uma cultura de formação baseada em desafios reais e impacto social.

Assim, conclui-se que o Biochallenge Brasil representa não apenas uma competição acadêmica, mas uma estratégia educacional com potencial de contribuir para o avanço da Engenharia Biomédica no Brasil, estimulando inovação responsável, engajamento estudantil e desenvolvimento de soluções alinhadas às necessidades da sociedade contemporânea.

V. Agradecimentos

Gostaríamos de expressar nossa gratidão à Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG, ao Instituto Nacional de Telecomunicações - INATEL, e ao eHealth Innovation Center pelo apoio técnico e recursos viabilizados para o desenvolvimento deste projeto.

Referências

[1] Rodolfo Herberto SCHNEIDER e Tatiana Quarti IRIGARAY. "O envelhecimento na atualidade: aspectos cronológicos, biológicos, psicológicos e sociais". Em: *Estudos de Psicologia (Campinas)* 25.4 (2008), pp. 585–593. DOI: 10.1590/S0103-166X2008000400013.

[2] INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Projeção da população 2018: número de habitantes do país deve parar de crescer em 2047*. Agência de Notícias IBGE. Disponível em: https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/21837-projecao-da-populacao-2018-numero-de-habitantes-do-pais-deve-parar-de-crescer-em-2047. Jul. de 2018.

[3] Renato VERAS. "Envelhecimento populacional contemporâneo: demandas, desafios e inovações". Em: *Revista de Saúde Pública* 43.3 (2009), pp. 548–554.

[4] Amanda Gabrielle do Vale Neves MACHADO et al. "O uso da tecnologia assistiva como facilitador do desempenho ocupacional de idosos". Em: *Pesquisas e Inovações em Ciências da Saúde e Biológicas*. São Paulo: Atena Editora, 2022.

[5] Universidade Federal de Pernambuco (UFPE). *Engenharia Biomédica ganha primeiro laboratório*. Disponível em: https://www.ufpe.br/agencia/index.php?option=com_content&view=article&id=38638. 2019.

[6] Rohini NANTHAKUMAR e Nivethika SIVAKUMARAN. "Role of Biomedical Engineering for Diagnose and Treatment". Em: *International Journal of Advances in Scientific Research and Engineering* 4.11 (2018). DOI: 10.31695/IJASRE.2018.32944.

[7] Maria Angela Boccara de Paula, Alessandra de Cássia Grilo e Maria Auxiliadora Ávila. "Engenharia biomédica: escolhas profissionais nas trajetórias docentes". Em: *Revista Internacional de Educação Superior* 9.00 (2022), e023013. DOI: 10.20396/riesup.v9i00.8660655. URL: https://periodicos.sbu.unicamp.br/ojs/index.php/riesup/article/view/8660655.

[8] Sociedade Brasileira de Engenharia Biomédica (SBEB). *Engenharia Biomédica no Brasil: ciência, mercado e o desafio da autossuficiência tecnológica*. Disponível em: https://sbeb.org.br/engenharia-biomedica-no-brasil-ciencia-mercado-e-o-desafio-da-autossuficiencia-tecnologica/. 2025.

[9] Sociedade Brasileira de Engenharia Biomédica (SBEB). *Por que é tão alta a evasão nos cursos de Engenharia Biomédica?* Disponível em: https://sbeb.org.br/por-que-e-tao-alta-a-evasao-nos-cursos-de-engenharia-biomedica/.

[10] Tongjie JU e Jiabin ZHU. "Exploring senior engineering students' engineering identity". Em: *International Journal of STEM Education* (2023).

[11] Fernanda Gobbi de Boer Garbin e Adriana Justin Cerveira Kampff. "O ensino e a aprendizagem de competências nos cursos superiores de engenharia: focos das pesquisas entre os anos 2001 e 2020". Em: *Educitec - Revista de Estudos e Pesquisas sobre Ensino Tecnológico* 7 (2021), e143021. DOI: 10.31417/educitec.v7.1430. URL: https://sistemascmc.ifam.edu.br/educitec/index.php/educitec/article/view/1430.

[12] M. N. de Moraes e P. A. Cardoso. "Jogos para ensino em Engenharia e desenvolvimento de habilidades". Em: *Revista Principia* 1.39 (2018), pp. 19–29. DOI: 10.18265/1517-03062015v1n39p19-29. URL: https://periodicos.ifpb.edu.br/index.php/principia/article/view/1769.

[13] Thiago Oliveira ROCHA, Edson Viana CARVALHO, Lara de Oliveira CARVALHO, Caio Eduardo Silva AMARAL e Polyane Alves SANTOS. "As olimpíadas científicas no desenvolvimento da educação brasileira". Em: *Congresso Nacional de Educação (CONEDU)*. 2016.

[14] VENTUREWELL. *DEBUT Challenge*. https://venturewell.org/debut/.

[15] VENTUREWELL. *BMEidea Competition*. https://venturewell.org/bmeidea/.

[16] SOCIEDADE BRASILEIRA DE ENGENHARIA BIOMÉDICA (SBEB). *Prêmio de Inovação em Engenharia Biomédica para o SUS 2025 está com inscrições abertas*. https://sbeb.org.br/premio-de-inovacao-em-engenharia-biomedica-para-o-sus-2025-esta-com-inscricoes-abertas/.

[17] Eugenio BRAVO et al. "Extracurricular student teams: learning gains in interdisciplinary projects". Em: *European Journal of Engineering Education* (2025).

[18] INSTITUTO NACIONAL DE TELECOMUNICAÇÕES (INATEL). *Projetos de inovação tecnológica são premiados na 6ª Semana da Engenharia Biomédica*. Disponível em: https://inatel.br/noticias/projetos-de-inovacao-tecnologica-sao-premiados-na-6-semana-da-engenharia-biomedica. Acesso em: 5 mar. 2026.

[19] INSTITUTO NACIONAL DE TELECOMUNICAÇÕES (INATEL). *Alunos do Inatel vencem a 2ª edição do Bio Challenge*. Disponível em: https://inatel.br/noticias/alunos-do-inatel-vencem-a-2-edicao-do-bio-challenge. Acesso em: 5 mar. 2026.

[20] INSTITUTO NACIONAL DE TELECOMUNICAÇÕES (INATEL). *1ª competição nacional de Engenharia Biomédica é do Inatel*. Disponível em: https://inatel.br/noticias/1-competicao-nacional-de-engenharia-biomedica-e-do-inatel. Acesso em: 5 mar. 2026.

[21] Sympla. *Biochallenge Brasil 2024 – Competição Nacional de Engenharia Biomédica*. Disponível em: https://www.sympla.com.br/evento/biochallenge-brasil-2024-competicao-nacional-de-engenharia-biomedica-inatel/2419685. Acesso em: 5 mar. 2026.

[22] Sociedade Brasileira de Engenharia Biomédica (SBEB). *Centros de Engenharia Biomédica*. Acesso em: 15 mar. 2026. 2024. URL: https://share.google/Mwz6tMz2xt5LauNI1.

AUTORES



Felipe Fonseca Vidal Prado Graduando em Engenharia da Computação pelo Instituto Nacional de Telecomunicações (INATEL). Atua como estagiário em Redes Móveis - CIDC Huawei.



Stefanie Barbosa Ancelmo Graduanda em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (INATEL). Técnica em Eletrônica com ênfase em Desenvolvimentos de Sistemas pela Escola Técnica de Eletrônica Francisco Moreira da Costa (ETE FMC) (2021). Atualmente atua como Bolsista de Iniciação científica no eHealth Inovattion Center.



Rita Elizabeth Santos de Almeida Possui graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações - Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center, centro de pesquisa de Engenharia Biomédica do Inatel.



Luma Rissatti Borges do Prado Graduada em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (Inatel), é mestranda em Engenharia Biomédica pela Escola Politécnica da Universidade de São Paulo (USP) e possui MBA em Gestão de Projetos e Metodologias Ágeis. Atua como Analista Sênior de Desenvolvimento de Produtos na Philips e é professora no curso de Engenharia Biomédica do Inatel, onde também é pesquisadora e orientadora no eHealth Innovation Center.



Elisa Renno Carneiro Dester Fisioterapeuta Pélvica, Mestre em Ciências da Motricidade Humana, Professora no Curso de Engenharia Biomédica do Inatel, Professora e Pesquisadora do eHealth Innovation Center - Inatel.

Acessibilidade Física e Sensorial no Ensino Superior: Um Estudo de Caso para a Promoção da Inclusão no Campus do Inatel

Luiz Otavio Ribeiro de Paiva, Arthur Felipe Vargas, Isabel Mendes, Fabiano Valias de Carvalho
Instituto Nacional de Telecomunicações - Inatel
luiz.paiva@ges.inatel.br , arthur.vargas@geb.inatel.br, isabel.f@dtel.inatel.br, fabiano@inatel.br

***Abstract*—This article presents a case study on accessibility at the National Institute of Telecommunications (Inatel) campus. The work approaches inclusion from a dual perspective: the removal of physical architectural barriers and the mitigation of sensory barriers, which especially affect neurodivergent individuals. Through a qualitative and exploratory methodology, which includes literature review, questionnaire analysis, and campus mapping, the study aims to identify obstacles in Building 1 and propose the creation of sensory regulation environments and the strengthening of inclusive practices.**

***Index Terms*—Accessibility, Inclusion, Higher education, Neurodivergence, Sensory rooms.**

***Resumo*—Este artigo apresenta um estudo de caso sobre a acessibilidade no campus do Instituto Nacional de Telecomunicações (Inatel). O trabalho aborda a inclusão sob uma dupla ótica: a remoção de barreiras arquitetônicas físicas e a mitigação de barreiras sensoriais, que afetam especialmente pessoas neurodivergentes. Através de uma metodologia qualitativa e exploratória, que inclui revisão bibliográfica, análise de questionários e mapeamento do campus, o estudo visa identificar obstáculos nos prédios do campus e propor a criação de ambientes de regulação sensorial e o fortalecimento de práticas inclusivas.**

***Palavras chave*—Acessibilidade; Inclusão, Ensino Superior, Neuro divergência, Salas Sensoriais.**

I. Introdução

A acessibilidade é definida como a possibilidade e condição de alcance para utilização, com segurança e autonomia, de espaços, mobiliários, equipamentos urbanos e sistemas de comunicação, por pessoas com deficiência ou com mobilidade reduzida [1]. Mais do que uma exigência, ela constitui um direito humano básico que assegura a cidadania e a participação social plena, eliminando quaisquer barreiras que limitem ou impeçam o convívio social e o acesso aos serviços essenciais. Com isso em mente, a acessibilidade no ensino superior destaca-se como um pilar fundamental para a inclusão, garantindo que o direito à educação seja exercido em sua totalidade. Contudo, muitas instituições enfrentam o desafio de transformar seus espaços físicos e pedagógicos para acolher a diversidade humana [2, 3].

No contexto brasileiro, observa-se um distanciamento entre a legislação vigente e a realidade prática das instituições. A falta de infraestrutura adequada, compreendida aqui como o cumprimento das normas técnicas de arquitetura (como rampas, pisos táteis e sinalização) [4], soma-se à ausência de suporte sensorial. Este último refere-se [5] à adequação de estímulos ambientais, como controle de iluminação e acústica, essenciais para evitar a desregulação de indivíduos com hipersensibilidade [6]. Essas lacunas erguem o que se denomina ”barreiras invisíveis”: obstáculos que, diferentemente de um degrau físico, não são percebidos visualmente pela maioria, mas isolam alunos neurodivergentes através do excesso de estímulos, impedindo sua permanência e êxito acadêmico. Para reduzir esses problemas, a legislação brasileira avançou significativamente com a Lei Brasileira de Inclusão (Lei nº 13.146/2015) [7], que assegura o direito à educação inclusiva em todos os níveis.

Nesse contexto, o presente estudo caracteriza-se como uma pesquisa de natureza qualitativa, com abordagem exploratória e delineamento de estudo de caso, tendo como objeto de análise o campus do Instituto Nacional de Telecomunicações (Inatel) [8], localizado em Santa Rita do Sapucaí–MG. A averiguação teve como propósito uma análise das condições de acessibilidade física e arquitetônica, motivada pela necessidade de alinhar o planejamento institucional às diretrizes do Desenho Universal, conceito que defende a criação de ambientes utilizáveis por todas as pessoas, sem necessidade de adaptação, conforme definido por Mace [9] e regulamentado no Brasil pelo Decreto nº 5.296/2004 [1] e também as normativas vigentes, como a ABNT NBR 9050 [4], que estabelece critérios de acessibilidade a edificações, mobiliário e espaços urbanos.

A identificação do problema aconteceu a partir de observações que evidenciaram a escassez de espaços acessíveis e de ambientes destinados à regulação sensorial no campus, em especial no Prédio 1, que é o local de maior fluxo acadêmico e centralidade de serviços. A partir do diagnóstico, o estudo propões o desenvolvimento teórico de diretrizes e modelos conceituais para espaços sensoriais adaptados e ações de capacitação contínua voltadas ao corpo docente e técnico-administrativo, focadas no letramento sobre neurodivergência e na eliminação de barreiras atitudinais. Essas propostas possuem caráter orientativo, a fim de ajudar a prefeitura do Campus na futura implementação prática dessas intervenções, contribuindo para um ambiente universitário mais inclusivo, equitativo e socialmente responsável.

II. Background

Para compreender a complexidade da inclusão no ensino superior, é necessário definir os conceitos que norteiam a acessibilidade e identificar os diferentes tipos de impedimentos que podem afetar a comunidade acadêmica.

A. Pessoas com Deficiência (PcD) e Neurodivergência

A deficiência é um conceito em evolução. No Brasil, o Censo Demográfico de 2022 [10] mostra que cerca de 18,6 milhões de pessoas (8,9% da população) possuem algum tipo de deficiência. A classificação, amparada pela legislação e literatura médica [1], inclui:

Deficiência Física/Motora: Alteração que compromete a função motora de um ou mais segmentos do corpo, manifestando-se como paraplegia, amputação, paralisia cerebral ou nanismo, entre outros [1].

Deficiência Visual: Engloba da baixa visão (mesmo com correção) à cegueira, exigindo recursos como piso tátil, Braille e leitores de tela.

Deficiência Auditiva: Perda auditiva bilateral de 41 decibéis (dB) ou mais [1], avaliada por audiograma, demandando alertas visuais e intérpretes de Libras.

Deficiência Intelectual e Neurodivergência: A primeira refere-se a limitações no funcionamento intelectual e adaptativo. Já a neurodivergência (como TEA e TDAH) envolve processamentos neurológicos atípicos. Destaca-se aqui a hipersensibilidade sensorial desse grupo, que se agrava em ambientes ruidosos ou visualmente poluídos [6, 5].

Quanto ao vocabulário, deve-se evitar termos pejorativos. Conforme Sassaki [11], utiliza-se Pessoa com Deficiência, pois a condição não define o indivíduo por completo (ex: usuário de cadeira de rodas, pessoa com deficiência visual).

B. Barreiras de Acessibilidade

Conforme a Lei Brasileira de Inclusão [7], barreiras são entraves que impedem a participação social e o exercício de direitos, classificadas em:

Urbanísticas e Arquitetônicas: Obstáculos em vias, espaços públicos e edifícios (ex: calçadas irregulares, escadas sem rampas e banheiros inadequados).

Nos Transportes e Comunicações: Veículos inacessíveis e dificuldades na troca de mensagens (ex: ausência de elevadores, legendas ou intérpretes).

Atitudinais e Tecnológicas: Preconceitos e plataformas digitais incompatíveis com tecnologias assistivas (como leitores de tela).

C. Tecnologia Assistiva

A Tecnologia Assistiva (TA) promove a funcionalidade e autonomia da PcD [7]. No contexto universitário [12], destacam-se:

Auxílios para estudo: Teclados ampliados, sorobã (ábaco adaptado para deficientes visuais) e linhas Braille.

Recursos Computacionais e Arquitetônicos: Leitores de tela, mouses adaptados, banheiros acessíveis e pisos táteis.

Acomodações Sensoriais: Abafadores de ruído, ferramentas de regulação tátil e controle luminotécnico.

Entender esses termos é essencial para propor melhorias. Vale ressaltar que, ao adequar estruturas para PcD's, beneficia-se toda a sociedade (idosos, pessoas com deficiências temporárias, etc.). A próxima seção apresenta as pesquisas base para este artigo.

III. Trabalhos Relacionados

A fundamentação deste estudo baseia-se na análise de pesquisas que abordam desde os desafios da infraestrutura física em campus universitários até as especificidades sensoriais de pessoas neurodivergentes.

A. Desafios da Acessibilidade Física e o Contexto Universitário

Diversos autores apontam que os obstáculos físicos tornaram-se uma rotina desgastante e normalizada. Os autores em [13] e em [3] destacam que problemas estruturais são reflexos de uma sociedade que ainda falha na manutenção de espaços inclusivos. A promoção da acessibilidade, portanto, não é apenas uma função social, mas um direito garantido [2]. Um estudo de caso na UNESP de Marília revelou que a inclusão social é frequentemente prejudicada por barreiras arquitetônicas e atitudinais [14]. O estudo concluiu que as instituições precisam intensificar o delineamento de suas ações, ouvindo o aluno com deficiência e seguindo rigorosamente as normas técnicas.

B. Neurodivergência e a Importância de Ambientes Sensoriais

Estudos mais recentes, como o apresentado em [6], identificam que adultos no espectro autista vivenciam o ambiente de forma distinta, relatando tanto hipersensibilidade quanto hipossensibilidade. Ambientes barulhentos e imprevisíveis foram citados como causas diretas de desregulação, ou seja, a dificuldade do sistema nervoso central em processar e organizar estímulos sensoriais excessivos, o que leva o indivíduo a um estado de estresse extremo, ansiedade, dor física e exaustão mental. Portanto, a criação de salas de acomodação sensorial surge como intervenção prática eficaz. Um exemplo nítido recente foi a implementação de espaços sensoriais na 76ª reunião anual da Sociedade Brasileira para o Progresso da Ciência (SBPC), sediada na Universidade Federal do Pará (UFPA). Conforme documentado em [15], a disponibilização de salas com isolamento acústico, luzes reguláveis e abafadores de ruído atendeu eficazmente participantes autistas, com TDAH e Síndrome de Tourette durante um evento de alto fluxo, comprovando que o acolhimento sensorial evita crises de sobrecarga (*meltdown/shutdown*) e garante a permanência do indivíduo no meio acadêmico.

A inovação tecnológica também tem se mostrado uma aliada, com pesquisas apontando o uso de Realidade Virtual (VR) para simular refúgios sensoriais, resultando em redução significativa de ansiedade [16].

IV. Metodologia

A metodologia foi dividida em três etapas principais:

A. Revisão da literatura

Foi realizada uma pesquisa bibliográfica nas bases de dados Google Acadêmico, SciELO e IEEE Xplore, buscando embasar teoricamente os conceitos de acessibilidade, neurodivergência e salas sensoriais. Para a seleção do material, foram utilizadas as palavras-chave: Acessibilidade no Ensino Superior, Neurodivergência, Salas Sensoriais e Barreiras Arquitetônicas, com recorte temporal dos últimos 4 anos. Do total de 22 artigos

encontrados, foram selecionados 13 que tratavam especificamente da aplicação prática desses conceitos em ambientes educacionais.

B. Coleta de Dados

A coleta de dados *in loco* foi realizada através de três instrumentos:

1) Reunião com a prefeitura do Campus: Realizada em 07 de maio de 2025, identificando como principal desafio a presença de "pedra portuguesa" nos calçamentos, cujo revestimento irregular dificulta a locomoção e a substituição complexa devido à extensão. Além disso, a reunião abrangeu o mapeamento dos projetos de acessibilidade atualmente em andamento no campus e o debate sobre lacunas de infraestrutura inclusiva observadas pelos pesquisadores, inclusive em locais onde já havia previsão institucional de adequação e melhora.
2) Questionário aos alunos: Aplicado via Google Forms à comunidade acadêmica de novembro de 2025 a fevereiro de 2026, obtendo 16 respostas válidas. As 10 questões do instrumento foram formuladas com base nos parâmetros da NBR 9050 e abrangeram os seguintes tópicos: avaliação geral da acessibilidade; nível de conforto no campus; facilidade de locomoção entre os prédios; adequação das instalações sanitárias e das salas de aula; clareza da sinalização; inclusão em eventos extracurriculares; eficácia dos comunicados institucionais; preparo dos professores e funcionários; liberdade para expressar necessidades de acessibilidade; e a identificação do prédio com melhor infraestrutura.
3) Entrevista exploratória: Realizada em 11 de junho de 2025 com um aluno neurodivergente, focando na experiência sensorial e identificando gatilhos de estresse no ambiente (ruídos, iluminação).

(*Nota: Houve uma tentativa de levantamento de dados junto ao Núcleo de Orientação Estudantil (NOE) em 28 de outubro de 2025, porém, não houve retorno institucional até o fechamento desta pesquisa*).

C. Análise Documental para Propostas de Inclusão

Para a estruturação das propostas de intervenção prática (capacitação e salas sensoriais), utilizou-se a metodologia de Análise Documental. Foram examinados os documentos oficiais do Ministério da Educação (Programa Incluir) e a Lei Brasileira de Inclusão [7]. A partir dessa análise, foram extraídas diretrizes que fundamentam a necessidade de práticas inclusivas integradas à infraestrutura do campus.

Na próxima seção são apresentados os resultados separados por eixos temáticos.

V. Resultados e Discussão

Os resultados apresentados a seguir estão divididos em duas partes: a primeira compreende as respostas obtidas pelo questionário online e a segunda apresenta a proposta estrutural das Salas de Acomodação Sensorial.

A. Primeira parte

O questionário online foi respondido por 16 membros da comunidade acadêmica. A análise dos dados foi subdividida em três eixos temáticos:

a.1) Eixo 1: Percepção Geral e Mobilidade

Inicialmente, buscou-se compreender a visão macro dos estudantes sobre o campus. Ao serem questionados sobre como a acessibilidade no campus era avaliada, a maioria dos respondentes demonstrou satisfação. Como observa-se na Figura 1, 62,5% classificaram como boa e 31,3% como muito boa. Esse sentimento positivo reflete-se na sensação de acolhimento, uma vez que a percepção de um ambiente físico transitável e com boa manutenção estrutural remete aos usuários a mensagem de que suas necessidades de mobilidade são respeitadas, gerando segurança e pertencimento.

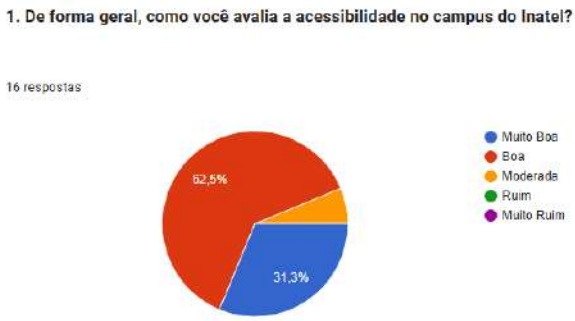


Fig. 1. Avaliação geral da acessibilidade.

A pergunta 2 questionava sobre o nível de conforto e acolhimento ao circular pelo campus: 62,5% avaliaram como muito boa, indicando que o ambiente físico, em sua maioria, não é percebido como hostil (Figura 2).



Fig. 2. Nível de conforto e acolhimento.

No quesito locomoção entre os prédios, presente na pergunta 3, 50% a consideraram muito boa e 43,8% como boa, como pode ser observado na Figura 3.



Fig. 3. Facilidade de locomoção entre prédios.

Um dado fundamental para a validação do estudo de caso foi a escolha do local de intervenção. A pergunta 11 indagou qual prédio era considerado com melhor acessibilidade. O prédio 1 foi o mais citado (37,5%), seguido pelo prédio 5 (25%) e prédio 2 (18,8%). Este dado corrobora a escolha do prédio 1 para a proposta da Sala Sensorial, pois, além de ser o local de maior fluxo e centralidade, já é reconhecido pelos alunos como referência de acesso, facilitando a implementação de melhorias (Figura 4).



Fig. 4. Prédio com melhor acessibilidade percebida.

a.2) Eixo 2: Infraestrutura Física e Sinalização

Apesar da avaliação geral positiva, a análise específica da infraestrutura revelou pontos de atenção que justificam a necessidade de monitoramento constante. Sobre as instalações sanitárias (pergunta 4), embora 56,3% avaliem como muito boa, 12,5% classificou como ruim. Isso indica que, para usuários com necessidades específicas, os banheiros adaptados podem não estar atendendo plenamente às normas da NBR 9050 (Figura 5).

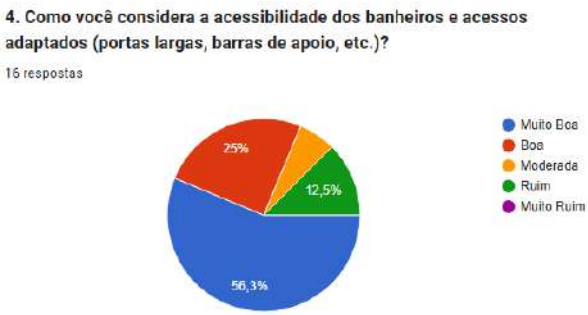


Fig. 5. Acessibilidade dos banheiros.

Cenário similar ocorre nas salas de aula (pergunta 5 - cadeiras, acústica, visibilidade), onde também houve registro de 12,5% de avaliação ruim. A questão acústica é particularmente relevante para a justificativa das salas sensoriais propostas neste trabalho (Figura 6).



Fig. 6. Acessibilidade nas salas de aula.

Quanto à sinalização do campus (visual, tátil ou sonora), a pergunta 6 mostrou que 50% consideram boa e 37,5% muito boa. Contudo, a ausência de sinalização sonora ou em Braille em todos os pontos ainda constitui uma barreira comunicacional para pessoas com deficiências visuais (Figura 7).



Fig. 7. Avaliação da sinalização.

a.3) Eixo 3: Aspectos Pedagógicos, Sociais e Comunicacionais

Este eixo apresentou os dados mais heterogêneos, evidenciando onde residem as maiores barreiras invisíveis e atitudinais. Na pergunta 7, sobre a inclusão em eventos e atividades extracurriculares, houve um registro preocupante: 12,5% avaliaram como muito ruim. Isso sugere que a vida social acadêmica pode estar excluindo parte dos alunos (Figura 8).

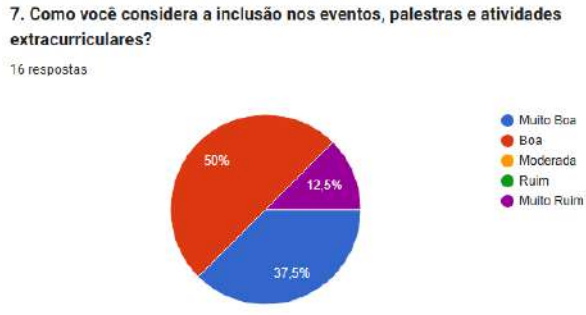


Fig. 8. Inclusão em eventos extracurriculares.

A comunicação institucional (pergunta 8) foi bem avaliada, mas com 12,5% considerando moderada e 6,3% ruim, ilustado na Figura 9.



Fig. 9. Clareza dos comunicados institucionais.

O dado mais crítico, que fundamenta a urgência de políticas institucionais de capacitação, está na pergunta 9, que questionou como o aluno avaliava a preparação dos professores e funcionários para lidar com neurodivergências e deficiências. Observa-se que 31,3% consideram a preparação apenas moderada. Diferente da infraestrutura física, que teve alta aprovação, o preparo humano apresenta uma lacuna clara (Figura 10).



Fig. 10. Preparo da equipe (professores e funcionários).

Por fim, a pergunta 10 investigou o clima organizacional, questionando sobre a liberdade para se expressar necessidades. O resultado foi positivo, com 50% sentindo-se muito bem e 37,5% bem. Isso indica que a instituição possui um canal aberto para o diálogo, o que é um facilitador para a implementação das propostas deste artigo (Figura 11).

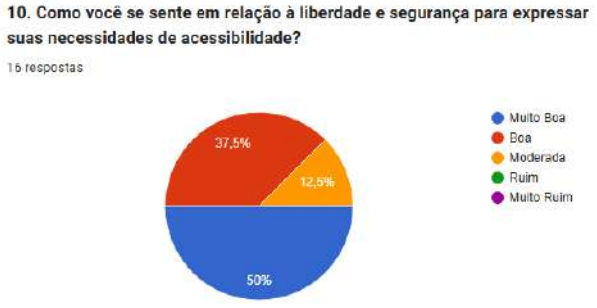


Fig. 11. Liberdade para expressar necessidades.

B. Segunda parte

Com base nas demandas por conforto acústico ilustradas na Figura 6 e no déficit no preparo da equipe para lidar com a neurodivergência da Figura 10, propõe-se a implementação de estratégias voltadas especificamente ao acolhimento sensorial.

b.1) Salas de Regulação Sensorial:

Sugere-se a adequação de um ambiente no prédio 1, validado pela Figura 4, focado na acomodação de alunos em momento de sobrecarga sensorial. Inspirado nas diretrizes exitosas adotadas na 76ª reunião da SBPC [15], este espaço não exige grandes reformas estruturais, baseando-se em tecnologias assistivas e design inclusivo. A sala deve conter iluminação controlada (luzes de LED com *dimmer* e paleta de cores neutras), isolamento acústico eficaz contra os barulhos externos, além de abafadores de ruído e mobiliário confortável (puffs, redes ou tapetes ponderados). A Figura 12 apresenta uma referência visual prática de como um ambiente de regulação sensorial pode ser estruturado, utilizando como exemplo o espaço implementado no Bioparque Pantanal [17].



Fig. 12. Exemplo prático de uma Sala de Acomodação Sensorial no Bioparque Pantanal.

b.2) Capacitação Institucional:

A eficácia do acolhimento depende de profissionais que identifiquem gatilhos de desregulação. Assim, recomenda-se a capacitação periódica da equipe do Campus sobre neurodivergência. Como referência de sucesso, a Pró-Reitoria de Inclusão e Pertencimento (PRIP) da USP [18] demonstrou que treinar docentes para acolher alunos neurodivergentes reduz significativamente as barreiras atitudinais, facilita adaptações metodológicas e promove um ambiente mais empático e com menor estresse acadêmico.

VI. Conclusão e Trabalhos Futuros

A relevância deste trabalho reside na ampliação do conceito de acessibilidade no ensino superior, demonstrando que a verdadeira inclusão transcende a construção de rampas e banheiros adaptados, englobando também a mitigação de barreiras sensoriais e atitudinais muitas vezes invisíveis. As intervenções propostas, especialmente a Sala de Acomodação Sensorial e o letramento contínuo da equipe, têm o potencial de impactar diretamente o bem-estar e a saúde mental dos alunos. Ao oferecer um ambiente previsível e acolhedor, a instituição atua diretamente no combate à evasão universitária, garantindo que o direito à educação seja exercido com dignidade, autonomia e equidade por alunos neurodivergentes ou com deficiência.

Como trabalhos futuros, sugere-se a implementação física do projeto piloto da Sala Sensorial no prédio 1 para a validação de sua eficácia prática por meio de avaliações dos usuários. Recomenda-se, também, a expansão do mapeamento arquitetônico e sensorial para os demais prédios da instituição e o desenvolvimento de soluções tecnológicas, como um aplicativo institucional que mapeie rotas acessíveis e indique o nível de

ruído dos ambientes do campus em tempo real, fornecendo maior autonomia à comunidade acadêmica.

VII. Agradecimentos

Gostaríamos de expressar nossa gratidão à Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG, ao Instituto Nacional de Telecomunicações - INATEL, e ao eHealth Innovation Center pelo apoio técnico e recursos viabilizados para o desenvolvimento deste projeto.

Referências

[1] Brasil. *Decreto nº 5.296, de 2 de dezembro de 2004: Regulamenta as Leis nº 10.048/2000 e nº 10.098/2000*. Brasília, DF, 2004.

[2] C. B. de Oliveira. "Jovens deficientes na universidade: experiências de acessibilidade?" Em: *Revista Brasileira de Educação* 18.55 (2013), pp. 961–984.

[3] J. C. da Silva e A. M. Pimentel. "Inclusão educacional da pessoa com deficiência visual no ensino superior". Em: *Cadernos Brasileiros de Terapia Ocupacional* 29 (2021), e2904.

[4] Associação Brasileira de Normas Técnicas - ABNT. *NBR 9050: Acessibilidade a edificações, mobiliário, espaços e equipamentos urbanos*. Rio de Janeiro, 2020.

[5] A. McDonnell et al. "Sensory and Relaxation Preferences to Inform the Design of Calming Spaces and Sensory Rooms for Autistic Adults". Em: *Autism in Adulthood* (2024). URL: https://www.liebertpub.com/doi/abs/10.1089/aut.2024.0088?journalCode=aut.

[6] A. E. Robertson e D. R. Simmons. "Understanding the Sensory Experiences of Autistic Adults: A Qualitative Study". Em: ().

[7] Brasil. *Lei nº 13.146, de 6 de julho de 2015: Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência*. Brasília, DF, 2015.

[8] Instituto Nacional de Telecomunicações - Inatel. *Portal Institucional*. Acesso em: 2026. 2026. URL: https://www.inatel.br.

[9] Ronald Mace. "Universal Design: Barrier Free Environments for Everyone". Em: *Designers West* 33.1 (1985), pp. 147–152.

[10] Instituto Brasileiro de Geografia e Estatística - IBGE. *Censo Demográfico 2022: Pessoas com deficiência*. Acesso em: 2026. 2023. URL: https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-demografico-2022.html.

[11] Romeu Kazumi Sassaki. *Inclusão: construindo uma sociedade para todos*. 7ª ed. Rio de Janeiro: WVA, 2009.

[12] Bárbara Larissa Alexandre Filgueira. *Guia para desenvolvimento de políticas institucionais de acessibilidade e inclusão das PcD nas Universidades Públicas Federais do Brasil*. 139 f. 2021.

[13] R. S. da S. Pereira et al. "People with acquired physical disabilities: From Activities of Living to rehabilitation nursing care". Em: *Texto & Contexto Enfermagem* 33 (2024), e20230362.

[14] Revista CIMEAC. *Estudo sobre Acessibilidade*. Acesso em 2025. 2025. URL: https://seer.uftm.edu.br/revistaeletronica/index.php/cimeac/article/view/4929.

[15] Miguel Costa Silva, Keila de Jesus Morais Lobato e Scheilla de Castro Abbud Vieira. "Salas de Acomodação Sensorial: Práticas Inclusivas para Pessoas Neurodiversas e com Transtorno no Processamento Sensorial". Em: *Anais do I Congresso Nacional de Práticas de Ensino na Educação Inclusiva (ICNPEEI)*. Belém, 2024. ISBN: 978-65-5465-106-6.

[16] K. Valencia et al. "A Virtual Reality-Based Sensory Room for Adults with Disabilities". Em: ().

[17] Bioparque Pantanal. *Sala de acomodação sensorial garante espaço reservado para autistas no Bioparque Pantanal*. Acesso em: 2026. 2022. URL: https://bioparquepantanal.ms.gov.br/sala-de-acomodacao-sensorial-garante-espaco-reservado-para-autistas-no-bioparque-pantanal/.

[18] Universidade de São Paulo - USP. *Pró-Reitoria de Inclusão e Pertencimento (PRIP)*. Acesso em: 2026. 2022. URL: https://prip.usp.br/.

Autores



Luiz Otavio Ribeiro de Paiva Graduando do 7º período de Engenharia de Software no Instituto Nacional de Telecomunicações. Atualmente atuando como estagiário pelo BRDC no Inatel/Huawei, sempre procurando expandir seus conhecimentos na área e tentando se tornar um profissional de excelência.



Arthur Felipe Vargas Graduando em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (Inatel) e, atualmente, bolsista de Iniciação Científica no eHealth Innovation Center.



Isabel Francine Mendes Doutoranda em Telecomunicações pelo Inatel (WAI-Lab), focada em soluções de visão computacional para sensoriamento de humanos e animais. Mestre pela USP em Sistemas de Informação com foco em acessibilidade digital.



Fabiano Valias de Carvalho Graduado em Engenharia Elétrica com ênfase em Telecomunicações e Eletrônica pelo Instituto Nacional de Telecomunicações (1995) e mestre em Engenharia Elétrica pela Universidade Federal de Itajubá (2002). É coordenador do laboratório eHealth Innovation Center e do curso de Graduação em Engenharia Biomédica.

Acessibilidade digital: Avaliação de Concordância do Portal Acadêmico do Inatel com ferramentas automatizadas apoiando-se nas normas da WCAG 2.2

Wagner Dourado Felix da Silva, Isabel Mendes, Rita Elizabeth Santos de Almeida, Fabiano Valias de Carvalho
Instituto Nacional de Telecomunicações - Inatel
wagner.silva@gec.inatel.br, isabel.f@dtel.inatel.br, r.elizabeth@inatel.br, fabiano@inatel.br

***Abstract*—Digital accessibility in educational environments is important for the inclusion of people with disabilities (PwD). This article presents an analysis and validation of the compliance of the Academic Portal (AP) of Inatel - National Institute of Telecommunications - based on the guidelines imposed by WCAG 2.2. Based on the lack of accessibility and the divergence between observed market practices and current laws requiring exemplary accessibility, this work performs a partial validation of six essential pages in Inatel's AP, chosen based on research conducted among a group of students at the institution. The tool used for page validation was WAVE, from WebAIM. It stands out for its visual efficiency in analyzing navigation difficulties and page code incompatibility. The observed events indicate only partial compliance, revealing crucial violations such as page contrast failures, lack of labels in forms, and absence of text descriptions in functional images. The results show the urgent need for semantic and structural corrections to ensure equal access to academic information for all users.**

***Index Terms*—Digital accessibility, WCAG 2.2, Inclusion, University education, Technical validation.**

***Resumo*—A acessibilidade digital em ambientes da educação é importante para que haja inclusão de Pessoas com deficiência (PcD). Este artigo mostra uma análise e validação de conformidade do Portal Acadêmico (PA) do Inatel - Instituto Nacional de Telecomunicações com base nas diretrizes impostas na WCAG 2.2. Com base na falta de acessibilidade e a divergência entre a prática observada no mercado e as leis atuais que exigem uma acessibilidade exemplar, o trabalho faz uma validação parcial de seis páginas essenciais no PA do Inatel, que foram escolhidas com base em uma pesquisa feita em um grupo de alunos da instituição. A ferramenta utilizada para validação das páginas foi a WAVE, da WebAIM. Ela se destaca por sua eficiência visual na análise de dificuldade na navegação e incompatibilidade de código da página. Os eventos observados indicam um cenário de conformidade apenas parcial, revelando violações cruciais como falhas no contraste da página, falta de rótulos em formulários e ausência de descrição em formato de texto em imagens funcionais. Os resultados mostram a necessidade urgente de correções semânticas e estruturais para garantir a igualdade no acesso às informações acadêmicas por todos os usuários.**

***Palavras chave*—Acessibilidade digital, WCAG 2.2, Inclusão, Ensino universitário, Validação técnica.**

I. Introdução

Atualmente, a internet se firmou como um sistema crucial para realizar atividades do dia a dia, desde acesso à sites educacionais quanto para movimentações financeiras e profissionais [1]. No contexto acadêmico, essa evolução digital explica uma questão importante: as Instituições de Ensino Superior (IES) passaram a usar portais eletrônicos como meio obrigatório de acompanhamento acadêmico. Nesse quesito, entra uma questão: quando existem portais que não são construídos com base na regras de acessibilidade, eles deixam de lado a sua missão de entregar suporte a qualquer usuário que esteja usando o portal e impõem dificuldades em relação à inclusão. Isso faz com que Pessoas com Deficiência (PcD) tenham limitações ao uso de tarefas essenciais no sistema acadêmico.

Diante de desafios de natureza semelhante, o World Wide Web Consortium (W3C), teve como objetivo definir diretrizes de acessibilidade em sites, publicando em 1999 a Web Content Accessibility Guidelines (WCAG) que se tornou referência internacional, fundamentando normas para a acessibilidade digital. Já no Brasil, a Lei Brasileira de Inclusão (LBI) [2] reafirma a obrigação de acessibilidade em sites mantidos pelo governo e empresas no país.

Para comprovar o uso da acessibilidade digital, este trabalho faz uma validação em páginas do Portal Acadêmico (PA) do Inatel - Instituto Nacional de Telecomunicações, localizado em Santa Rita do Sapucaí - MG. A pesquisa entende como é necessário uma verificação inicial e semiautomatizada, considerando que validações manuais demoram um longo prazo de desenvolvimento. Com o intuito de realizar uma verificação do quão acessível é o portal às normas impostas pela WCAG 2.2 [3], sendo a versão mais atualizada e recomendada pelo W3C, a ferramenta WAVE [4] foi escolhida, por ser de fácil entendimento e com curva de aprendizado pequena.

Este artigo está organizado da seguinte maneira: a Seção II apresenta um background; a Seção III aborda trabalhos relacionados; a Seção IV detalha a Metodologia adotada; a Seção V exibe os resultados e discussão; E finalmente são apresentados a conclusão e os trabalhos aconselháveis no futuro.

II. Background

Com o intuito de entender a avaliação de acessibilidade que está sendo proposta neste artigo, será preciso contextualizar alguns conceitos essenciais sobre o que é deficiência, suas dificuldades na parte digital e as regras técnicas que orientam a construção de uma internet mais acessível.

A. Acessibilidade digital e a Pessoa com Deficiência

Quando é falado de acessibilidade digital, refere-se à capacidade da web ser consumida por todas as pessoas e isso não leva em conta suas limitações físicas, sensoriais ou cognitivas. Segundo a LBI, toda pessoa que tem um impedimento físico, mental, intelectual ou sensorial que atrapalhe em interações rotineiras é considerada Pessoa com Deficiência. Essas limitações podem ser graves a ponto de sua participação na comunidade ser totalmente desigual e diferente de outras pessoas [2]. Segundo dados da Organização Mundial da Saúde (WHO) é estimado que 1.3 bilhões de pessoas, algo em torno de 16% da população global, vivam com algum tipo de deficiência. Isso torna a acessibilidade um tema global de direitos humanos de caráter obrigatório [5]. De acordo com o W3C com a WCAG 2.2, são separados em quatro grupos as limitações que podem atrapalhar o uso da internet e navegação no dia a dia [3]:

Visual: Nesse grupo, são incluídas deficiências como a baixa visão até a cegueira total e o daltonismo. O conteúdo visual não detalhado com imagens sem texto alternativo e páginas com contraste fora do padrão são as maiores dificuldades para este grupo.

Auditiva: Aqui pode-se incluir pessoas com surdez parcial ou total. Vídeos e conteúdos sem legendas são impossíveis de entender para este grupo. É importante ressaltar que no Brasil existe a Língua Brasileira de Sinais, a Libras, admitida de forma legal como uma forma de comunicação. Isso configura como mais uma forma eficiente e acessível de incluir PcDs auditivos no país [2].

Motora: Toda limitação na coordenação motora, como o caso de amputação ou tremores. A principal dificuldade que é encontrada nesse caso é o uso do mouse e teclado, que são primordiais para navegação na internet [3].

Cognitiva e Intelectual: É um termo que se usa quando uma pessoa apresentas certas limitações no seu funcionamento intelectual, tanto para aprendizagem ou desempenho de tarefas diversas como a comunicação, cuidado pessoal e relacionamento social com outras pessoas. Atrasos em alguns movimentos também são incluídos, como a dislexia. As adversidades podem incluir textos longos e confusos, páginas muito coloridas e o uso de figura de linguagem e botões que mudam de posição na tela [3].

Dentro desse contexto, ainda vale ressaltar a existência de mais um caso que pode ser chamado de deficiência múltipla, onde há a associação em uma mesma pessoa de duas ou mais deficiências primárias, sendo ainda mais difícil a adaptação na sociedade [6]. Um exemplo são PcDs física e intelectual, que possuem limitações motoras e dificuldade cognitivas combinadas [5]. PcDs física e intelectual poderão ter problemas tanto na parte de navegação com o mouse e teclado no portal como na capacidade de entender um texto longo.

B. Tecnologias Assistivas

Tecnologia Assistiva, ou TA, são métodos, técnicas ou produtos que auxiliam no cotidiano de PcDs [7, 8]. Há 12 categorias de TA sendo separadas por tipo de auxílio, indo desde um comunicador para pessoas com limitações de fala até programas que processam os dados que o navegador oferece e convertem o conteúdo em áudio ou em Braille para dificuldade visual [8].

C. As Diretrizes WCAG

A acessibilidade digital teve seu padrão e normas definidas pelo W3C com a Web Accessibility Initiative (WAI). Com isso, foi criado o principal documento que estabelece normas e diretrizes de caráter cruciais internacionalmente que é a WCAG [3], que é recomendada tecnicamente por ser um documento em constante aprimoramento e melhoras relevantes no ramo da acessibilidade. A versão 2.0, apresentada em 2008, trouxe o modelo de critérios testáveis, enquanto a 2.1, divulgada em 2018, ampliou para normas que se aplicam em dispositivos móveis. Atualmente, a recomendação atual em uso é a WCAG 2.2, sendo publicada no final de 2023. Esta versão adicionou novas regras focadas especialmente em aprimorar a praticidade para usuários com baixa visão e dificuldades cognitivas [3].

No que se refere a um site acessível, a WCAG sugere três camadas cruciais que devem ser levadas em conta:

1) **Conhecendo os Pilares POUR:** Qualquer página de acesso precisa atender obrigatoriamente a quatro regras essenciais que o W3C define [3]:
   a) Perceptível:O usuário precisa perceber pelos seus sentidos a informação apresentada, como por exemplo os textos alternativos nas imagens.
   b) Operável: Qualquer que seja o dispositivo usado, a página tem que ser ativa com o usuário, como por exemplo a navegação de um site deve ser possível tanto por teclado quanto por mouse.
   c) Compreensível: A navegação na página deve ter sentido lógico e compreensivo, tanto seus textos e imagens quanto formulários.
   d) Robusta: Tecnologias como Leitores de telas devem interpretar perfeitamente o código, como por exemplo hierarquia de títulos, idioma da página e texto explicativo nos campos de formulários, por exemplo.
2) **Diretrizes da WCAG:** A WCAG estruturou normas grandes chamadas Diretrizes, que ficam logo embaixo dos princípios, e regras testáveis de acordo com uma especificação, chamada de Critérios de Sucesso. Essa numeração segue uma hierarquia de uma forma rigorosa. Um exemplo seria uma validação mostrar que foi violado o critério 1.4.3. Isso quer dizer que o erro está no Princípio 1, de perceptível, com a diretriz 4 de Distinguível, e à regra de testes número 3 que aborda problema de contraste mínimo. As treze diretrizes estão classificadas dentro dos Pilares POUR na WCAG 2.2 [3].
3) **Níveis de conformidade:** Os Critérios de Sucesso são separados em três níveis pela WCAG [3]:
   a) Nível A: É a conformidade mais básica da WCAG, porém, a que mostra as barreiras mais significativas de um site inacessível.
   b) Nível AA: Para atender o nível AA, o site deve atingir todos os critérios do nível A, além de ser funcional e compreensível para a maioria das pessoas com ou sem deficiência.

c) Nível AAA: Esse é o nível máximo de conformidade, que torna o site acessível a um número máximo de usuários com deficiência ou não, idosos, analfabetos e ainda aqueles que desconhecem o idioma nativo do site.

D. Ferramentas Automatizadas para Avaliação

Com o propósito de verificar a concordância junto à WCAG 2.2, há a possibilidade de validação semiautomática através de ferramentas de acessibilidade digital. Elas não só trabalham para encontrar erros que violam a WCAG como também analisam a parte do código e se o mesmo está estruturado e bem montado. Algumas ferramentas são:

Google Lighthouse: Ferramenta de código aberto, produz uma pontuação em forma escalar de 0 a 100 e se baseia em acessibilidade e testes de desempenho [9].

axe DevTools: Desenvolvida pela Deque Systems e segundo o próprio site da fabricante possui mais de 3 bilhões de downloads, é conhecida por sua técnica da árvore DOM e análise através dela, reconhecendo erros de sintaxes que podem bloquear TAs como leitores de tela [10].

Web Accessibility Evaluation Tool - WAVE: Criada pela WebAIM, difere das anteriores por sua abordagem visual. Ela injeta ícones diretamente na interface da página, permitindo identificar erros e localizá-los no código, bem como uma explicação do erro e a diretriz da WCAG que foi violado. [4].

Neste artigo, a ferramenta WAVE foi a escolhida para a validação do PA por apresentar características mais didáticas em comparação às demais ferramentas mostradas. No aspecto visual, o WAVE destaca-se por sua abordagem direta, evidenciando as dificuldades de acessibilidade na própria página do usuário. Isso o diferencia do Lighthouse, que fornece apenas uma pontuação escalar, e do axe DevTools, cujos relatórios são baseados em linhas de código no terminal e não possuem um apelo visual didático.

III. Trabalhos Relacionados

A revisão da literatura foi importante pois, embora a legislação brasileira seja avançada, a implantação no ponto de vista técnico em portais de instituições segue sendo algo insuficiente. No que diz respeito a sites universitários, as pesquisas convergem para um relatório de não conformidade constante. Os autores em [11], ao verificarem portais de universidades federais, e em [12], ao investigar base de dados científicas, mostraram que a parte acadêmica fracassa em pontos básicos de navegação acessível. Ambos os estudos deixam claro que, apesar das universidades serem exemplo e polos de conhecimento, suas plataformas digitais ignoram e erram princípios fundamentais da WCAG, como a falta de descrição de imagens e hierarquia de cabeçalhos de forma inadequada, o que provoca uma exclusão grande de PcDs.

Examinando e destacando causas de falhas como essas, os autores em [13] investigaram como a validação de acessibilidade é realizada na indústria de software. Eles identificaram que a principal dificuldade está na cultura de testes. Existe uma dependência muito grande de verificações manuais que são vagas e, junto a isso, há uma ignorância sobre as diretrizes que são cobradas. Uma das conclusões do estudo é que a falta de ferramentas automatizadas no processo de qualidade acaba resultando em projetos que chegam ao usuário final com erros sérios tanto na parte de sintaxe como na parte semântica.

Este trabalho buscou entender e aprofundar a carência de metodologia e os problemas abordados que mostra na análise dos autores em [11], com uma contextualização sobre o estudo de [13]. Diferente dos estudos anteriores que focaram em portais públicos ou na cultura industrial como um todo, esta pesquisa adotou o uso de ferramenta semiautomatizada para validar um PA privado de uso frequente. A proposta foi mostrar o quanto algumas páginas do PA são inacessíveis para uma parcela estudantil e apresentar possíveis correções para essas páginas.

IV. Metodologia

Para esse artigo, além do contexto teórico foi adotada uma abordagem de quantidade e qualidade que tem fundamento na metodologia da Ciência da Computação sugerida pelo autor Wazlawick [14]. A parte prática foi dividida em dois tópicos:

A. Identificação do instrumento para estudo e escolha de uma amostra

O PA é composto por muitas páginas com propostas diferentes e com opções dinâmicas, como por exemplo a necessidade de entrar com um login e senha. Nesse contexto apenas algumas páginas foram selecionadas para a validação de acessibilidade digital. A ideia inicial foi buscar auxílio do setor de Tecnologia da Informação da instituição para identificar os sites de maior acesso e fornecer mapas de calor. Essa ideia logo foi descartada pelo tempo hábil que seria demandado ao setor e a escrita desse artigo.

A segunda ideia foi a realização de uma pesquisa junto aos estudantes da instituição sobre quais páginas eram mais acessadas. Essa pesquisa aconteceu em 3 de dezembro de 2025 e 9 alunos dos cursos de graduação responderam ao questionário. Essa dinâmica foi feita através de aplicativo de mensagem e foi perguntando quais eram as três páginas mais importantes e obrigatórias para o seu cotidiano acadêmico.

A Figura 1 mostra como foi a distribuição dos votos, deixando claro a predominância da página ”Notas”(escolhida por 7 dos 9 participantes) e um empate de votos entre ”Quadro de Horários”, ”Bolsa Finatel”e ”Frequência”.

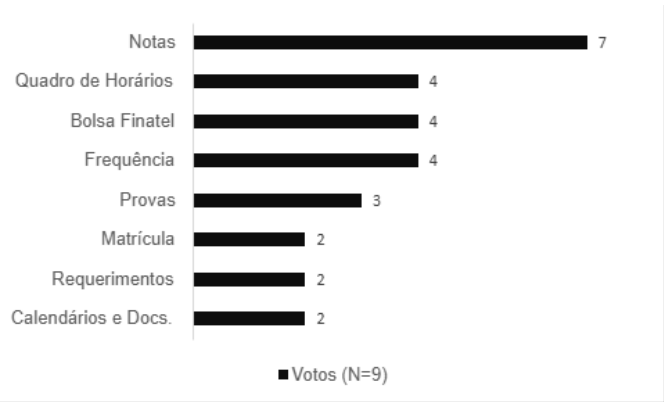


Fig. 1. Quantidade de votos nas páginas do PA.

A decisão final foi para seis opções das disponibilizadas na pesquisa:

Página de Login: É a porta de entrada do PA e obrigatório para acessá-lo. Sem essa página, não há continuidade na navegação;

Página Principal: É para onde a página de login redireciona. É uma interface inevitável e tem-se a obrigação de acessá-la;

Notas: Observada como página de maior importância na pesquisa, com 7 votos, devido à necessidade de seu acompanhamento frequente entre todos os alunos;

Quadro de Horários: Primordial para organizar o dia a dia acadêmico e ficou com 4 votos. Com ela, os estudantes têm a organização para as aulas e sua localização;

Bolsa Finatel: Por conta do número elevado de alunos bolsistas na faculdade, é uma página de crucial acesso e através dela as bolsas socioeconômicas são fornecidas e renovadas;

Frequência: A instituição possui regras como uma quantidade de frequência nas aulas para aprovação. Sendo assim, o controle do aluno sobre esse fator deve ser levado a sério para sua formação. Ficou com 4 votos no total.

B. Verificação e Análise

As validações foram realizadas no navegador Google Chrome em sua versão convencional e sem nenhum outro auxílio, nem extensões como ADBlock (bloqueador de anúncios automáticos), que poderia impactar no resultado final da análise da página. Todas as seis páginas selecionadas foram verificadas, de forma separada, utilizando a ferramenta de acessibilidade WAVE. Todos os erros e problemas descobertos foram registrados e ordenados como: crítico que impede o usuário de navegar; sério, que não bloqueia mas dificulta o acesso; moderados e mínimos, sendo um alerta e não necessariamente um erro. Essas nomenclaturas aos erros foram adotados por ser o padrão imposto pela Deque Systems. Os resultados encontrados estão exibidos na próxima seção.

V. Resultados e Conclusão

A avaliação revelou um cenário de concordância apenas parcial. O PA apresenta falhas de Perceptibilidade, onde há informações que não podem ser percebidas pelos sentidos do usuário, e Operabilidade, que causa dificuldade do usuário ao interagir com os elementos da interface [3]. A figura 2 sumariza a quantidade de erros em cada página testada, sendo categorizados em Erros Estruturais e Erros de Contraste. A separação específica dos problemas de contraste justifica-se pela sua alta incidência em todas as interfaces, o que preocupa quando se olha as normas da WCAG 2.2. Além disso, foi observado que páginas com grande volume de dados ordenados, como o Quadro de Horários, focam em vários alertas de formatação, sugerindo que houve um desenvolvimento de código não detalhado.

A violação de acessibilidade digital mais recorrente foi a falta de contraste de cores entre o texto e o fundo. Este erro impossibilita a leitura para usuários com baixa visão, daltonismo e até mesmo usuários com visão normal em ambientes muito iluminados [3]. Esse tipo de erro é detalhado pela WCAG e sua identificação é 1.4.3, que descreve sobre contraste mínimo e requer uma razão matemática definida de 4.5:1 para textos regulares. Além disso, detectou-se a violação do critério 1.4.11

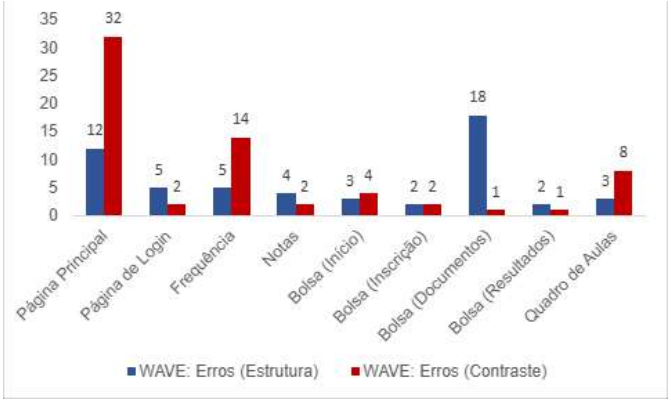


Fig. 2. Comparativo entre Erros Estruturais e Erros de Contraste na ferramenta WAVE.

onde aborda o contraste não-textual que exige a razão de 3:1 para componentes gráficos. Essa infração dificulta a utilização de elementos de tela, como por exemplo itens clicáveis.

Outra dificuldade séria é a falta do elemento *alt* nas imagens. Quando um aluno com cegueira acessa o PA através de Leitor de Tela, o *software* procura por essa descrição e se não achar, o leitor entrega para o usuário a palavra imagem ou o nome do arquivo, limitando o estudante de entender o que há na interface [12, 13]. A Figura 3 expõe um exemplo claro na Página Principal, que foi capturado pelo WAVE, mostrando diretamente a violação do Critério 1.1.1 de conteúdo não textual da WCAG 2.2 [3]

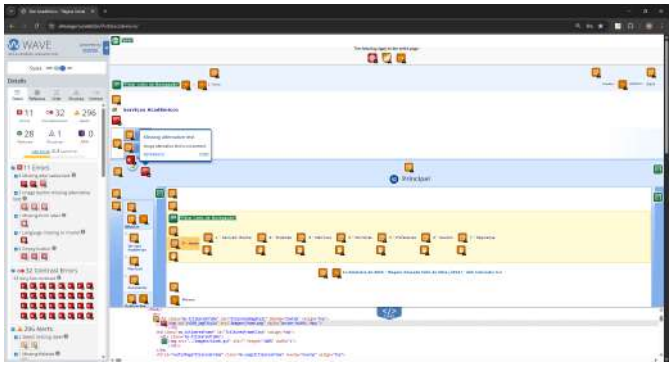


Fig. 3. Detecção de falhas sérias de contraste e falta de texto alternativo na Página Principal.

Em relação ao código em si, a análise revelou falhas graves na estrutura semântica das páginas. Uma das características da linguagem HTML é que a estrutura de cabeçalho é feita por marcações como `<h1>` para o título principal da página, `<h2>` para o subtítulo e assim por diante. [15]. A verificação através do WAVE apontou que o PA quebra essa ordem pulando de um `<h2>` direto para um `<h4>`. Para um usuário que navega pelos títulos usando o teclado, essa quebra pode causar confusão e falta de sentido na navegação. Como pode ser visto na figura 4, a página Quadro de Horários gerou diversos erros, incluindo erros de cabeçalhos, o que viola o critério 1.3.1 de Informações e Relações [3].

Outra dificuldade crucial na operabilidade foi detectada nas áreas de inserção de dados, como o login: os campos de digitação não possuem rótulos associados no código. Com isso, o leitor de tela anunciará apenas um campo de edição

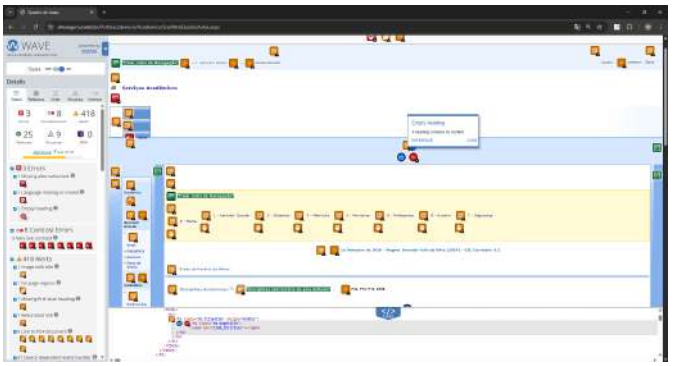


Fig. 4. Alto índice de alertas estruturais na página de Quadro de Horários.

em branco [13], resultando em um caso onde o estudante não saberá se deve digitar CPF, matrícula ou senha. A figura 5 ilustra essa falha na página de Login, violando o critério 3.3.2 sobre Rótulos ou Instruções [3].

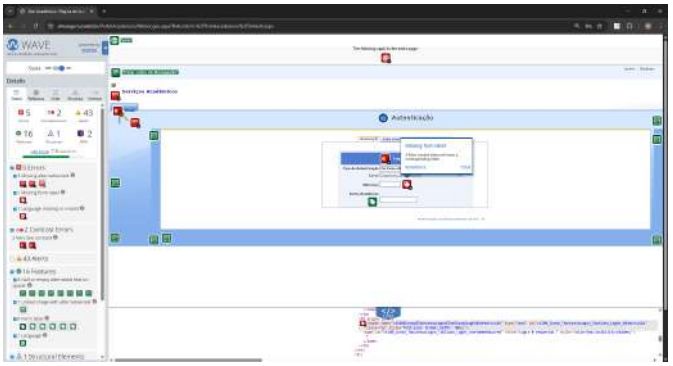


Fig. 5. Identificação de campos de formulário sem rótulos *labels* na página de Login.

Diante do diagnóstico de irregularidade de acessibilidade revelado, existem alguns elementos a ressaltar: recomenda-se um planejamento de melhorias com base nos problemas relatados e com uma atenção nos erros mais sérios e de caráter crítico. A solução encontrada será eficiente, pois são erros em abundância que podem ser tratados simultaneamente. A Tabela I resume as correções de código necessárias para diminuir as limitações, alinhando a estrutura do PA aos critérios exigidos pela WCAG 2.2 [3].

TABELA I
SUGESTÕES PARA AJUSTES TÉCNICOS BASEADO NA WCAG 2.2

URGÊNCIA	DIFICULDADE LOCALIZADA	CRITÉRIO EM FALTA DA WCAG 2.2	SOLUÇÃO RECOMENDADA
Crítica	Formulários sem rótulos *labels*	3.3.2 de Rótulos ou Instruções	Associar tag `<label>` aos campos `<input>` para identificação por leitores de tela.
Crítica	Contraste insuficiente em texto/fundo	1.4.3 sobre Contraste Mínimo	Ajustar CSS para atingir razão mínima de 4.5:1 de texto normal e 3:1 de texto grande.
Alta	Imagens úteis sem descrição	1.1.1 de Conteúdo Não Textual	Inserir atributo `alt` descritivo em imagens operacionais e `alt=` nas decorativas.
Média	Hierarquia de títulos imprecisos	1.3.1 de Informações e Relações	Reordenar níveis de cabeçalho `<h1>` a `<h6>` sucessivamente, sem pular níveis.

A. Conclusão

A acessibilidade digital deixa de ser um elemento opcional e passa a ser fundamental para a garantia do direito à educação para toda a sociedade. A análise feita em uma amostra de páginas do PA do Inatel mostra que o sistema não é funcional para todas as pessoas, impondo limitações relevantes, principalmente para PcDs. Os resultados obtidos através da ferramenta WAVE demonstraram que nenhuma das páginas selecionadas atingiu um nível de concordância exemplar. As violações mais frequentes encontradas são a falta de contraste, ausência de contextualização e a clareza de código e imagens sem texto alternativos. Essas infrações são falhas importantes e de alto impacto, segundo normas da WCAG.

Conclui-se que os ajustes do PA aos padrões da WCAG 2.2 é tecnicamente executável e de caráter imediato. Além de ser uma melhoria atraente para a instituição, é um passo necessário para que o Inatel confirme seu compromisso com a inclusão e a diversidade no ensino superior em engenharia e tecnologia.

B. Trabalhos futuros

Este estudo confirma as revelações dos autores de [11] e [12], provando que a irregularidade sistemática observada em portais de universidades federais também ocorre nos PAs de instituições de ensino privado. Dessa forma, trabalhos futuros com uma validação mais robusta devem ser considerados em todo o portal acadêmico.

VI. Agradecimentos

Gostaríamos de expressar nossa gratidão à Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG, ao Instituto Nacional de Telecomunicações - INATEL, e ao eHealth Innovation Center pelo apoio técnico e recursos viabilizados para o desenvolvimento deste projeto.

Referências

[1] Manuel Castells. *A sociedade em rede*. Vol. 1. A Era da Informação: Economia, Sociedade e Cultura. São Paulo: Paz e Terra, 1999.

[2] Brasil. *Lei nº 13.146, de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)*. Diário Oficial da União, Brasília, DF, 7 jul. 2015. 2015.

[3] World Wide Web Consortium (W3C). *Web Content Accessibility Guidelines (WCAG) 2.2*. https://www.w3.org/TR/WCAG22/. Out. de 2023.

[4] WebAIM. *WAVE Web Accessibility Evaluation Tool*. ttps://wave.webaim.org/. 2024.

[5] World Health Organization. *World Report on Disability*. Geneva: WHO, 2023.

[6] PARANÁ. *Deficiência Múltipla*. https://www.desenvolvimentosocial.pr.gov.br/Pagina/Deficiencia-Multipla. 2026.

[7] Brasil. Subsecretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência. *Comitê de Ajudas Técnicas: Tecnologia Assistiva*. Brasília, 2009.

[8] Ana Sofia Gala. *Tecnologias assistivas: o que são e exemplos*. https://www.handtalk.me/br/blog/tecnologias-assistivas/. Jun. de 2023.

[9] Google Developers. *Lighthouse Overview*. https://developer.chrome.com/docs/lighthouse/overview/. 2024.

[10] Deque Systems. *axe DevTools*. https://www.deque.com/axe/devtools/. 2024.

[11] Lucas M. M. Ribeiro e Marcela J. da Silva. “Análise de conformidade com acessibilidade digital em websites das universidades federais brasileiras à luz da e-MAG”. Em: *Anais do 11º Workshop de Computação Aplicada em Governo Eletrônico (WCGE)*. Belém: Sociedade Brasileira de Computação, 2019, pp. 80–89.

[12] Luan Soares Silva. “Acessibilidade web em portais de periódicos científicos de universidades federais brasileiras”. Diss. de mestr. Florianópolis: Universidade Federal de Santa Catarina, 2020, p. 154.

[13] Isabel Francine Mendes, Wajdi Aljedaani e Marcelo Medeiros Eler. “Exploring knowledge and practices in digital accessibility assessment in the Brazilian software industry”. Em: *Universal Access in the Information Society* 25.1 (2026), p. 6.

[14] Raul Sidnei Wazlawick. *Metodologia de pesquisa para ciência da computação*. 2ª ed. Rio de Janeiro: Elsevier, 2014.

[15] WebAIM. *Screen Reader User Survey #10 Results*. https://webaim.org/projects/screenreadersurvey10/. 2024.

AUTORES



Wagner Dourado Felix da Silva Possui formação técnica em Eletrônica pela Escola Técnica de Eletrônica Francisco Moreira da Costa (ETE FMC). Atualmente, é graduando em Engenharia de Computação pelo Instituto Nacional de Telecomunicações (INATEL), com previsão de conclusão em 2027. Atua como bolsista de Iniciação Científica no eHealth Innovation Center, onde desenvolve pesquisas voltadas à inovação tecnológica. Atua nas áreas de Acessibilidade Digital e Engenharia de Computação e com interesse em especialização em Inteligência Artificial.



Isabel Francine Mendes Doutoranda em Telecomunicações pelo Inatel (WAI-Lab), focada em soluções de visão computacional para sensoriamento de humanos e animais. Mestre pela USP em Sistemas de Informação com foco em acessibilidade digital.



Rita Elizabeth Santos de Almeida Possui formação técnica em eletrônica, com ênfase em equipamentos médicos, na Escola Técnica de Eletrônica Francisco Moreira da Costa - ETE FMC (2018) e graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações - Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center, centro de pesquisa de Engenharia Biomédica do Inatel.



Fabiano Valias de Carvalho Graduado em Engenharia Elétrica com ênfase em Telecomunicações e Eletrônica pelo Instituto Nacional de Telecomunicações (1995) e mestre em Engenharia Elétrica pela Universidade Federal de Itajubá (2002). Atua há mais de 30 anos no Inatel, onde exerce atividades docentes e de gestão acadêmica. Desde 2010, é coordenador do laboratório eHealth Innovation Center e do curso de Graduação em Engenharia Biomédica, tendo participado de sua concepção, estruturação e implantação, contribuindo para a consolidação e o reconhecimento nacional do curso.

Desenvolvimento de uma Ferramenta para Apoio aos Postos de Saúde da Família no Município de Santa Rita do Sapucaí

Luan Robert Aparecido Santos, Rita Elizabeth Santos de Almeida, Joyce Hemily Amorim Vieira,
Victor Arruda Gorgal, Fabiano Valias de Carvalho, Pedro Paulo de Souza Silva
Instituto Nacional de Telecomunicações - Inatel
luan.robert@ges.inatel.br, r.elizabeth@inatel.br, joyce.vieira@geb.inatel.br, victor.arruda@ges.inatel.br,
fabiano@inatel.br , pedrocirurgiao2@gmail.com

***Abstract*—Primary Health Care (PHC) in Brazil, mainly organized through the Family Health Strategy (FHS), faces challenges related to service organization, long waiting times, and communication difficulties between users and healthcare teams. This study presents the development of an integrated technological tool, available as a mobile application and web platform, designed for Family Health Programs (FHPs) in the municipality of Santa Rita do Sapucaí, Brazil. The solution includes patient, community health worker, and physician profiles, integrating appointment scheduling, communication, and notification functionalities. Functional tests indicated consistent system performance, highlighting its potential to support care organization and strengthen communication within PHC.**

***Index Terms*—Mobile Application, Primary Health Care, Family Health Unit.**

***Resumo*—A Atenção Primária à Saúde (APS), organizada no Brasil principalmente por meio da Estratégia Saúde da Família (ESF), enfrenta desafios relacionados à organização dos atendimentos, longas filas de espera e dificuldades de comunicação entre usuários e equipes de saúde. Este trabalho apresenta o desenvolvimento de uma ferramenta tecnológica integrada, acessível por aplicativo móvel e plataforma web, voltada ao Programa de Saúde da Família (PSF) do município de Santa Rita do Sapucaí (MG). A solução contempla perfis de paciente, agente comunitário de saúde e médico, integrando funcionalidades de agendamento, comunicação e notificações. Testes funcionais indicaram funcionamento consistente do sistema, com potencial para apoiar a organização dos atendimentos e fortalecer a comunicação na APS.**

***Palavras chave*—Aplicativo, Atenção Primária à Saúde, Posto de Saúde da Família.**

I. Introdução

A Estratégia de Saúde da Família (ESF) é o principal modelo de organização da Atenção Primária à Saúde (APS) no Brasil e representa a principal porta de entrada da população no Sistema Único de Saúde (SUS) [1]. A ESF é responsável por ações de promoção da saúde, prevenção de doenças e acompanhamento contínuo dos pacientes, especialmente em contextos de maior vulnerabilidade socioeconômica. Sabe-se, que a ampliação da cobertura da ESF contribui com a diminuição da desigualdade social e melhora os indicadores de saúde[2, 3].

Apesar de sua ampla cobertura, a APS enfrenta desafios relacionados ao acesso a consultas e procedimentos especializados após os encaminhamentos realizados pelas equipes do Programa Saúde da Família (PSF). Problemas relacionados ao mecanismos de encaminhamento e retorno entre os serviços, junto com longos tempos de espera e a distribuição desigual de serviços e especialistas, dificultam a continuidade do cuidado e reduzem a capacidade de resolução da atenção primária [4, 5]. Esses obstáculos contribuem para o agravamento de quadros clínicos e para o aumento da demanda por serviços de urgência [6].

Para aprimorar o acesso e a organização dos serviços na APS, o Ministério da Saúde implementou o e-SUS APS, uma plataforma nacional voltada à padronização dos registros de atendimentos [7]. Estudos indicam que sua utilização contribui para a melhoria da qualidade dos registros clínicos e para fortalecer o cuidado contínuo. No entanto, limitações na integração com os sistemas municipais de regulação, na capacitação dos profissionais e nas diferenças de infraestrutura tecnológica entre os municípios reduzem o impacto da ferramenta na redução do tempo de espera por consultas especializadas. [8, 9].

Outra iniciativa relevante é o programa Regula Mais Brasil, desenvolvido no âmbito do PROADI-SUS, que utiliza estratégias de telerregulação para análise e priorização dos encaminhamentos realizados pela APS [10]. A telerregulação pode ajudar a reduzir as filas de espera e os encaminhamentos desnecessários, aumentando a capacidade de resolução da atenção primária. No entanto, a efetividade dessas estratégias depende da adesão dos municípios, da disponibilidade de infraestrutura tecnológica e da integração com os prontuários eletrônicos locais, o que pode gerar diferenças na implementação entre as regiões.[11, 12].

Nesse contexto, foi proposto o desenvolvimento de uma ferramenta tecnológica integrada, acessível por meio de aplicativo móvel e plataforma *web*, voltada aos PSFs do município de Santa Rita do Sapucaí (MG). A solução teve como objetivo apoiar a organização dos atendimentos, reduzir filas e absenteísmo, fortalecer a comunicação entre pacientes e equipes de saúde e contribuir para a melhoria da gestão dos fluxos assistenciais na APS.

II. Metodologia

O presente estudo caracterizou-se como uma pesquisa aplicada, com o objetivo de desenvolver uma ferramenta tecnológica acessível por meio de aplicativo móvel e plataforma

web, possibilitando sua utilização em diferentes dispositivos e ambientes digitais. A pesquisa foi conduzida no contexto da saúde, com foco na APS, especificamente nos PSFs, tendo como cenário o município de Santa Rita do Sapucaí (MG). A metodologia adotada compreendeu o levantamento de requisitos e necessidades dos usuários, seguido pelo desenvolvimento da ferramenta proposta.

A. Levantamento de requisitos

O levantamento de requisitos foi realizado a partir de uma demanda apresentada por um médico da família. A problemática identificada incluiu a existência de longas filas de espera para atendimento, elevados índices de absenteísmo por parte dos pacientes e dificuldades no compartilhamento de informações entre os profissionais de saúde e a comunidade. De acordo com o profissional, tais fatores são associados a atrasos na procura por atendimento, insatisfação da população em relação aos serviços ofertados e fragilização do vínculo entre pacientes e equipes de saúde, além de limitações no acesso contínuo a informações em saúde. Esses aspectos foram relevantes para a definição dos requisitos funcionais da ferramenta desenvolvida.

Após a identificação da problemática, foi definido o escopo do projeto, que consistiu no desenvolvimento de uma ferramenta tecnológica de uso simplificado, estruturada a partir de três perfis de acesso: paciente, agente comunitário de saúde e médico responsável.

O perfil do paciente contemplou funcionalidades relacionadas ao acompanhamento do cuidado, ao acesso à informação e à comunicação com a equipe de saúde, incluindo agendamento de consultas, comunicação via chat com o médico, recebimento e visualização de notificações, bem como consulta à data e ao horário dos atendimentos agendados.

O perfil do agente comunitário de saúde incluiu funcionalidades voltadas ao suporte ao paciente e ao apoio às ações da equipe, tais como visualização do grupo de pacientes sob sua responsabilidade, envio de notificações individuais ou coletivas e criação de perfis de pacientes e médicos no sistema.

Por fim, o perfil do médico contemplou funcionalidades relacionadas à gestão das informações clínicas e à interação com pacientes, incluindo confirmação de agendamentos, visualização das consultas programadas, envio de notificações e comunicação por meio de chat.

Foi estabelecido, em conjunto com o profissional de saúde, que a criação das contas de acesso ao sistema seria realizada exclusivamente por um agente comunitário de saúde, com o objetivo de assegurar maior controle, confiabilidade e segurança das informações.

B. Arquitetura do sistema

Para a definição da arquitetura do sistema, as tecnologias selecionadas basearam-se nos critérios de portabilidade e escalabilidade. O *framework* Flutter, em conjunto com a linguagem Dart, foi utilizado para o desenvolvimento do *front-end* da aplicação, visando à construção de uma solução multiplataforma a partir de uma base de código única.

A biblioteca Riverpod foi empregada para estruturar o controle, a organização e o compartilhamento dos dados da aplicação de forma centralizada, permitindo o gerenciamento reativo do estado entre os diferentes componentes da interface. Além disso, o gerenciamento de dependências foi realizado por meio de *providers*, responsáveis por definir e controlar o acesso aos serviços da aplicação, bem como pela inicialização e pelo controle do ciclo de vida dos dados, evitando a manipulação direta do estado.

C. Infraestrutura e Persistência de Dados

Para o desenvolvimento do *back-end*, foi utilizada a plataforma Firebase, adotada como um modelo de *Back-end as a Service* (BaaS). Esse modelo foi empregado com o objetivo de disponibilizar infraestrutura e serviços necessários ao funcionamento da aplicação, incluindo autenticação de usuários, banco de dados em tempo real, armazenamento de arquivos e hospedagem, eliminando a necessidade de implementação de um servidor próprio. A persistência de dados foi estruturada no Cloud Firestore, um banco de dados NoSQL orientado a documentos, que organiza as informações em coleções e documentos hierarquicamente estruturados.

A Figura 1 apresenta a modelagem do banco de dados implementado no Cloud Firestore, estruturada em coleções independentes e relacionadas por meio de identificadores únicos. A organização adotada seguiu o modelo NoSQL orientado a documentos, de forma a atender aos requisitos de flexibilidade estrutural, escalabilidade e sincronização em tempo real exigidos pela aplicação.

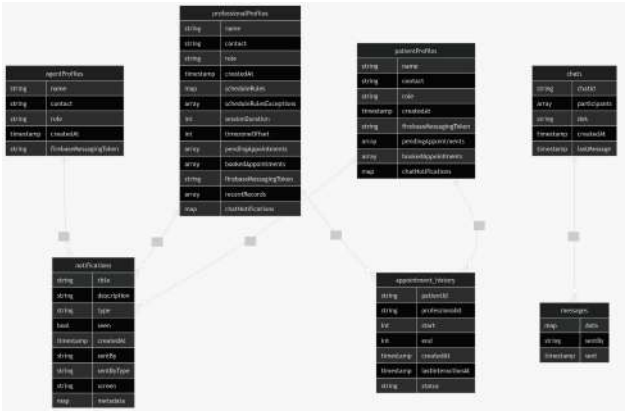


Fig. 1. Esquema do Firestore

A estrutura do banco de dados contemplou inicialmente coleções destinadas aos perfis de usuários: agentProfiles, professionalProfiles e patientProfiles. A coleção agentProfiles foi definida para armazenar os dados dos agentes comunitários de saúde, incluindo nome, informações de contato, função , data de criação do registro e *token* do Firebase Cloud Messaging, utilizado para o envio de notificações.

A coleção professionalProfiles foi destinada ao armazenamento das informações dos profissionais de saúde. Além dos dados cadastrais básicos, como nome, contato, função e data de criação, essa coleção incluiu campos específicos para o gerenciamento da agenda, como por exemplo regras de atendimento (scheduleRules), exceções de agenda (scheduleRulesExceptions), duração das consultas (sessionDuration) e fuso horário (timezoneOffset). Também foram definidos campos para o registro de consultas pendentes e confirmadas, bem como para

o armazenamento do *token* de notificações e controle de alertas de chat.

De forma semelhante, a coleção patientProfiles foi estruturada para armazenar os dados dos pacientes, incluindo informações cadastrais, *token* de notificação e listas de consultas pendentes e confirmadas, possibilitando o acompanhamento dos agendamentos e a comunicação com o profissional responsável.

O relacionamento entre pacientes e profissionais foi implementado por meio da coleção appointmentHistory, destinada ao registro das consultas realizadas ou agendadas. Cada documento dessa coleção foi definido contendo os identificadores do paciente (patientId) e do profissional (professionalId), além de informações temporais, como horário de início e término, data de criação, última interação e *status* da consulta.

A comunicação entre os usuários foi estruturada a partir da coleção chats, na qual cada documento representa uma conversa e armazena o identificador do chat, a lista de participantes, a data de criação e a última mensagem enviada. Associada a essa coleção, a subcoleção *messages* foi definida para o armazenamento individual das mensagens trocadas, contendo o conteúdo da mensagem, o identificador do remetente e o horário de envio.

Adicionalmente, foi implementada a coleção *notifications*, destinada ao gerenciamento das notificações enviadas pelo sistema. Cada documento dessa coleção foi estruturado contendo título, descrição, tipo, *status* de visualização, data de criação, identificação do remetente, tipo de remetente, tela de destino e metadados complementares.

A autenticação e o controle de acesso foram implementados por meio do Firebase Auth. Essa ferramenta foi empregada para validação de credenciais, gerenciamento de sessões e definição dos perfis de acesso.

Além disso, foram configuradas as *Security Rules* do Firebase, responsáveis por definir regras de leitura e escrita no banco de dados, restringindo o acesso às informações de acordo com o perfil do usuário (paciente, agente comunitário de saúde ou médico). Essas regras foram adotadas com o objetivo de limitar o acesso a dados sensíveis, como informações cadastrais, registros de conversas e dados de agendamento, em conformidade com os princípios da Lei Geral de Proteção de Dados (LGPD).

A funcionalidade de comunicação em tempo real entre médico e paciente foi implementada por meio da integração de *streams* da linguagem Dart com os *snapshots* do Cloud Firestore, possibilitando a atualização contínua das informações exibidas na interface. Para o envio de alertas e confirmações de consultas, foi utilizado o Firebase Cloud Messaging (FCM).

D. Projeto de Interface e Experiência do Usuário (UI/UX)

O projeto da interface e da experiência do usuário foi conduzido a partir de critérios de usabilidade e clareza visual. Para a elaboração dos *layouts* das telas, foram adotados critérios relacionados à padronização visual e organização da informação.

Entre os critérios considerados no projeto da interface, pode-se mencionar:

- Definição de paleta de cores com contraste adequado, composta predominantemente por branco, azul escuro e preto, utilização de tipografia sem serifa, com tamanho mínimo de 16px, visando à legibilidade das informações apresentadas.
- Estruturação de fluxos de navegação simplificados, com ênfase no processo de agendamento de consultas, organizado em etapas sequenciais e objetivas.
- Estabelecimento de hierarquia visual entre os diferentes perfis de usuário, com diferenciação das interfaces destinadas ao médico, ao agente comunitário de saúde e ao paciente, de acordo com as funcionalidades disponíveis para cada perfil.

No que se refere à estrutura da aplicação, foram definidas e implementadas as seguintes telas principais:

- **Tela de *Login***: destinada à autenticação do usuário, contendo campos para inserção de credenciais (e-mail e senha) e validação do acesso conforme o perfil previamente cadastrado no sistema.
- **Tela Inicial (*Dashboard*)**: responsável por apresentar as funcionalidades disponíveis para cada perfil de usuário, organizadas por meio de botões ou cartões de acesso.
- **Tela de Agendamento de Consultas (Paciente)**: desenvolvida para permitir a seleção de datas e horários disponíveis, a confirmação de agendamentos, a visualização de consultas previamente marcadas, o acesso ao canal de comunicação por chat com o médico e a visualização de notificações recebidas.
- **Tela de Gestão de Consultas (Médico)**: estruturada para possibilitar a visualização das consultas agendadas, a confirmação de atendimentos, o gerenciamento de horários disponíveis, a comunicação via chat com pacientes e o envio de notificações.
- **Tela de Gerenciamento de Pacientes (Agente Comunitário de Saúde)**: destinada à visualização dos pacientes sob responsabilidade do agente, à criação de cadastros e ao envio de notificações individuais ou em grupo.

III. Resultados e Discussões

O desenvolvimento da aplicação resultou em uma plataforma funcional acessível por meio de aplicativo móvel e versão *web*, contemplando os três perfis definidos na fase de levantamento de requisitos: paciente, agente comunitário de saúde e médico.

A aplicação foi submetida a testes iniciais, de integração, segurança e usabilidade, realizados em ambiente local por meio de emulador Android Studio, navegador *web* e dispositivos físicos. Esses testes tiveram como objetivo verificar o correto funcionamento dos principais fluxos do sistema, a integração entre o *front-end* desenvolvido em Flutter e os serviços do Firebase, bem como o controle de acesso entre os diferentes perfis de usuários.

As funcionalidades avaliadas incluíram criação de contas, autenticação, agendamento de consultas, comunicação via chat e envio de notificações. A validação contemplou ainda a aplicação das *Security Rules* do Cloud Firestore, por meio

da simulação de tentativas de acesso indevido entre perfis distintos.

Adicionalmente, foram conduzidos testes de usabilidade com integrantes da equipe, considerando aspectos como clareza das interfaces, tempo de resposta e facilidade de navegação, com o objetivo de avaliar a interação do usuário com a aplicação.

A partir dessas avaliações, os principais módulos do sistema são apresentados e discutidos a seguir.

A. *Gerenciamento de Consultas*

O módulo de gerenciamento de consultas permitiu ao paciente realizar agendamentos por meio da seleção de datas e horários disponíveis. Após a confirmação, as informações foram armazenadas no Cloud Firestore e sincronizadas com o perfil do médico responsável.

A sincronização dos dados possibilitou que o médico visualizasse as consultas agendadas em sua agenda, organizada cronologicamente, permitindo a visualização estruturada dos atendimentos programados.

Para o agente comunitário de saúde, o sistema permitiu uma visão dos pacientes sob sua responsabilidade, incluindo informações sobre consultas agendadas, viabilizando o acompanhamento dos atendimentos previstos.

As Figuras 2 e 3 apresentam as telas de agendamento disponibilizadas para os perfis de paciente e médico, respectivamente.



Fig. 2. Tela de agendamento – Paciente

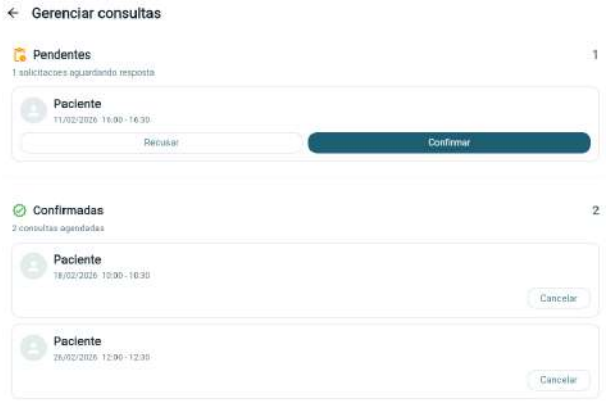


Fig. 3. Tela de agendamento – Médico

B. *Chat em Tempo Real*

A funcionalidade de chat permitiu a troca de mensagens em tempo real entre os usuários do sistema, com atualização automática da interface a cada nova mensagem registrada no banco de dados.

Durante o envio das mensagens, o conteúdo foi armazenado no Cloud Firestore e disponibilizado ao destinatário sem a necessidade de recarregamento da interface. Essa exibição instantânea foi possível devido ao mecanismo de escuta contínua dos dados.

A interface do chat apresentou organização visual das mensagens, com diferenciação clara entre remetente e destinatário, conforme ilustrado na Figura 4.



Fig. 4. Interface do chat em tempo real

C. *Sistema de Notificações*

O sistema de notificações *push* foi implementado por meio do Firebase Cloud Messaging (FCM), possibilitando o envio de lembretes e avisos aos usuários da aplicação.

Durante os testes realizados, as notificações foram entregues corretamente aos dispositivos cadastrados, sem atrasos significativos entre o envio e o recebimento. Os avisos foram exibidos tanto para confirmações de consultas quanto para comunicações informativas.

As Figuras 5 e 6 apresentam, respectivamente, as interfaces de envio e recebimento das notificações.



Fig. 5. Interface de envio de notificações

D. *Estrutura de Dados e Segurança*

A estrutura do banco de dados foi organizada em coleções distintas para usuários, consultas, mensagens e notificações, permitindo separação lógica das informações e controle de acesso por meio de identificadores únicos.

As *Security Rules* do Firebase foram configuradas para restringir o acesso aos dados conforme o perfil do usuário autenticado. Durante os testes, verificou-se que pacientes tiveram



Fig. 6. Interface de recebimento de notificações

acesso apenas aos seus próprios registros e que as conversas no chat permaneceram restritas aos respectivos pares médico–paciente.

Todo o tráfego de dados entre cliente e servidor ocorreu por meio do protocolo HTTPS, e os dados armazenados permaneceram criptografados nos servidores do Firebase, assegurando proteção durante a transmissão e o armazenamento.

E. Sistema de Autenticação

O mecanismo de autenticação foi implementado utilizando o Firebase Authentication, permitindo validação segura das credenciais por meio de e-mail e senha.

Após a autenticação bem-sucedida, o sistema identificou o perfil do usuário e realizou o redirecionamento automático para a interface correspondente. Em casos de falha na autenticação, o acesso às funcionalidades permaneceu bloqueado.

A interface da tela de *login* é apresentada na Figura 7.



Fig. 7. Tela de login

IV. Conclusão

No contexto da APS foi identificado algumas problemáticas relacionadas as longas filas de espera, ao absenteísmo e às dificuldades de comunicação entre pacientes e profissionais de saúde do PSF. Essas limitações impactam a organização dos fluxos assistenciais e a continuidade do cuidado, especialmente em municípios de pequeno porte e com restrições de infraestrutura e recursos especializados.

Com o objetivo de contribuir para esse cenário, foi desenvolvido uma ferramenta tecnológica acessível por meio de aplicativo móvel e versão *web*, voltada aos PSFs do município de Santa Rita do Sapucaí (MG). A solução contempla três perfis de acesso: paciente, agente comunitário de saúde e médico, e incorpora funcionalidades como autenticação de usuários, agendamento e confirmação de consultas, comunicação por chat em tempo real e envio de notificações e informativos educativos.

A arquitetura adotada, baseada no uso do *framework* Flutter para o desenvolvimento da interface e a utilização do Firebase para autenticação, persistência de dados e notificações, possibilitou a sincronização das informações em tempo real e o controle de acesso conforme os diferentes perfis de usuários. Os testes realizados indicaram funcionamento consistente dos principais fluxos da aplicação, com adequada integração entre os módulos implementados e aplicação correta das regras de segurança.

Assim, o sistema desenvolvido mostrou potencial como ferramenta de apoio à organização dos atendimentos e à comunicação entre usuários e equipes de saúde na APS. Como trabalhos futuros, sugere-se ampliar as funcionalidades da plataforma, incluindo um sistema de controle de filas para redistribuição de vagas remanescentes, além da realização de estudos com usuários em ambiente real de uso, a fim de avaliar o impacto da ferramenta na rotina dos serviços de saúde.

V. Agradecimentos

Gostaríamos de expressar nossa gratidão à Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG, ao Instituto Nacional de Telecomunicações - INATEL, e ao eHealth Innovation Center pelo apoio técnico e recursos que viabilizaram para o desenvolvimento deste projeto.

Referências

[1] Brasil. Ministério da Saúde. *Atenção Primária à Saúde*. Acesso em: 24 fev. 2026. 2023. URL: https://www.gov.br/saude/pt-br/composicao/saps.

[2] Barbara Starfield. “Primary care: an increasingly important contributor to effectiveness, equity, and efficiency of health services. SESPAS report 2012”. Em: *Gaceta Sanitaria* 26 (2012), pp. 20–26. ISSN: 0213-9111. DOI: https://doi.org/10.1016/j.gaceta.2011.10.009. URL: https://www.sciencedirect.com/science/article/pii/S0213911111003876.

[3] James Macinko, Frederico C. Guanais e Maria de Fátima Marinho de Souza. “Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002”. Em: *Journal of Epidemiology & Community Health* 60.1 (2006), pp. 13–19. DOI: 10.1136/jech.2005.038323. URL: https://jech.bmj.com/content/60/1/13.

[4] Ligia Giovanella, Cassia Maria Franco e Patty Fidelis Almeida. “Política nacional de atenção básica: para onde vamos?” Em: *Cadernos de Saúde Pública* 36.8 (2020), e00077020. DOI: 10.1590/0102-311X00077020. URL: https://www.scielo.br/j/csp/.

[5] João Ricardo Silva, Maria Alice Santos e Paula Regina Souza. “Tempo de espera para acesso a serviços especializados no Sistema Único de Saúde”. Em: *Revista de Saúde Pública* 55 (2021), p. 87. DOI: 10.11606/s1518-8787.2021055003415. URL: https://www.scielo.br/j/rsp/.

[6] Owen O’Donnell, Eddy van Doorslaer, Ravi Rannan-Eliya et al. “Access to health care in developing countries: breaking down demand side barriers”. Em: *BMC Health Services Research* 7 (2007), p. 14. DOI: 10.1186/1472-6963-7-14. URL: https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-7-14.

[7] Brasil. Ministério da Saúde. *e-SUS Atenção Primária à Saúde (e-SUS APS)*. Sistema de informação em saúde. Acesso em: 24 fev. 2026. 2023. URL: https://sisaps.saude.gov.br/esus/.

[8] Taciana Silveira Passos, Tiago Machado de Alencar, Alessandra Page Brito, Ayrton Martins Pereira Sena, Débora Bleza Santos, Agilran Araújo Barreto e Lídia Glasielle de Oliveira Silva. "Utilização dos relatórios do e-SUS da Atenção Primária à Saúde na rotina de trabalho de equipes de Saúde da Família". Em: *Saúde em Debate* 48.Especial 2 (2024). Acesso em: 24 fev. 2026, e8917. URL: https://www.saudeemdebate.org.br/sed/article/view/8917.

[9] Clemens S. Kruse, Charles Kristof, Ben Jones, Edward Mitchell e Angelica Martinez. "Evaluating barriers to adopting health information technology". Em: *BMC Medical Informatics and Decision Making* 18 (2018), pp. 1–14. DOI: 10.1186/s12911-018-0669-6. URL: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-018-0669-6.

[10] Brasil. Ministério da Saúde. *Regula Mais Brasil*. Programa no âmbito do PROADI-SUS. Acesso em: 24 fev. 2026. 2023. URL: https://www.gov.br/saude/pt-br/assuntos/proadi-sus.

[11] C. D. Souza et al. "Telerregulação no Sistema Único de Saúde". Em: *Ciência & Saúde Coletiva* (2022). URL: https://www.scielo.br/j/csc/.

[12] D. O. Franco et al. "Regulation and telehealth challenges in health systems". Em: *BMC Health Services Research* (2021). URL: https://bmchealthservres.biomedcentral.com/.

AUTORES



Luan Robert Aparecido Santos Graduando em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel). Atualmente, é voluntário de Iniciação Científica no eHealth Innovation Center. Atua profissionalmente como Engenheiro de Software na Levty e é co-founder da BullVest, startup voltada ao desenvolvimento de um sistema de controle financeiro, gamificado com simulação e estudo.



Rita Elizabeth Santos de Almeida Possui graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações - Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center, centro de pesquisa de Engenharia Biomédica do Inatel.



Joyce Hemily Amorim Vieira Graduanda de Engenharia Biomédica no Inatel, técnica em Informática e estagiária do FabLab Inatel, com atuação em atividades de laboratório e suporte a projetos de fabricação digital.



Victor Arruda Gorgal Graduando em Engenharia de Software no Instituto Nacional de Telecomunicações.



Fabiano Valias de Carvalho Graduado em Engenharia Elétrica com ênfase em Telecomunicações e Eletrônica pelo Instituto Nacional de Telecomunicações (1995) e mestre em Engenharia Elétrica pela Universidade Federal de Itajubá (2002). Desde 2010, é coordenador do laboratório eHealth Innovation Center e do curso de Graduação em Engenharia Biomédica, tendo participado de sua concepção, estruturação e implantação, contribuindo para a consolidação e o reconhecimento nacional do curso.



Pedro Paulo de Souza Silva Médico de família, graduado em Medicina pela Fundação Tricordiana de Educação (2013), com residência em Cirurgia Geral pela Fundação Hospital Deraldo Guimarães (2016) e especialização em Psiquiatria pela Faculdade Unyleya. Possui graduações prévias em Fisioterapia e Enfermagem e experiência como professor universitário nas áreas de bioquímica, fisiologia e farmacologia.

Classificação de Imagens Colposcópicas com Redes Neurais Convolucionais para Detecção de Lesões Cervicais

Gislaine Abreu Mendonça, Lívia Toledo Gandini, José Andery Carneiro, Márcio Erik Franco & Paulo Roberto Maia
Instituto Nacional de Telecomunicações - Inatel
gislaine.abreu@geb.inatel.br, livia.toledo@gec.inatel.br, jose.andery@inatel.br, marcioerikfr.med@hotmail.com & paulo.maia@inatel.br

***Abstract*—Early detection of cervical cancer through colposcopy is subject to individual judgment during clinical evaluation. This study investigates the use of Artificial Intelligence as a diagnostic support tool for colposcopic examination. A model based on the ResNet-18 architecture was used to classify colposcopic images into four diagnostic categories: normal, LSIL, HSIL, and cancer. Public databases were used for training, applying *transfer learning*, *data augmentation* techniques, and class balancing strategies such as Focal Loss. The objective is to develop a more accurate and accessible diagnostic support tool for medical practice.**

***Index Terms*—Cervical Cancer, Colposcopy, Convolutional Neural Networks, ResNet-18.**

***Resumo*—A detecção precoce do câncer de colo do útero por colposcopia está sujeita ao julgamento individual durante a avaliação clínica. Este trabalho investiga o uso de Inteligência Artificial como ferramenta de apoio ao diagnóstico colposcópico. Foi utilizado um modelo baseado na arquitetura ResNet-18 para classificar imagens colposcópicas em quatro categorias diagnósticas: normal, LSIL, HSIL e câncer. Para o treinamento, foram utilizadas bases de dados públicas, aplicando técnicas de *transfer learning*, *data augmentation* e estratégias de balanceamento de classes, como Focal Loss. O objetivo é desenvolver uma ferramenta de apoio ao diagnóstico mais precisa e acessível para a prática médica.**

***Palavras chave*—Câncer de colo de útero, Colposcopia, Redes Neurais Convolucionais, Resnet-18.**

I. Introdução

O câncer do colo do útero permanece como uma das principais causas de morbimortalidade feminina, sobretudo em países de baixa e média renda, onde barreiras ao diagnóstico e ao tratamento são frequentes [1, 2]. Além do câncer, a infecção pelo Vírus da Imunodeficiência Humana (HIV) também apresenta alta prevalência, com as mulheres representando cerca de metade da população mundial infectada [1]. Fatores biológicos e sociais como a anatomia do trato reprodutor, desigualdade de gênero, acesso limitado a métodos de prevenção e condições socioeconômicas desfavoráveis aumentam o risco de transmissão e dificultam o acesso à profilaxia pré-exposição (PrEP) e a terapia antirretroviral (TARV) [2, 3].

O diagnóstico precoce é, portanto, fundamental para reduzir a morbimortalidade associada a ambas as condições. Embora o rastreamento do câncer do colo do útero seja tradicionalmente realizado pelo exame citopatológico de Papanicolau, a colposcopia destaca-se como método essencial para a detecção de lesões intraepiteliais escamosas precursoras do câncer, apresentando melhor desempenho, especialmente em mulheres vivendo com HIV, que possuem maior risco de progressão para lesões de alto grau (CIN2+) [4]. Entretanto, trata-se de um exame dependente do operador, cuja acurácia é influenciada pela subjetividade do julgamento clínico, resultando em variabilidade diagnóstica e maior probabilidade de resultados falso-negativos e falso-positivos, particularmente em regiões com escassez de especialistas [5, 1].

Diante dessas limitações, a integração de técnicas de Inteligência Artificial à colposcopia surge como alternativa promissora para padronizar avaliações e ampliar o acesso a diagnósticos mais precisos. Métodos baseados em Redes Neurais Convolucionais (CNNs, do inglês Convolutional Neural Networks), como o Avaliação *Visual Automatizada* (AVE, do inglês Automated Visual Evaluation), têm apresentado resultados encorajadores na diferenciação entre lesões de baixo e alto grau, embora ainda enfrentem desafios relacionados à limitação geográfica dos bancos de imagens e à escassez de validações clínicas [5, 6, 7].

Diversos estudos têm investigado a aplicação de técnicas de Inteligência Artificial na análise de imagens colposcópicas. Métodos baseados em CNNs têm apresentado resultados promissores na detecção automática de lesões cervicais [8, 9]. O trabalho de Hu et al. [9] propôs um sistema de AVE, alcançando área sob a curva de 0,91 na detecção de lesões precancerígenas. Outros estudos também exploram arquiteturas profundas como ResNet e Faster R-CNN, demonstrando desempenho comparável ou superior ao de especialistas humanos na análise de imagens médicas [10, 11, 9]. Apesar dos avanços, desafios como variabilidade dos dados, limitações geográficas dos bancos de imagens e escassez de validação clínica ainda persistem [5, 9], motivando o desenvolvimento de novas abordagens mais robustas e generalizáveis.

Assim, este estudo tem como objetivo desenvolver um modelo de Inteligência Artificial capaz de classificar imagens colposcópicas em quatro classes diagnósticas: normal, lesão intraepitelial de baixo grau (LSIL), lesão intraepitelial de alto grau (HSIL) e câncer, utilizando bancos de dados públicos, técnicas de aumento de dados e estratégias de otimização do treinamento, a fim de ampliar a capacidade de generalização do modelo e contribuir para um diagnóstico mais preciso e acessível, especialmente em regiões com recursos limitados.

II. Referencial Teórico

A. Redes Neurais Convolucionais e ResNet-18

As CNNS tem sido amplamente utilizadas na análise de imagens médicas, aplicadas em tarefas como detecção de doenças, segmentação de artefatos de interesse e melhoria da qualidade de imagens em áreas como radiologia, dermatologia e patologia [8]. Essas redes são projetadas para lidar com dados com estrutura espacial, extraindo automaticamente características relevantes por meio de filtros calibrados durante o treinamento da própria rede no contexto específico da aplicação [8, 10].

Uma CNN é composta principalmente por camadas convolucionais, de *pooling* (camada de agrupamento/subamostragem) e totalmente conectadas, como pode ser visto na Figura 1. A camada convolucional é responsável por aplicar filtros que identificam padrões como bordas, texturas e formas, gerando mapas de características [8]. A camada de *pooling* reduz a dimensionalidade desses mapas, diminuindo o custo computacional, sendo o *max pooling* (uma operação que reduz o tamanho da imagem na rede neural) a operação mais utilizada [8]. Já a camada totalmente conectada, comumente a última da rede, integra todas as características extraídas e propagadas até esta etapa e realiza a decisão final do modelo, produzindo como saída a atribuição de uma classe (classificação) com sua respectiva probabilidade ou nível de confiança, geralmente obtido por funções como *softmax* (última camada da rede), ou a estimativa de um valor contínuo (regressão), dependendo do objetivo do problema.

Ao longo do tempo, diferentes arquiteturas foram desenvolvidas, como LeNet, VGGNet, ResNet e ConvNeXt [8, 10]. A ResNet, proposta por He et al. (2015), introduziu o aprendizado residual por meio das skip connections, permitindo o treinamento de redes mais profundas e reduzindo o problema do desaparecimento do gradiente [11].

Estudos recentes indicam que as ResNets continuam apresentando excelente equilíbrio entre desempenho e custo computacional, além de robustez e estabilidade de treinamento, o que justifica sua ampla aplicação em imagens médicas [12].

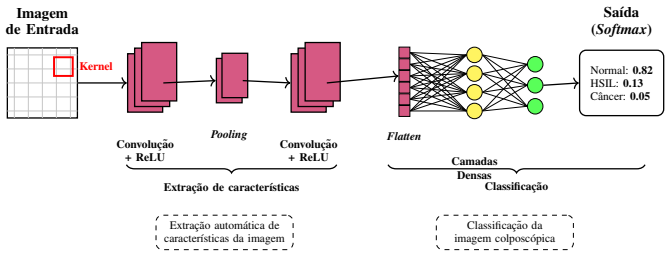


Fig. 1. Arquitetura simplificada de uma Rede Neural Convolucional (CNN) aplicada à classificação de imagens colposcópicas.

B. HPV e Câncer de Colo do Útero

O câncer do colo do útero é uma das neoplasias ginecológicas (isto é, tumores formados pela proliferação anormal e descontrolada de células) mais prevalentes mundialmente, com maior incidência em países em desenvolvimento, estando fortemente associado à infecção persistente pelo Papilomavírus Humano (HPV) [1]. O HPV é um vírus de DNA da família Papillomaviridae, com mais de 200 genótipos identificados, sendo cerca de 14 classificados como de alto risco oncogênico. Os tipos HPV-16 e HPV-18 são responsáveis por aproximadamente 70% dos casos de câncer [13]. Embora a maioria das infecções seja eliminada espontaneamente, a persistência viral pode levar ao desenvolvimento da Neoplasia Intraepitelial Cervical (CIN) e, posteriormente, ao carcinoma invasivo [13]. O avanço da doença ocorre porque o vírus produz duas proteínas nocivas (E6 e E7) que desregulam a multiplicação das células e causam danos ao seu DNA. São exatamente essas alterações celulares que os exames preventivos buscam identificar para realizar o diagnóstico precoce [6].

b.1) Exame de Colposcopia

A colposcopia é um exame complementar utilizado para avaliação ampliada do epitélio cervical, permitindo a identificação de lesões precursoras do câncer [7]. O procedimento envolve a aplicação de ácido acético (3–5%) e solução de Lugol, que evidenciam alterações epiteliais por contraste de coloração [7]. Entretanto, as imagens colposcópicas apresentam elevada variabilidade devido a fatores como iluminação, aquisição e características anatômicas. A Figura 2 apresenta exemplos de imagens geradas por esse exame. Além disso, a interpretação é dependente do examinador, podendo ocorrer variações interobservador e intra-observador, o que impacta a acurácia diagnóstica [14].

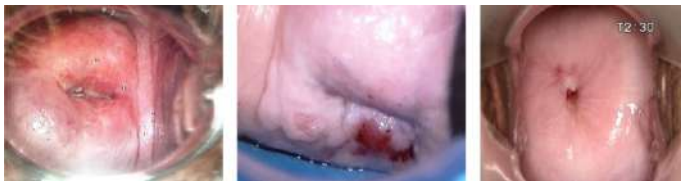


Fig. 2. Exemplos de imagens colposcópicas do acervo dos autores.

C. Aplicação de Métodos Computacionais na Colposcopia

O avanço das tecnologias de processamento de imagens e do aprendizado de máquina possibilitou o desenvolvimento de sistemas de apoio ao diagnóstico em colposcopia [9]. Sob a perspectiva da engenharia, o problema pode ser formulado como uma tarefa de classificação de imagens, com o objetivo de distinguir tecidos normais de lesões cervicais associadas ao HPV. Devido à complexidade visual dessas imagens, as CNNs tem demonstrado desempenho superior aos métodos tradicionais de extração manual de características. Como exemplo prático, o estudo conduzido por Hu et al. [9] desenvolveu um AVE utilizando a arquitetura Rede Neural Convolucional baseada em regiões mais rápida (Faster R-CNN, do inglês Region-based Convolutional Neural Network). Os resultados experimentais indicaram que a rede atingiu uma área sob a curva de 0,91 na detecção de lesões pré-cancerígenas e câncer, superando a interpretação original feita por especialistas humanos 0,69, no mesmo cenário de avaliação. Esses achados evidenciam a capacidade dos modelos baseados em deep learning de atingir precisão comparável ou superior à de colposcopistas experientes, reforçando seu potencial como ferramenta robusta de apoio à decisão clínica [9].

III. Metodologia

Inicialmente, foi explorado o dataset AnnoCerv, publicado no final de 2023. Este banco de dados é composto por 527 imagens de colposcopia cervical, extraídas de 100 prontuários médicos de pacientes da Romênia, tendo sido desenvolvido por pesquisadores da Universidade de Medicina e Farmácia Grigore T. Popa e da Universidade Sapientia. As imagens são acompanhadas de máscaras de segmentação detalhadas, que foram anotadas e validadas por médicos especialistas em ginecologia e colposcopia, visando diferenciar tecidos cervicais saudáveis de alterações patológicas. Os dados foram devidamente anonimizados em conformidade com as diretrizes de ética médica para uso retrospectivo de prontuários, sendo disponibilizados publicamente para pesquisa sob licença Creative Commons. O conjunto apresenta grande variabilidade de iluminação, contraste e características anatômicas incluindo capturas originais com ácido acético, solução de iodo e filtro verde, o que representa um desafio computacional adicional e realista para modelos de classificação.

Para viabilizar o treinamento supervisionado, realizou-se a reorganização das imagens com base nas máscaras: imagens contendo pixels positivos foram rotuladas como "lesão" (totalizando 121 imagens), enquanto as demais foram classificadas como "normal" (85 imagens). Também foram conduzidos testes considerando as categorias subdivididas, cuja distribuição consistiu em 27 imagens para câncer, 59 para HSIL e 35 para LSIL. Para a realização dos experimentos, o banco de dados foi dividido em 80% das imagens para a etapa de treinamento e 20% para a etapa de validação (teste). Entretanto, apesar da relevância do dataset, os resultados obtidos não apresentaram desempenho satisfatório para a tarefa proposta, motivando a ampliação da base de dados utilizada no estudo.

Dessa forma, foi desenvolvido um modelo de Inteligência Artificial para reconhecimento de níveis de lesão escamosa intraepitelial a partir de imagens colposcópicas. Foram selecionados quatro bancos de dados com base em critérios de adequação metodológica: MSCI - A Multistate Dataset for Colposcopy Image Screening, ShuffleNet Attention Extend, IARC Colposcopy Image Bank e IARC Visual Inspection with Acetic Acid (VIA) Image Bank. O número de imagens provenientes de cada base foi, respectivamente, 28, 80, 565 e 324. O conjunto consolidado foi dividido em 80% para treinamento (901 imagens) e 20% para validação (206 imagens).

O pré-processamento incluiu padronização automática de orientação, redimensionamento para 256×256 pixels e aplicação de técnicas de aumento de dados, com o objetivo de ampliar a variabilidade e reduzir o overfitting. As estratégias de aumento de dados incluíram rotações horizontais e angulares (-10° a 10°), variações de brilho (-20% a 20%), ajustes de tonalidade e saturação (-10% a 10%), adição de ruído e borramento (até 0,5%) e recortes com zoom de até 20%. Para cada imagem original, foram geradas duas novas sintéticas, com aplicação de uma ou mais das estratégias descritas acima, com as escolhas da técnica e do seu respectivo parâmetro feitas de forma aleatória. A arquitetura escolhida foi a ResNet-18, pré-treinada por 90 épocas no ImageNet-1K, devido ao seu bom equilíbrio entre desempenho e custo computacional. Foi adotada a estratégia de *transfer learning*, mantendo-se congeladas as camadas convolucionais iniciais e ajustando-se as camadas finais para a tarefa de classificação. A camada totalmente conectada foi atualizada para uma nova camada compatível com o número de classes do problema. O treinamento foi realizado em máquina virtual (Google Colaboratory) com linguagem Python, utilizando a biblioteca PyTorch.

Como técnicas de otimização do treinamento da rede, foram utilizados:

- **Cross-Entropy Loss:** para medir a discrepância entre predições e rótulos reais;
- **Adam Optimizer:** para ajuste adaptativo da taxa de aprendizado;
- **Focal Loss:** para mitigar o desbalanceamento entre classes;
- **Weighted Random Sampler:** para balanceamento durante o treinamento.

O processo experimental foi conduzido em duas etapas: inicialmente, avaliou-se o desempenho do modelo em diferentes números de épocas sem aumento de dados; posteriormente, foram testadas diferentes combinações de técnicas de aumento de dados, visando identificar a configuração que proporciona melhor desempenho e maior capacidade de generalização.

IV. Resultados e Discussão

A. Testes com o banco AnnoCerv

A Tabela I apresenta os resultados obtidos com o banco AnnoCerv. Após 10 épocas de treinamento, o modelo atingiu acurácia global de 0,3884 e uma média ponderada de F1-score entre as classes de 0,3153, indicando desempenho limitado na classificação multiclasse.

A classe Câncer apresentou alta precisão (1,0000), porém recall muito baixo (0,0370), evidenciando dificuldade na detecção efetiva dos casos positivos. A classe HSIL obteve métricas próximas ao aleatório, enquanto a LSIL não foi identificada em nenhum caso. A classe Normal apresentou melhor desempenho (recall de 0,7294), sugerindo viés para a classe majoritária.

Os resultados indicam impacto significativo do desbalanceamento e da variabilidade das imagens, comprometendo a capacidade de generalização do modelo e justificando a adoção de estratégias e bases complementares nas etapas seguintes.

TABELA I
Resultados do Teste com o banco AnnoCerv

Classe	Precisão	Recall	F1-Score	Número real de amostras
Câncer	1.0000	0.0370	0.0714	27
HSIL	0.2615	0.2881	0.2742	59
LSIL	0.0000	0.0000	0.0000	35
Normal	0.4429	0.7294	0.5511	85

B. Testes com o banco montado I

O primeiro treinamento foi realizado sem a utilização de técnicas de *data augmentation*, com o objetivo de estabelecer

um cenário base para analisar o comportamento do modelo e identificar em qual número de épocas são obtidas as melhores métricas de desempenho. Na Tabela II pode ser observado que, ao longo do treinamento, o desempenho no conjunto de treino decresce com o aumento do número de épocas, enquanto as métricas no conjunto de validação apresentam melhora progressiva, indicando maior capacidade de generalização. Com base na análise das métricas de validação, verificou-se que o melhor desempenho foi obtido dentro de 100 épocas, dentre os valores avaliados, sendo este adotado como o número ideal de épocas para os experimentos subsequentes. Esse experimento permitiu compreender a dinâmica de aprendizado do modelo em um cenário sem aumento dos dados, servindo como referência para quantas épocas usar em treinamentos subsequentes e demais comparações.

C. Testes com o banco montado II

O segundo treinamento, apresentado na Tabela III, buscou comparar o desempenho do modelo sob diferentes estratégias de aumento de dados. Cada versão explorou diferentes configurações de rotação, exposição, saturação, brilho e ruído.

A análise paralela das Tabelas III e IV evidencia que o desempenho global do modelo não reflete, de forma isolada, sua capacidade em identificar as diferentes classes clínicas. Observa-se uma tendência consistente do modelo em reconhecer com maior precisão padrões associados a lesões de maior gravidade, especialmente a classe Câncer, que apresentou os maiores valores de F1-score entre as versões avaliadas, atingindo 0,75 na versão 02. A classe CIN 2 também demonstrou desempenho relativamente superior quando comparada às categorias de menor gravidade, Normal e CIN 1, que apresentaram métricas inferiores e maior instabilidade entre as versões. Destaca-se a redução expressiva do recall para a classe Normal nas versões que empregaram maior número de técnicas de aumento de dados. Esse comportamento sugere uma tendência do modelo a superestimar as lesões, possivelmente influenciada por desbalanceamento de amostras ou maior complexidade na delimitação dos padrões de decisão entre tecidos normais e alterações de baixo grau.

A quarta versão apresentou o melhor desempenho global de validação (55%), além da maior acurácia no treinamento (95%), indicando elevada capacidade de ajuste aos dados. Embora a validação global tenha sido superior, a distribuição das métricas por classe mostrou desempenho menos equilibrado quando comparado à versão 02. A segunda versão, por sua vez, apresentou métricas globais mais consistentes entre treino e validação. Além disso, obteve o melhor desempenho na identificação da classe Câncer e desempenho semelhante para CIN 2, categorias consideradas de maior impacto direto na conduta terapêutica. Esse resultado indica que, apesar de não apresentar a maior acurácia global, a versão 02 demonstrou maior sensibilidade e equilíbrio na detecção das lesões de maior gravidade. Dessa forma, para a definição da configuração final, foram considerados o desempenho global, a capacidade de generalização, o equilíbrio entre as métricas de treinamento e validação e, principalmente, a capacidade de identificação das classes clinicamente mais críticas. Com base nesses critérios, optou-se pela segunda versão do modelo, por apresentar desempenho consistente e equilíbrio entre os requisitos estabelecidos. Além disso, esses achados evidenciam que técnicas de aumento de dados podem beneficiar classes específicas, mas não garantem melhorias uniformes, especialmente em cenários de dados limitados e desbalanceados.

TABELA II
RESULTADO TESTE BANCO DE DADOS MONTADO I

Número de épocas		Loss	Accuracy	Recall
10	Treino	0.78	0.85	0.75
	Validação	1.02	0.47	0.50
25	Treino	1.17	0.67	0.45
	Validação	0.85	0.52	0.65
100	Treino	1.45	0.45	0.4
	Validação	0.85	0.6	0.7

TABELA III
RESULTADO TESTE BANCO DE DADOS MONTADO II

Versão		Loss	Accuracy
01	Treino	0.2	0.8
	Validação	0.7	0.45
02	Treino	0.4	0.65
	Validação	0.65	0.5
03	Treino	0.1	0.85
	Validação	0.65	0.45
04	Treino	0.1	0.95
	Validação	0.65	0.55

TABELA IV
RESULTADO TESTE BANCO DE DADOS MONTADO II - MÉTRICAS DETALHADAS

Versão	Métricas	Normal	CIN 1	CIN 2	Câncer
01	Precision	0.38	0.38	0.4	0.8
	Recall	0.65	0.65	0.5	0.45
	F1-Score	0.55	0.55	0.48	0.58
02	Precision	0.43	0.55	0.44	0.75
	Recall	0.65	0.43	0.4	0.75
	F1-Score	0.57	0.45	0.42	0.75
03	Precision	0.38	0.4	0.4	0.75
	Recall	0.25	0.65	0.37	0.55
	F1-Score	0.3	0.45	0.38	0.7
04	Precision	0.35	0.4	0.43	0.75
	Recall	0.42	0.6	0.55	0.5
	F1-Score	0.37	0.45	0.5	0.55

V. CONCLUSÃO

O presente artigo teve como objetivo o desenvolvimento de um modelo de Inteligência Artificial capaz de classificar imagens colposcópicas em quatro classes diagnósticas: normal, LSIL, HSIL e câncer, utilizando bancos de dados open source,

técnicas de aumento de dados e estratégias de otimização do treinamento. A pesquisa possibilitou a análise do desempenho do modelo em um cenário caracterizado por número limitado de amostras e restrições de recursos computacionais e de desenvolvimento.

Os resultados obtidos indicam que as técnicas de data augmentation e as estratégias de balanceamento contribuíram para a melhoria da capacidade de generalização do modelo, especialmente nas classes clinicamente mais relevantes, como HSIL e câncer. Além disso, em aplicações clínicas, a avaliação individual por classe é essencial, uma vez que o desempenho em cada categoria diagnóstica influencia diretamente a definição da conduta terapêutica, justificando, portanto, a preferência por modelos que apresentem maior sensibilidade e melhor desempenho discriminativo nas classes críticas.

Entretanto, considerando a sensibilidade da tarefa e sua relevância direta no contexto clínico e na vida das pacientes, torna-se evidente que a aplicação de modelos de Inteligência Artificial para a classificação de imagens colposcópicas exige elevados níveis de confiabilidade e robustez. Erros de classificação podem impactar diretamente decisões médicas, influenciando desde a realização de exames adicionais até o atraso ou a antecipação de tratamentos. Desse modo, embora o modelo apresente resultados promissores em ambiente experimental, ele não demonstra potencial de aplicabilidade clínica imediata, sendo essencial a utilização de bases de dados mais amplas e diversificadas, bem como a realização de avaliações clínicas rigorosas, a fim de garantir segurança, boa interpretabilidade e contribuição efetiva para o suporte ao diagnóstico médico.

VI. Agradecimentos

Os autores expressam seu profundo agradecimento ao Instituto Nacional de Telecomunicações (Inatel) pelo suporte institucional e pela infraestrutura tecnológica disponibilizada, que foram fundamentais para a realização deste trabalho. Agradecem especialmente ao e-Health Innovation Center, centro de pesquisas em Engenharia Biomédica do Inatel, pelo apoio técnico especializado, pelo ambiente de pesquisa colaborativo e pelo constante incentivo à inovação tecnológica aplicada à saúde. Este trabalho foi realizado com o suporte financeiro da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e da Fundação Inatel (FINATEL), através da concessão de bolsas de Iniciação Científica. O apoio destas instituições foi indispensável para a viabilização da investigação e para o alcance dos resultados científicos apresentados neste artigo.

Referências

[1] World Health Organization. *Cervical cancer – Fact sheet*. https://www.who.int/news-room/fact-sheets/detail/cervical-cancer. Acesso em 15/12/2025.

[2] T. Girum et al. "Gender disparity in epidemiological trend of HIV/AIDS infection and treatment in Ethiopia". Em: *Archives of Public Health* 76 (2018), p. 51. DOI: 10.1186/s13690-018-0299-8. URL: https://pubmed.ncbi.nlm.nih.gov/30237881/.

[3] L. S. Massad et al. "Outcome after negative colposcopy among women with HIV infection". Em: *Obstetrics & Gynecology* 111.2 (2008), pp. 341–348. URL: https://pubmed.ncbi.nlm.nih.gov/18515509/.

[4] J. W. Sellors et al. "Assessment of the cervix after acetic acid wash: interobserver agreement using digital images". Em: *Journal of Lower Genital Tract Disease* 26.1 (2022), pp. 1–9. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC8732320/.

[5] Z. Xue et al. "Artificial intelligence–based cervical cancer screening using colposcopy images". Em: *Frontiers in Medicine* 11 (2024), pp. 1–12. URL: https://pubmed.ncbi.nlm.nih.gov/39280205/.

[6] C. D. Hernández-Silva et al. "HPV and Cervical Cancer: Molecular and Immunological Aspects". Em: *Viruses* 16.3 (2024). URL: https://www.mdpi.com/1999-4915/16/3/327/pdf.

[7] World Health Organization. *Colposcopy and treatment of cervical intraepithelial neoplasia: a beginners' manual*. https://iris.who.int/server/api/core/bitstreams/98beead0-4aca-431a-88a9-c0c30f68ecf3/content. Acesso em 15/12/2025.

[8] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan e P. Ilono. "Deep Convolutional Neural Networks in Medical Image Analysis: A Review". Em: *Information* 16 (2025), p. 195. DOI: 10.3390/info16030195.

[9] L. Hu et al. "An Observational Study of Deep Learning and Automated Evaluation of Cervical Images for Cancer Screening". Em: *Journal of the National Cancer Institute* 111.9 (2019), pp. 923–932. DOI: 10.1093/jnci/djy225.

[10] M. Krichen. "Convolutional Neural Networks: A Survey". Em: *Computers* 12 (2023), p. 151. DOI: 10.3390/computers12080151.

[11] K. He, X. Zhang, S. Ren e J. Sun. "Deep residual learning for image recognition". Em: *arXiv preprint arXiv:1512.03385* (2015). URL: https://arxiv.org/abs/1512.03385.

[12] X. Chen et al. "Battle of Backbones: A Large-Scale Comparison of Pretrained Models Across Computer Vision Tasks". Em: *arXiv preprint arXiv:2206.08016* (2022). URL: https://arxiv.org/abs/2206.08016.

[13] A. T. Lorenzi, A. Longatto-Filho e K. Syrjänen. "Human papillomavirus (HPV) screening and cervical cancer burden". Em: *Virology Journal* (2015). URL: https://www.nature.com/articles/nrdp201686.

[14] *Artigo open access – PubMed Central (PMC6179004)*. https://pmc.ncbi.nlm.nih.gov/articles/PMC6179004/.

Autores



Gislaine Abreu Mendonça É graduanda do 9º período de Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (INATEL). Possui experiência prática no setor de tecnologia, tendo atuado como estagiário na área de telecomunicações da Nokia, empresa parceira da

instituição. Atualmente, concentra seus estudos e pesquisas acadêmicas no desenvolvimento de soluções em Inteligência Artificial, Visão Computacional e Deep Learning aplicadas ao diagnóstico médico e análise de imagens clínicas.



Livia Toledo Gandini é graduanda do 5° período de Engenharia de Computação no Instituto Nacional de Telecomunicações (INATEL). Atua em projetos de iniciação científica voltados à área de Visão Computacional, desenvolvendo e aplicando métodos de Inteligência Artificial e Deep Learning para análise de imagens médicas. Suas atividades acadêmicas concentram-se principalmente em visão computacional e no desenvolvimento de ferramentas computacionais para apoio ao diagnóstico em saúde.



José Andery Carneiro Graduado em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (INATEL, 2019), mestre em Computação Aplicada pela Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (USP - 2023). Professor do INATEL, lecionando nos cursos de graduação das Engenharias, em especial na Engenharia Biomédica. Professor da ETE FMC no curso técnico em Eletrônica com ênfase em Desenvolvimento de Sistemas. Suas principais áreas de pesquisa são Visão Computacional e Inteligência Artificial aplicadas à imagens médicas, e Práticas de Ensino-Aprendizagem.



Paulo Roberto Maia Graduado em Engenharia de Controle de Automação pela Universidade Federal de Itajubá (Unifei) em 2003, especialista em Qualidade e Produtividade pela Unifei em 2011, mestre em Ciências em Engenharia de Produção pela Unifei em 2013, doutor em Ciências em Engenharia de Produção pela Unifei em 2020. Professor do Inatel. Linha de pesquisa: Estatística aplicada as áreas de saúde, ciências exatas e humanas e otimização de processos.



Márcio Erik Franco Graduação em Medicina pela Universidade do Vale do Sapucaí (2007). Residência médica em Ginecologia e Obstetrícia pelo Hospital das Clínicas Samuel Libânio (2013). Especialização em Ginecologia Oncológica pelo Instituto Mario Penna (2015). Mestrado em Ciências Aplicadas à Saúde pela Universidade do Vale do Sapucaí (2018). Especialização em cirurgia oncológica minimamente invasiva no Hospital Sírio-Libanês (2018-2019). Atualmente é membro do Serviço de Ginecologia Oncológica do Hospital das Clínicas Samuel Libânio, médico plantonista em Ginecologia e Obstetrícia do Hospital das Clínicas Samuel Libânio e docente da disciplina de saúde da mulher na Universidade do Vale do Sapucaí (2019).

Técnicas de Esteganálise Estatística para Detecção de Mensagens Ocultas em Imagens Digitais

Douglas Miguel Ferreira, Igor Nogueira Olivio, Ítalo Augusto de Sousa Tacca, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
douglas.ferreira@inatel.br, igor.nogueira@gec.inatel.br, italo.augusto@inatel.br, evandro.cesar@inatel.br

***Abstract*—Statistical steganalysis is crucial for detecting hidden messages in digital images. Classical techniques, such as Chi-Square, RS Analysis, and Sample Pair Analysis, identify changes in pixels and bit planes caused by LSB insertions. However, modern methods in the frequency domain, like DCT and DWT, offer greater resistance to detection due to the subtler distribution of noise. Furthermore, the integration of machine learning, particularly convolutional neural networks (CNNs), has shown great effectiveness in detecting complex and subtle patterns, surpassing traditional methods. However, these methods require large datasets and high computational costs. The combination of statistical and data-driven approaches represents the most promising direction for the evolution of steganalysis.**

***Index Terms*—Steganalysis, LSB, DCT, DWT, Machine Learning.**

***Resumo*—A esteganálise estatística é essencial para detectar mensagens ocultas em imagens digitais. Técnicas clássicas, como Chi-Square, RS Analysis e Sample Pair Analysis, identificam alterações em pixels e planos de bits causadas por inserções LSB. No entanto, métodos modernos no domínio da frequência, como DCT e DWT, oferecem maior resistência à detecção devido à distribuição mais sutil do ruído. Além disso, a combinação de aprendizado de máquina, especialmente com redes neurais convolucionais (CNNs), tem mostrado grande eficácia na detecção de padrões complexos e sutis, superando métodos tradicionais. No entanto, esses métodos demandam grandes bases de dados e alto custo computacional. A integração de abordagens estatísticas e baseadas em dados representa a direção mais promissora para a evolução da esteganálise.**

***Palavras chave*—Esteganálise, LSB, DCT, DWT, Aprendizado de Máquina.**

I. Introdução

A esteganografia digital tem como objetivo ocultar informações em mídias aparentemente inofensivas, como imagens, áudios e vídeos, de forma imperceptível ao usuário. Diferentemente da criptografia, que protege o conteúdo da mensagem, a esteganografia busca ocultar a própria existência da comunicação [1]. Enquanto a criptografia transforma dados legíveis em representações cifradas, preservando a evidência de que há uma mensagem protegida, a esteganografia atua no nível estrutural da mídia, explorando redundâncias estatísticas e perceptuais para inserir informações sem alterar significativamente suas características visuais ou estatísticas [2].

No contexto de imagens digitais, técnicas clássicas como a substituição do bit menos significativo, conhecida como LSB, exploram a baixa sensibilidade do sistema visual humano a pequenas variações nos níveis de intensidade dos pixels. Em imagens RGB de 24 bits, cada pixel é composto por três canais de 8 bits, permitindo que os bits menos significativos sejam modificados com impacto visual mínimo [3]. Entretanto, tais modificações introduzem alterações sutis na distribuição estatística dos valores dos pixels, especialmente nos planos de bits, o que torna possível sua detecção por métodos apropriados de esteganálise [4].

O crescimento do uso dessas técnicas, tanto para fins legítimos quanto maliciosos, impulsionou o desenvolvimento da esteganálise, área dedicada à detecção de mensagens ocultas [5]. A esteganálise não necessariamente busca recuperar o conteúdo escondido, mas sim determinar se uma mídia contém ou não informação embutida. Esse processo é, em essência, um problema de decisão binária, no qual a hipótese nula corresponde à ausência de esteganografia e a hipótese alternativa indica a presença de dados ocultos. A confiabilidade dessa decisão depende da capacidade de modelar o comportamento estatístico natural das imagens e identificar desvios provocados pela inserção da mensagem.

Técnicas estatísticas clássicas, como Chi-Square, RS Analysis e Sample Pair Analysis, foram amplamente estudadas para detectar padrões anômalos decorrentes da modificação sistemática dos bits menos significativos [5]. Essas abordagens exploram propriedades como frequência de ocorrência de pares de valores, correlação entre pixels vizinhos e regularidade nos planos de bits. Quando uma mensagem é inserida de forma sequencial, como ocorre em implementações simples de LSB, a aleatoriedade natural da imagem tende a ser alterada, gerando assinaturas detectáveis [4].

Além dos testes estatísticos específicos, métricas de qualidade amplamente utilizadas em processamento de imagens, como o erro médio quadrático, a relação sinal ruído de pico e o índice de similaridade estrutural, também desempenham papel relevante na análise de imagens suspeitas. O erro médio quadrático quantifica a energia média da diferença pixel a pixel entre a imagem original e a esteganografada. A relação sinal ruído de pico expressa essa diferença em termos logarítmicos, enquanto o índice de similaridade estrutural avalia a preservação de luminância, contraste e estrutura espacial [6]. Embora essas métricas não sejam, isoladamente, detectores de esteganografia, variações sistemáticas em seus valores podem indicar a presença de modificações artificiais.

Em paralelo às técnicas clássicas no domínio espacial, métodos modernos de esteganografia passaram a atuar no domínio da frequência, utilizando transformadas como DCT e DWT para inserir dados em coeficientes de média frequência, especialmente em imagens no formato JPEG [7, 8]. Tais

abordagens tendem a apresentar maior robustez a compressão e filtragem, mas ainda produzem alterações estatísticas mensuráveis nos coeficientes transformados, abrindo espaço para técnicas de esteganálise específicas para esse domínio [6].

Recentemente, abordagens baseadas em aprendizado de máquina e redes neurais convolucionais têm sido empregadas para modelar padrões complexos de alteração estatística introduzidos por diferentes algoritmos de esteganografia. Modelos como EfficientNet aplicados à esteganálise demonstram elevada capacidade de generalização na detecção de imagens esteganografadas, superando, em diversos cenários, métodos puramente analíticos [9]. Essa transição marca uma evolução da esteganálise baseada em regras estatísticas explícitas para modelos orientados por dados, treinados a partir de grandes conjuntos de imagens rotuladas.

Este artigo apresenta uma análise estruturada das principais técnicas de esteganálise aplicadas a imagens digitais, com foco em métodos estatísticos clássicos, métricas de qualidade e avaliação de limiares de detecção. Também são discutidas limitações inerentes ao processo de decisão, como a ocorrência de falsos positivos e falsos negativos, bem como tendências modernas fundamentadas em aprendizado de máquina e análise automatizada de padrões.

II. Fundamentos da Esteganálise

A esteganálise pode ser definida como o processo de identificar indícios de modificação artificial em uma mídia digital, sugerindo a presença de informação oculta [5]. Diferentemente da esteganografia, cujo objetivo é maximizar a imperceptibilidade e a capacidade de inserção, a esteganálise busca modelar o comportamento estatístico natural das mídias digitais e detectar desvios introduzidos por algoritmos de ocultação [6].

Em termos formais, o problema da esteganálise pode ser modelado como um teste de hipóteses. Seja H_0 a hipótese de que a imagem é original e H_1 a hipótese de que a imagem contém mensagem oculta. O objetivo do analisador é decidir entre H_0 e H_1 com base em um conjunto de características extraídas da imagem. A qualidade dessa decisão está diretamente relacionada à sensibilidade do método às alterações introduzidas pelo processo de inserção e à robustez contra variações naturais da imagem.

De forma geral, os ataques podem ser classificados em:

- **Ataques visuais**: baseados na inspeção direta da imagem, buscando distorções, artefatos ou padrões anômalos [5].
- **Ataques estatísticos**: exploram alterações na distribuição dos pixels e nos bits menos significativos [4].

Os ataques visuais constituem a abordagem mais elementar de esteganálise. Eles envolvem a análise direta da imagem, eventualmente com o auxílio de ferramentas que permitem ampliação, realce de contraste e visualização de planos de bits. Em técnicas baseadas em LSB, por exemplo, a inspeção dos planos menos significativos pode revelar padrões artificiais quando a inserção é realizada de forma sequencial ou com alta taxa de ocupação. Regiões homogêneas da imagem são particularmente sensíveis, pois pequenas variações introduzidas nos níveis de intensidade tornam-se mais perceptíveis nessas áreas [2].

Apesar de sua simplicidade, ataques visuais apresentam limitações significativas. Métodos modernos de esteganografia distribuem a mensagem de forma pseudoaleatória ou utilizam estratégias adaptativas, como padrões invertidos de LSB, com o objetivo de reduzir artefatos perceptíveis [10]. Nessas situações, a inspeção visual isolada torna-se insuficiente, exigindo abordagens quantitativas mais sofisticadas.

Os ataques estatísticos, por sua vez, fundamentam-se na premissa de que imagens naturais apresentam regularidades estatísticas específicas, decorrentes de propriedades físicas do processo de aquisição e de características estruturais da cena. A inserção de dados, especialmente quando realizada de maneira sistemática, tende a perturbar tais regularidades. Técnicas clássicas como o teste Qui-Quadrado, RS Analysis e Sample Pair Analysis exploram variações na frequência de pares de valores e na distribuição dos bits menos significativos para identificar tais perturbações [5].

No caso específico da técnica LSB, a substituição direta do bit menos significativo tende a equalizar a distribuição de zeros e uns nesse plano de bits, aproximando-a de uma distribuição uniforme quando a taxa de inserção é elevada. Essa uniformização constitui uma assinatura estatística detectável por testes apropriados [4]. Além disso, a correlação entre pixels vizinhos pode ser reduzida artificialmente, uma vez que a modificação do LSB introduz ruído de alta frequência não correlacionado com a estrutura original da imagem.

Outro aspecto relevante é a relação entre capacidade de ocultação e detectabilidade. Técnicas simples de inserção, como o LSB sequencial, tendem a introduzir alterações estatísticas detectáveis, especialmente quando a capacidade de ocultação é elevada. À medida que a taxa de inserção aumenta, cresce a probabilidade de distorção mensurável nos histogramas, nas distribuições condicionais de pares de pixels e nas métricas de qualidade da imagem [3].

Em contrapartida, métodos que operam no domínio da frequência, como aqueles baseados em DCT ou DWT, buscam inserir dados em coeficientes menos perceptíveis e estatisticamente menos relevantes, reduzindo a magnitude das alterações observáveis no domínio espacial [7, 8]. Ainda assim, tais métodos não são imunes à esteganálise, pois modificações sistemáticas nos coeficientes transformados também podem ser modeladas e detectadas por técnicas estatísticas específicas [6].

Dessa forma, os fundamentos da esteganálise se apoiam na modelagem estatística rigorosa do comportamento natural das imagens digitais e na identificação de padrões de inconsistência introduzidos pelos algoritmos de ocultação. A eficácia de um método de detecção depende da sua capacidade de capturar essas inconsistências sem incorrer em taxas elevadas de erro de classificação.

III. Métodos Estatísticos Clássicos

Diversos algoritmos estatísticos foram propostos para detectar esteganografia no domínio espacial. Essas abordagens baseiam-se na análise da distribuição dos níveis de intensidade dos pixels e na modelagem das propriedades estatísticas naturais de imagens digitais. A hipótese central é que imagens não modificadas apresentam padrões estatísticos específicos, enquanto a inserção de dados, sobretudo por substituição direta de

LSB, introduz perturbações mensuráveis nessas distribuições [5, 6].

A. Chi-Square

O teste Chi-Square é um dos métodos clássicos empregados para detectar esteganografia baseada em LSB. Esse teste avalia a distribuição dos pares de valores de pixels, particularmente pares cuja diferença é igual a um, como $(2k, 2k+1)$. Em imagens naturais, a frequência desses pares tende a seguir uma distribuição não uniforme, influenciada por características da cena e do processo de aquisição.

Quando ocorre inserção por substituição do bit menos significativo, há uma tendência de equalização das frequências dentro desses pares, pois valores originalmente distintos passam a ser ajustados para acomodar bits da mensagem secreta. Essa equalização aproxima as frequências observadas das frequências esperadas sob hipótese de distribuição uniforme, reduzindo a variabilidade natural da imagem [5]. Formalmente, o teste calcula a estatística:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{1}$$

onde O_i representa a frequência observada do par i e E_i a frequência esperada sob hipótese de distribuição natural. Valores elevados da estatística χ^2 indicam desvio significativo entre distribuição observada e esperada, sugerindo possível presença de esteganografia. O método apresenta boa eficácia contra inserções sequenciais e de alta taxa de ocupação, mas sua sensibilidade diminui quando a inserção é realizada de forma pseudoaleatória ou com baixa capacidade [4].

B. RS Analysis

A RS Analysis, conhecida como Regular-Singular Analysis, constitui um método estatístico mais refinado para detecção de LSB. Essa técnica divide a imagem em grupos de pixels e aplica uma função de discriminação que mede a regularidade local. Em seguida, realiza-se uma operação de inversão controlada nos bits menos significativos e avalia-se a variação da função de discriminação [5].

Grupos são classificados como regulares quando a função de discriminação aumenta após a inversão, e como singulares quando a função diminui. Em imagens naturais, existe uma relação característica entre a quantidade de grupos regulares e singulares. A inserção de dados por LSB tende a alterar essa relação, aproximando as proporções de grupos regulares e singulares.

A análise consiste em comparar as proporções antes e depois da inversão. Diferenças estatisticamente significativas entre esses grupos indicam possível inserção de dados [5]. Uma vantagem da RS Analysis em relação ao teste Chi-Square é sua maior robustez frente a inserções pseudoaleatórias, pois considera propriedades estruturais locais da imagem, e não apenas frequências globais.

C. Sample Pair Analysis

O método Sample Pair Analysis explora a correlação estatística entre pares de pixels adjacentes. Em imagens naturais, pixels vizinhos apresentam forte dependência estatística devido à continuidade espacial da cena. A substituição sequencial de LSB introduz ruído independente, reduzindo essa correlação [5].

O algoritmo modela a probabilidade de ocorrência de pares de valores específicos e estima a taxa de inserção com base nas transições observadas entre estados possíveis dos pares. A partir dessa modelagem, é possível inferir não apenas a presença de esteganografia, mas também estimar aproximadamente a proporção de bits inseridos.

Esse método apresenta elevada sensibilidade para detecção de LSB sequencial, sobretudo quando a taxa de inserção é moderada ou alta. Contudo, sua eficácia diminui quando a inserção ocorre em regiões selecionadas adaptativamente ou quando técnicas de embaralhamento são empregadas para reduzir padrões detectáveis [4].

De modo geral, os métodos estatísticos clássicos são particularmente eficazes contra técnicas espaciais simples. Entretanto, quando aplicados a métodos no domínio da frequência, como aqueles baseados em DCT ou DWT, sua eficácia é reduzida, uma vez que as alterações deixam de ocorrer diretamente nos níveis de intensidade dos pixels e passam a afetar coeficientes transformados [7, 8]. Nesses casos, tornam-se necessárias técnicas de esteganálise específicas para o domínio da frequência ou abordagens baseadas em aprendizado de máquina [6].

IV. Histogramas e Análise Estatística

Histogramas representam a distribuição de níveis de intensidade de uma imagem. Em imagens em escala de cinza, o histograma descreve a frequência de ocorrência de cada nível de intensidade no intervalo discreto $[0, 255]$. Em imagens coloridas, histogramas podem ser analisados separadamente para cada canal RGB. Essa representação constitui uma ferramenta fundamental na esteganálise estatística, pois permite observar alterações globais na distribuição dos pixels [6].

Técnicas de esteganografia mal projetadas podem introduzir descontinuidades perceptíveis no histograma [5]. No caso da substituição direta de LSB, a modificação sistemática dos bits menos significativos tende a alterar a frequência relativa de valores consecutivos. Como resultado, pares de intensidades adjacentes podem apresentar distribuição mais uniforme do que a observada em imagens naturais. Essa equalização é particularmente evidente quando a taxa de inserção é elevada, produzindo um efeito de suavização no histograma.

Em imagens naturais, a distribuição dos níveis de intensidade geralmente apresenta comportamento não uniforme, influenciado por características da cena, iluminação e propriedades do sensor. A inserção de dados por LSB reduz parcialmente essa irregularidade, aproximando a distribuição de uma forma estatisticamente mais balanceada. Esse fenômeno pode ser explorado por métodos estatísticos que avaliam a diferença entre a distribuição observada e modelos esperados para imagens não modificadas [4]. Além da análise direta do histograma, métricas quantitativas são amplamente utilizadas para avaliar o impacto da inserção de dados sobre a qualidade da imagem:

- **MSE (Mean Squared Error)**: mede o erro médio quadrático entre a imagem original I e a imagem esteganografada I_s. É definido por:

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(I(i,j) - I_s(i,j) \right)^2, \quad (2)$$

onde M e N correspondem às dimensões da imagem.
- **PSNR (Peak Signal-to-Noise Ratio)**: avalia o nível de ruído introduzido, sendo calculado como:

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad (3)$$

em que MAX representa o valor máximo possível do pixel, tipicamente 255 para imagens de 8 bits.
- **SSIM (Structural Similarity Index)**: mede similaridade estrutural entre duas imagens, considerando luminância, contraste e estrutura espacial [6].

O MSE quantifica a energia média da diferença pixel a pixel, sendo sensível a pequenas alterações introduzidas pela esteganografia. O PSNR, por sua natureza logarítmica, fornece uma medida relativa da intensidade do ruído em relação ao sinal máximo possível. Já o SSIM busca capturar similaridade perceptual, avaliando preservação de padrões estruturais da imagem.

Valores elevados de PSNR e SSIM indicam alta imperceptibilidade visual, mas não garantem indetectabilidade estatística. Uma imagem pode apresentar diferenças numéricas mínimas em relação à original, mantendo PSNR acima de 40 dB e SSIM próximo de 1, e ainda assim exibir alterações detectáveis nos planos de bits ou nas distribuições condicionais de pares de pixels. Em outras palavras, imperceptibilidade visual não implica preservação das propriedades estatísticas naturais da imagem [5, 6].

Portanto, a análise de histogramas e métricas quantitativas deve ser interpretada como parte de um conjunto mais amplo de técnicas de esteganálise. Enquanto métricas de qualidade avaliam distorção global, métodos estatísticos específicos são necessários para detectar modificações sutis e estruturadas introduzidas por algoritmos de ocultação.

V. Avaliação de Limiar de Detecção

Ferramentas automáticas de esteganálise utilizam limiares para classificar imagens como suspeitas ou não. Em geral, um detector produz uma estatística ou escore de decisão $T(I)$ para uma imagem I, e essa medida é comparada a um limiar τ. Se $T(I) > \tau$, a imagem é classificada como esteganografada, caso contrário, é considerada limpa. A escolha adequada de τ é determinante para o desempenho do sistema, especialmente em cenários reais com grande variabilidade de imagens [6]. O ajuste inadequado desses limiares pode resultar em:

- **Falsos positivos**: imagens normais classificadas como esteganografadas.
- **Falsos negativos**: imagens com dados ocultos classificadas como limpas.

Em termos estatísticos, essas situações correspondem, respectivamente, à rejeição incorreta da hipótese nula H_0 e à aceitação incorreta de H_0 quando a hipótese alternativa H_1 é verdadeira. A taxa de falsos positivos está associada ao erro do tipo I, enquanto a taxa de falsos negativos está relacionada ao erro do tipo II. Em sistemas de esteganálise, ambos os erros possuem impacto relevante, dependendo do contexto de aplicação, como investigações forenses digitais ou monitoramento automatizado de tráfego de imagens [4].

Estudos práticos demonstram que limiares muito baixos aumentam a sensibilidade do detector, elevando a taxa de falsos positivos. Nesse cenário, pequenas variações naturais na distribuição estatística de imagens legítimas podem ser interpretadas como indícios de esteganografia. Esse comportamento é particularmente crítico quando se utiliza testes estatísticos globais, como Chi-Square, pois imagens com histogramas atípicos, mas legítimos, podem gerar valores elevados da estatística de teste [5].

Por outro lado, limiares altos reduzem a sensibilidade do sistema, gerando falsos negativos. Imagens com baixa taxa de inserção ou com técnicas adaptativas de ocultação podem apresentar desvios estatísticos sutis, insuficientes para ultrapassar um limiar excessivamente conservador. Esse problema torna-se ainda mais pronunciado quando métodos no domínio da frequência são empregados, pois as alterações tendem a ser distribuídas de forma menos detectável no domínio espacial [7, 8].

Esse comportamento evidencia o *trade-off* entre sensibilidade e especificidade em sistemas de detecção automática. Sensibilidade elevada implica maior capacidade de detectar imagens esteganografadas, ao custo de maior incidência de alarmes falsos. Especificidade elevada reduz a taxa de falsos positivos, porém pode comprometer a capacidade de identificar inserções discretas. Uma abordagem comum para avaliar esse compromisso consiste na utilização de curvas ROC (*Receiver Operating Characteristic*), que relacionam a taxa de verdadeiros positivos à taxa de falsos positivos para diferentes valores de limiar. A área sob a curva ROC fornece uma medida agregada do desempenho do classificador, permitindo comparar diferentes técnicas de esteganálise [6]. Logo, a definição do limiar de detecção não deve ser arbitrária, mas baseada em validação experimental sobre bases de dados representativas. A escolha ideal depende do contexto operacional, do custo associado a cada tipo de erro e do perfil das técnicas de esteganografia que se deseja detectar.

VI. Limitações das Técnicas Clássicas

Embora métodos estatísticos clássicos sejam eficazes contra LSB sequencial e variações diretas no domínio espacial, técnicas modernas baseadas em DCT e DWT apresentam maior resistência à detecção [6]. A principal razão reside no fato de que a inserção de dados deixa de ocorrer diretamente nos níveis de intensidade dos pixels e passa a afetar coeficientes no domínio da frequência, modificando propriedades menos perceptíveis e menos evidentes sob análise estatística global.

Em algoritmos baseados na Transformada Discreta do Cosseno, amplamente utilizados em imagens JPEG, a imagem é dividida em blocos de 8×8 pixels, os quais são convertidos em coeficientes de frequência. A inserção de dados costuma ocorrer em coeficientes de média frequência, evitando tanto as componentes de baixa frequência, que concentram a maior

parte da energia visual, quanto as de alta frequência, mais suscetíveis à compressão [7]. Essa estratégia reduz a magnitude das alterações observáveis no domínio espacial, tornando ineficazes detectores baseados exclusivamente na análise de histogramas de pixels.

De forma análoga, métodos fundamentados na Transformada Discreta Wavelet exploram a decomposição multirresolução da imagem em sub-bandas de aproximação e detalhe. A inserção pode ser realizada em bandas específicas, como LH, HL ou HH, distribuindo o impacto estatístico de maneira mais uniforme e adaptativa [8, 11]. Como consequência, as alterações introduzidas tendem a ser mascaradas pelas próprias características de textura e borda da imagem.

Inserções realizadas no domínio da frequência tendem a distribuir melhor o ruído, dificultando análises baseadas apenas em histogramas ou testes simples como Chi-Square e RS Analysis. Enquanto a substituição direta de LSB frequentemente equaliza distribuições de valores consecutivos, modificações em coeficientes DCT ou DWT produzem alterações indiretas no domínio espacial, menos previsíveis e menos concentradas em planos de bits específicos [6].

Outra limitação relevante das técnicas clássicas reside na suposição de independência ou de modelos estatísticos simplificados para imagens naturais. Imagens reais apresentam ampla variabilidade decorrente de fatores como compressão prévia, ruído do sensor, filtragem e pós-processamento. Essa variabilidade pode gerar padrões semelhantes aos produzidos por esteganografia, aumentando a taxa de falsos positivos quando modelos estatísticos não capturam adequadamente a complexidade da distribuição natural [5].

Além disso, a crescente sofisticação dos algoritmos de ocultação exige abordagens mais robustas de detecção. Técnicas adaptativas, como variantes invertidas de LSB com padrões dinâmicos, foram propostas com o objetivo de preservar propriedades estatísticas locais e reduzir assinaturas detectáveis [10]. Tais métodos procuram minimizar alterações em regiões homogêneas e concentrar a inserção em áreas de maior variabilidade, onde o ruído adicional tende a ser menos distinguível do ruído natural.

Diante desse cenário, torna-se evidente que técnicas puramente analíticas, baseadas em regras fixas ou estatísticas isoladas, apresentam limitações estruturais frente a algoritmos modernos de esteganografia. A evolução da área aponta para a necessidade de métodos capazes de modelar relações estatísticas complexas e dependências de alta ordem, o que tem impulsionado o uso de aprendizado de máquina e arquiteturas profundas na esteganálise contemporânea [9, 6].

VII. Aprendizado de Máquina na Esteganálise

O uso de aprendizado de máquina (*Machine Learning*) e aprendizado profundo (*Deep Learning*) tem se destacado na esteganálise moderna [9, 6]. Diferentemente das abordagens estatísticas clássicas, que dependem de hipóteses explícitas sobre a distribuição dos pixels ou coeficientes transformados, modelos baseados em dados aprendem automaticamente padrões discriminativos a partir de conjuntos de treinamento rotulados.

Na formulação supervisionada do problema, a esteganálise é tratada como uma tarefa de classificação binária. Dado um conjunto de imagens $\{I_i\}$ rotuladas como limpas ou esteganografadas, o modelo aprende uma função de decisão $f(I)$ capaz de estimar a probabilidade de presença de mensagem oculta. Essa abordagem permite capturar dependências estatísticas de alta ordem e relações não lineares difíceis de modelar analiticamente.

Redes neurais convolucionais (*Convolutional Neural Networks*, CNNs) têm demonstrado desempenho superior na detecção de esteganografia, especialmente em cenários onde as alterações introduzidas são sutis e distribuídas [9]. As CNNs operam por meio de filtros convolucionais que exploram padrões locais na imagem, aprendendo automaticamente representações hierárquicas que podem evidenciar ruído artificial, correlações anômalas entre pixels e inconsistências estruturais.

Uma vantagem significativa dessas arquiteturas é a capacidade de extração automática de características. Enquanto métodos clássicos dependem de estatísticas predefinidas, como frequência de pares ou variância local, modelos profundos ajustam seus parâmetros de forma a maximizar a separabilidade entre imagens limpas e esteganografadas. Essa característica tende a aumentar a taxa de detecção e reduzir falsos positivos, particularmente quando o treinamento é realizado com bases de dados representativas e diversificadas [6]. Além do domínio espacial, modelos de aprendizado profundo também podem ser aplicados à análise de coeficientes no domínio da frequência, permitindo a detecção de técnicas baseadas em DCT e DWT. Essa flexibilidade torna o aprendizado de máquina uma ferramenta relevante frente à crescente complexidade dos algoritmos de ocultação [7, 8].

No entanto, modelos baseados em ML exigem grandes bases de dados para treinamento, contendo tanto imagens limpas quanto esteganografadas por diferentes algoritmos e taxas de inserção. A generalização do modelo depende da diversidade e qualidade desses dados. Além disso, o treinamento de redes profundas envolve elevado custo computacional, tanto em termos de processamento quanto de memória, especialmente quando arquiteturas profundas e conjuntos de alta resolução são utilizados [9]. Outro desafio reside na interpretabilidade. Diferentemente de testes estatísticos clássicos, cujo funcionamento pode ser formalmente descrito, modelos profundos operam como sistemas paramétricos complexos, dificultando a análise direta dos critérios utilizados para a decisão. Essa característica pode ser relevante em contextos forenses, nos quais justificativas técnicas claras são necessárias.

Apesar dessas limitações, o aprendizado de máquina representa uma evolução natural da esteganálise, permitindo a modelagem de padrões complexos e adaptando-se à evolução contínua das técnicas de esteganografia. A integração entre métodos estatísticos tradicionais e modelos baseados em dados constitui uma direção promissora para sistemas de detecção mais robustos e generalizáveis.

VIII. Conclusão

A esteganálise estatística desempenha papel central na detecção de mensagens ocultas em imagens digitais, especi-

almente diante da ampla utilização de técnicas espaciais como LSB. Métodos clássicos, como Chi-Square, RS Analysis e Sample Pair Analysis, demonstram eficácia na identificação de alterações na distribuição dos pixels e nos planos de bits. No entanto, sua eficiência é reduzida frente a técnicas modernas baseadas no domínio da frequência, como DCT e DWT. A análise de histogramas e o uso de métricas como MSE, PSNR e SSIM complementam a avaliação, embora alta imperceptibilidade visual não implique necessariamente indetectabilidade estatística. A definição adequada de limiares de decisão é essencial para equilibrar sensibilidade e especificidade, minimizando falsos positivos e falsos negativos. Com a crescente sofisticação dos algoritmos de ocultação, abordagens baseadas em aprendizado de máquina tornam-se cada vez mais relevantes. Modelos como redes neurais convolucionais apresentam maior capacidade de generalização e detecção de padrões complexos. Assim, a integração entre técnicas estatísticas clássicas e métodos baseados em dados representa o caminho mais promissor para sistemas robustos de esteganálise.

IX. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] Neil F. Johnson e Sushil Jajodia. “Exploring steganography: Seeing the unseen”. Em: *Computer* 31.2 (1998), pp. 26–34. DOI: 10.1109/MC.1998.4655281.

[2] Célio Albuquerque, Eduardo Pagani Julio e Wagner Gaspar Brazil. “Esteganografia: a arte das mensagens ocultas”. Em: *Fonte* (2007), pp. 101–109.

[3] Deepika Bansal e Rita Chhikara. “A Study on Steganography Techniques”. Em: *International Journal of Engineering Research & Technology (IJERT)* 3.02 (2014).

[4] João Souza e Maria Silva. “Um Estudo de Técnicas de Esteganálise em Estego-Imagens com Texto Embutido com LSB”. Em: *Brazilian Journal of Digital Forensics and Security* (2023).

[5] Souvik Bhattacharyya, Indradip Banerjee e Gautam Sanyal. “A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier”. Em: *Journal of Global Research in Computer Science* 2.4 (2011), pp. 1–16.

[6] O. Evsutin, A. Melman e R. Meshcheryakov. “Digital Steganography and Watermarking for Digital Images: A Review of Current Research Directions”. Em: *IEEE Access* 8 (2020), pp. 166589–166611. DOI: 10.1109/ACCESS.2020.3022779.

[7] Pratap Chandra Mandal. “A Study of Steganography Techniques Using Discrete Wavelet Transform”. Em: *Journal of Global Research in Computer Science* 5.5 (2014), pp. 7–14.

[8] Moh Moh Zan e Nyein Aye. “A Modified High Capacity Image Steganography using Discrete Wavelet Transform”. Em: *International Journal of Engineering Research & Technology (IJERT)* 2.8 (2013). DOI: 10.17577/IJERTV2IS80678.

[9] Gustavo Possatti, Otávio Andrade e outros. “EfficientNets aplicadas à esteganálise em imagens digitais”. Em: *Revista de Engenharia e Pesquisa Aplicada* 6.2 (2021), pp. 362–372.

[10] Supriadi Rustad et al. “Inverted LSB image steganography using adaptive pattern to improve imperceptibility”. Em: *Journal of King Saud University - Computer and Information Sciences* 34.6, Part B (2022), pp. 3559–3568. DOI: 10.1016/j.jksuci.2020.12.017.

[11] H. S. Majunatha Reddy e K. B. Raja. “High Capacity and Security Steganography Using Discrete Wavelet Transform”. Em: *International Journal of Computer Science and Security (IJCSS)* 3.6 (2010), pp. 462–472.

Análise Comparativa de Técnicas de Esteganografia em Imagens Digitais: Domínio Espacial versus Domínio da Frequência

Douglas Miguel Ferreira, Gustavo Silva Marques Ribeiro, Guilherme Pedro Aquino, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
douglas.ferreira@inatel.br, gustavo.s@ges.inatel.br, guilhermeaquino@inatel.br, evandro.cesar@inatel.br

***Abstract*—Digital steganography aims to conceal information within seemingly harmless media while preserving its imperceptibility. This work presents a comparative analysis between spatial-domain and frequency-domain techniques applied to digital images. It discusses the fundamentals of the steganographic model and its main requirements, including imperceptibility, capacity, robustness, and security, as well as the specific characteristics of the LSB technique, the Discrete Cosine Transform (DCT), and the Discrete Wavelet Transform (DWT). Spatial methods are shown to offer high embedding capacity and low computational complexity, but lower robustness against compression and manipulation. In contrast, frequency-based techniques demonstrate greater resistance and improved structural preservation, although they require higher computational cost. It is concluded that there is no universally superior technique, and the choice depends on the application context and the required level of security.**

***Index Terms*—Digital Steganography, Spatial Domain, Frequency Domain, LSB, DCT/DWT.**

***Resumo*—A esteganografia digital tem como objetivo ocultar informações em mídias aparentemente inofensivas, preservando sua imperceptibilidade. Este trabalho apresenta uma análise comparativa entre técnicas no domínio espacial e no domínio da frequência aplicadas a imagens digitais. Discutem-se os fundamentos do modelo esteganográfico, seus requisitos principais, incluindo imperceptibilidade, capacidade, robustez e segurança, e as características específicas da técnica LSB, da Transformada Discreta do Cosseno (DCT) e da Transformada Discreta Wavelet (DWT). Observa-se que métodos espaciais oferecem alta capacidade e baixa complexidade computacional, porém menor robustez frente a compressão e manipulação. Já técnicas em frequência demonstram maior resistência e melhor preservação estrutural, embora exijam maior custo computacional. Conclui-se que não há uma técnica universalmente superior, sendo a escolha dependente do contexto de aplicação e do nível de segurança requerido.**

***Palavras chave*—Esteganografia Digital, Domínio Espacial, Domínio da Frequência, LSB, DCT/DWT.**

I. Introdução

A ampla disseminação de imagens, vídeos, áudios e documentos eletrônicos em ambientes corporativos, governamentais e pessoais ampliou a superfície de exposição de informações sensíveis, exigindo soluções cada vez mais sofisticadas para proteção de dados. Entre as principais técnicas de segurança da informação destacam-se a criptografia, a marca d'água digital (*watermarking*) e a esteganografia. Enquanto a criptografia visa transformar o conteúdo em uma forma ilegível para terceiros não autorizados, a esteganografia busca ocultar a própria existência da mensagem [1, 2].

A criptografia atua diretamente na confidencialidade da informação, convertendo dados legíveis em representações cifradas por meio de algoritmos e chaves criptográficas. Embora eficaz na proteção do conteúdo, sua utilização evidencia a presença de informação sensível, uma vez que o dado cifrado é claramente identificável como protegido. Por outro lado, a marca d'água digital tem como principal objetivo assegurar autoria, autenticidade e rastreabilidade de conteúdos digitais, inserindo identificadores visíveis ou invisíveis na mídia [2].

Nesse contexto, a esteganografia diferencia-se por seu caráter dissimulatório. A esteganografia digital consiste na inserção de informações secretas em mídias aparentemente inofensivas, como imagens, áudios e vídeos, de forma imperceptível ao sistema visual ou auditivo humano [1, 3]. O princípio fundamental dessa técnica é manter as características estatísticas e perceptuais da mídia original, de modo que a presença da mensagem oculta não desperte suspeitas [1, 2]. Assim, diferentemente da criptografia, que protege o conteúdo, a esteganografia procura mascarar a comunicação, tornando-a invisível ao observador comum [4].

Historicamente, técnicas de ocultação de mensagens remontam à Antiguidade, mas foi no contexto digital que a esteganografia ganhou relevância técnica e científica [4, 1]. Com o avanço do processamento digital de sinais e da compressão de imagens, surgiram abordagens cada vez mais refinadas para inserção de dados em diferentes domínios de representação da mídia [5, 2]. No caso específico de imagens digitais, as técnicas de esteganografia exploram propriedades estruturais da representação binária dos pixels, permitindo a incorporação de informações adicionais com impacto visual mínimo [1].

Diversas técnicas foram propostas ao longo das últimas décadas, sendo tradicionalmente classificadas em dois grandes grupos: métodos no domínio espacial e métodos no domínio da frequência [6, 2]. Técnicas espaciais, como o LSB (*Least Significant Bit*), modificam diretamente os bits menos significativos dos pixels da imagem, explorando a baixa sensibilidade do sistema visual humano a pequenas variações de intensidade [1, 7]. Essa abordagem apresenta simplicidade de implementação e alta capacidade de inserção, porém tende a ser menos robusta frente a operações como compressão com perdas, redimensionamento ou filtragem [8, 9].

Em contrapartida, métodos baseados em transformadas, como a Transformada Discreta do Cosseno (DCT) e a Trans-

formada Discreta Wavelet (DWT), realizam a inserção da informação no domínio da frequência, alterando coeficientes associados a componentes espectrais da imagem [10, 11, 12, 13]. Essas técnicas são amplamente utilizadas em imagens comprimidas e oferecem maior resistência à compressão e a determinados tipos de processamento digital [9, 2]. Entretanto, apresentam maior complexidade computacional e exigem um conhecimento mais aprofundado da estrutura matemática das transformadas envolvidas [11].

Diante desse cenário, torna-se fundamental compreender as diferenças estruturais, funcionais e de desempenho entre os métodos espaciais e os métodos em frequência. Portanto, este artigo apresenta uma análise comparativa estruturada entre técnicas de esteganografia no domínio espacial e no domínio da frequência, discutindo seus princípios teóricos, mecanismos de inserção, vantagens, limitações e aplicações práticas [2, 5]. A comparação é conduzida com base em critérios técnicos consolidados na literatura, como imperceptibilidade, robustez, capacidade de ocultação e complexidade computacional [8, 14, 15].

II. Fundamentos da Esteganografia

O modelo clássico de esteganografia envolve quatro elementos principais: mensagem secreta, algoritmo de inserção, mídia de cobertura (*cover media*) e mídia estego (*stego media*) [1, 2]. A mensagem é incorporada à mídia original por meio de um algoritmo específico, produzindo um arquivo estego que é transmitido por um canal de comunicação. No receptor, o processo inverso é realizado, recuperando-se a informação oculta [5].

De forma mais detalhada, o processo pode ser descrito em etapas [3, 4]. Inicialmente, o emissor seleciona a mídia de cobertura, por exemplo, uma imagem digital. Em seguida, a mensagem secreta é preparada, podendo consistir em texto, arquivo binário ou qualquer outro dado digital. O algoritmo de esteganografia realiza então a inserção dos bits da mensagem em posições estratégicas da mídia, seja no domínio espacial (alterando diretamente pixels) ou no domínio da frequência (modificando coeficientes de transformadas) [6, 11]. O resultado é a mídia estego, visualmente semelhante à original, porém contendo a informação oculta.

Após a geração da mídia estego, o arquivo é transmitido por um canal de comunicação convencional. No lado do receptor, conhecendo o algoritmo utilizado e, quando aplicável, uma chave secreta, executa-se o processo de extração, reconstruindo a mensagem original [1, 2]. Em muitos sistemas, a inserção e extração dependem de parâmetros previamente compartilhados, como posição dos bits, número de planos utilizados ou sequência pseudoaleatória de distribuição dos dados [7].

Em muitos cenários práticos, a esteganografia pode ser combinada com criptografia, aumentando significativamente o nível de segurança [2, 9]. Nesse caso, a mensagem é primeiramente cifrada e, em seguida, inserida na mídia de cobertura. Assim, mesmo que um adversário detecte a presença de dados ocultos por meio de técnicas de esteganálise, o conteúdo permanecerá protegido por mecanismos criptográficos, adicionando uma camada extra de confidencialidade ao sistema.

O projeto de um sistema esteganográfico eficiente envolve o atendimento a um conjunto de requisitos fundamentais que determinam sua eficácia e aplicabilidade [2, 5]. Os principais critérios são:

- **Imperceptibilidade**: a modificação introduzida deve ser invisível ou inaudível ao usuário [1]. No contexto de imagens digitais, isso implica que alterações nos pixels ou nos coeficientes de frequência não devem gerar artefatos visuais perceptíveis, nem distorções estatísticas evidentes [6]. Trabalhos comparativos utilizam métricas como PSNR e análise estatística para avaliar esse critério [8, 9].
- **Capacidade**: refere-se à quantidade de dados que pode ser ocultada na mídia de cobertura [5]. Técnicas espaciais tendem a oferecer maior capacidade bruta, enquanto métodos no domínio da frequência geralmente priorizam robustez em detrimento da taxa de inserção [11, 12].
- **Robustez**: corresponde à resistência da informação oculta a operações comuns de processamento digital, como compressão com perdas ou filtragem [9]. Métodos baseados em transformadas, como DCT e DWT, apresentam maior robustez frente a esses processos quando comparados às técnicas puramente espaciais [13, 11].
- **Segurança**: diz respeito à dificuldade de detecção da presença de dados ocultos por técnicas de esteganálise [14, 15]. Sistemas seguros devem preservar características estatísticas naturais da mídia, minimizando indícios detectáveis por análises estatísticas e técnicas de detecção automatizadas [2].

Na prática, esses requisitos frequentemente entram em conflito [2, 5]. O aumento da capacidade pode comprometer a imperceptibilidade, maior robustez pode implicar maior complexidade computacional e maior segurança pode exigir estratégias de inserção mais sofisticadas e custosas. Assim, o desenvolvimento de técnicas esteganográficas envolve um equilíbrio entre esses parâmetros, dependendo do cenário de aplicação. Esses critérios serão utilizados como base comparativa neste trabalho, servindo como referência para a análise estruturada das técnicas no domínio espacial e no domínio da frequência [8, 2].

III. Técnicas no Domínio Espacial

As técnicas no domínio espacial realizam a inserção da informação diretamente sobre os valores dos pixels da imagem, sem a necessidade de transformações matemáticas para o domínio da frequência [1, 2]. Nessas abordagens, a modificação ocorre na própria representação binária dos componentes de cor, explorando a redundância perceptual do sistema visual humano [1]. De modo geral, são métodos de implementação simples, com baixo custo computacional e alta taxa de inserção de dados [6, 5].

A. Técnica LSB

A técnica LSB (*Least Significant Bit*) é uma das abordagens mais tradicionais e amplamente estudadas em esteganografia digital [1, 6]. Ela consiste na substituição do bit menos significativo de cada canal RGB por bits da mensagem secreta [1]. Como os bits menos significativos possuem menor peso na

representação da cor, sua modificação tende a ser imperceptível ao olho humano [1].

Em uma imagem RGB de 24 bits por pixel, cada pixel é composto por três canais de 8 bits (vermelho, verde e azul), permitindo 256 níveis de intensidade por canal [1]. Ao substituir apenas o bit menos significativo de cada canal, altera-se o valor máximo de intensidade em apenas uma unidade, o que dificilmente provoca variações perceptíveis na imagem resultante [6]. Assim, considerando a utilização de um LSB por canal, é possível modificar até 3 bits por pixel, oferecendo elevada capacidade de inserção [5].

A capacidade total de ocultação está diretamente relacionada à resolução da imagem [5]. Por exemplo, em uma imagem de 800×600 pixels, há 480.000 pixels. Utilizando 3 bits por pixel, obtém-se uma capacidade aproximada de 1.440.000 bits para armazenamento da mensagem secreta. Essa característica torna o LSB particularmente atrativo em aplicações que exigem alta taxa de inserção [6]. Além da substituição simples de 1 bit por canal, existem variantes que utilizam dois ou mais bits menos significativos, aumentando ainda mais a capacidade [6]. Contudo, quanto maior o número de bits alterados, maior a probabilidade de introdução de artefatos visuais ou distorções estatísticas detectáveis [8].

B. Vantagens

As principais vantagens da técnica LSB incluem:

- **Simplicidade de implementação**: o algoritmo é direto, exigindo apenas operações básicas de manipulação de bits [1];
- **Baixa complexidade computacional**: não há necessidade de transformadas ou cálculos matemáticos complexos, sendo adequado para aplicações em tempo real [6];
- **Alta capacidade de ocultação**: a inserção pode explorar múltiplos canais e múltiplos bits por pixel [5];
- **Flexibilidade de aplicação**: pode ser utilizada em diferentes formatos de imagem que preservam diretamente os valores dos pixels [2].

Essas características tornam o LSB uma técnica amplamente empregada em estudos acadêmicos [1, 2], ferramentas didáticas e cenários onde o processamento da imagem após a inserção é mínimo.

C. Limitações

Entretanto, o LSB apresenta vulnerabilidades significativas que comprometem sua segurança e robustez em ambientes reais:

- **Alta suscetibilidade à compressão com perdas (ex.: JPEG)**: como a informação está armazenada diretamente nos pixels, qualquer processo que altere os valores espaciais pode corromper ou eliminar os dados ocultos [9];
- **Vulnerabilidade a ataques estatísticos**, como Chi-Square, RS Analysis e Sample Pair Analysis [14, 15], que exploram alterações na distribuição dos bits menos significativos;
- **Baixa robustez frente a manipulações simples**, como redimensionamento, filtragem, recorte ou adição de ruído [8];
- **Alterações detectáveis em histogramas e padrões estatísticos**, especialmente quando grandes quantidades de dados são inseridas [8].

Do ponto de vista da esteganálise, a substituição sequencial e determinística dos LSBs tende a introduzir padrões estatísticos anômalos, quebrando a aleatoriedade natural da imagem [14, 15]. Essa característica facilita a detecção por ferramentas automatizadas que analisam planos de bits ou distribuições de frequência dos valores de pixel.

Para mitigar essas limitações, foram propostas técnicas aprimoradas, como o LSB invertido com padrões adaptativos e a inserção pseudoaleatória baseada em chave secreta [7]. Nessas abordagens, os bits da mensagem são distribuídos segundo uma sequência aparentemente aleatória, reduzindo padrões detectáveis e aumentando a segurança do sistema [7]. Ainda assim, métodos puramente espaciais permanecem menos robustos quando comparados às técnicas baseadas no domínio da frequência [2, 11].

IV. Técnicas no Domínio da Frequência

As técnicas no domínio da frequência realizam a inserção da informação após a aplicação de transformadas matemáticas que convertem a imagem do domínio espacial para uma representação espectral [2, 5]. Diferentemente das abordagens espaciais, que modificam diretamente os valores dos pixels, os métodos em frequência alteram coeficientes associados a componentes de baixa, média ou alta frequência [11, 13].

De modo geral, essas técnicas apresentam maior robustez frente a operações como compressão com perdas, filtragem e recompressão, pois exploram a própria estrutura interna dos algoritmos de codificação de imagem [9, 2]. Em contrapartida, exigem maior complexidade computacional e conhecimento detalhado da modelagem matemática envolvida [11].

A. Esteganografia Baseada em DCT

A Transformada Discreta do Cosseno (DCT) é amplamente utilizada no padrão JPEG e constitui uma das bases mais importantes para esteganografia em imagens comprimidas [10, 2]. No processo de compressão JPEG, a imagem é inicialmente convertida do espaço de cores RGB para YCbCr, separando luminância e crominância. Em seguida, o algoritmo divide a imagem em blocos de 8×8 pixels, aplicando a DCT a cada bloco para convertê-lo em 64 coeficientes de frequência [10].

Após a transformação, os coeficientes passam por um processo de quantização, etapa responsável pela compressão com perdas [10]. É justamente após essa quantização que ocorre, na maioria das abordagens esteganográficas, a inserção da informação secreta [2]. Geralmente, modificam-se os bits menos significativos de coeficientes de média frequência, evitando tanto os coeficientes de baixa frequência quanto os de alta frequência [2, 9].

Ferramentas clássicas, como o JSteg, utilizam o princípio de substituição sequencial dos LSBs dos coeficientes DCT quantizados não nulos para incorporar a mensagem [1]. Essa abordagem apresenta maior robustez em comparação ao LSB espacial, pois atua diretamente nos coeficientes utilizados pelo próprio algoritmo JPEG [2]. Assim, pequenas recompressões

tendem a preservar parcialmente os dados inseridos, desde que a taxa de compressão não seja excessiva [9].

Entre as principais características da esteganografia baseada em DCT destacam-se:

- Maior resistência à compressão com perdas no formato JPEG [9];
- Melhor preservação das características visuais da imagem [2];
- Maior dificuldade de detecção por inspeção visual simples [1].

Entretanto, a técnica ainda pode ser vulnerável a métodos estatísticos avançados de esteganálise que exploram anomalias na distribuição dos coeficientes DCT, especialmente quando a taxa de inserção é elevada [14, 15].

B. Esteganografia Baseada em DWT

A Transformada Discreta Wavelet (DWT) constitui outra abordagem relevante no domínio da frequência [11, 13]. Diferentemente da DCT, que opera em blocos fixos de 8×8, a DWT realiza uma decomposição multirresolução da imagem, separando-a em sub-bandas de frequência com diferentes níveis de detalhamento [11].

A aplicação bidimensional da DWT resulta em quatro sub-bandas principais:

- **LL** (baixa frequência – aproximação): concentra a maior parte da energia e representa a estrutura global da imagem [11];
- **LH** (detalhe vertical);
- **HL** (detalhe horizontal);
- **HH** (detalhe diagonal).

A banda LL contém as informações mais relevantes da imagem no domínio espacial, sendo menos indicada para modificações significativas [13]. Já as bandas de detalhe (LH, HL e HH) podem ser exploradas para ocultação de dados, pois alterações nessas regiões tendem a causar menor impacto perceptual [12].

Uma vantagem importante da DWT é a possibilidade de inserção adaptativa, distribuindo a informação de maneira mais homogênea entre diferentes escalas e regiões da imagem [12]. Isso contribui para maior imperceptibilidade e resistência a operações como filtragem e compressão [13]. Além disso, por não depender exclusivamente de blocos fixos como na DCT, a DWT reduz a ocorrência de artefatos associados a fronteiras de blocos [11]. Comparativamente, métodos baseados em DWT apresentam:

- Maior resistência a compressão e filtragem [13];
- Melhor preservação estrutural da imagem [11];
- Maior complexidade computacional [11].

Apesar de sua robustez superior, técnicas baseadas em DWT exigem maior poder de processamento e algoritmos mais elaborados para controle da distorção e manutenção das propriedades estatísticas da imagem [12].

V. Análise Comparativa

A Tabela I sintetiza as principais diferenças entre as técnicas analisadas. A comparação evidencia diferenças estruturais entre métodos espaciais e em frequência [2, 5]. A técnica LSB destaca-se pela alta capacidade e baixa complexidade, pois atua diretamente nos bits menos significativos dos pixels [1]. Entretanto, essa simplicidade implica baixa robustez: compressão com perdas, redimensionamento ou filtragem podem comprometer a mensagem oculta [9, 8]. Já as abordagens baseadas em DCT e DWT modificam coeficientes no domínio espectral. Na DCT, a inserção ocorre nos coeficientes quantizados utilizados pelo padrão JPEG, conferindo maior resistência à recompressão [10, 2]. A DWT, por sua decomposição multirresolução em sub-bandas, permite inserção adaptativa em regiões menos perceptíveis, resultando em maior robustez estrutural [11, 13].

TABELA I
Comparação entre Técnicas de Esteganografia

Critério	LSB	DCT	DWT
Domínio	Espacial	Frequência	Frequência
Capacidade	Alta	Média	Média
Robustez	Baixa	Média	Alta
Resistência à Compressão	Baixa	Média	Alta
Complexidade	Baixa	Média	Alta
Vulnerabilidade Estatística	Alta	Média	Baixa

Quanto à capacidade, métodos espaciais tendem a inserir mais bits por pixel, enquanto técnicas em frequência restringem a modificação a coeficientes específicos para preservar imperceptibilidade [12]. Em termos computacionais, LSB exige apenas manipulação de bits, ao passo que DCT e DWT demandam cálculo de transformadas e reconstrução da imagem [11]. Do ponto de vista da qualidade e detecção, o LSB pode apresentar PSNR elevado sob baixa taxa de inserção, mas tende a introduzir padrões estatísticos detectáveis, facilitando ataques de esteganálise [14, 15]. Técnicas em frequência distribuem as alterações de forma mais coerente com a estrutura da imagem, preservando melhor características estatísticas e estruturais [2].

De forma geral, observa-se que técnicas espaciais priorizam simplicidade e capacidade, enqaunto métodos baseados em DCT buscam equilíbrio entre robustez e complexidade e abordagens baseadas em DWT favorecem robustez e preservação estrutural, ao custo de maior processamento [2, 11]. Assim, a escolha da técnica depende do cenário de aplicação e do compromisso desejado entre imperceptibilidade, robustez, capacidade e custo computacional [5].

VI. Conclusão

A análise realizada demonstrou que as técnicas de esteganografia em imagens digitais apresentam características distintas conforme o domínio de atuação. Métodos no domínio espacial, como o LSB, destacam-se pela simplicidade, baixa complexidade computacional e elevada capacidade de inserção, sendo adequados para ambientes controlados. Contudo, apresentam limitações significativas quanto à robustez e resistência a ataques estatísticos e compressão com perdas. Por outro lado, técnicas baseadas em DCT e DWT oferecem maior robustez, melhor preservação estrutural e maior resistência à manipulação da mídia, ao custo de maior complexidade computacional. Assim, a escolha da técnica mais apropriada depende do contexto de aplicação, do nível de segurança desejado e das restrições operacionais do sistema.

VII. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] Neil F. Johnson e Sushil Jajodia. "Exploring steganography: Seeing the unseen". Em: *Computer* 31.2 (1998), pp. 26–34. DOI: 10.1109/MC.1998.4655281.

[2] O. Evsutin, A. Melman e R. Meshcheryakov. "Digital Steganography and Watermarking for Digital Images: A Review of Current Research Directions". Em: *IEEE Access* 8 (2020), pp. 166589–166611. DOI: 10.1109/ACCESS.2020.3022779.

[3] Célio Albuquerque, Eduardo Pagani Julio e Wagner Gaspar Brazil. "Esteganografia: a arte das mensagens ocultas". Em: *Fonte* (jul. de 2007). Acesso em: 12 dez. 2025, pp. 101–109. URL: http://profs.ic.uff.br/~celio/papers/fonte07.pdf (acesso em 12/12/2025).

[4] Eduardo Pagani Julio, Wagner Gaspar Brazil e Célio Vinicius Neves Albuquerque. "Esteganografia e suas Aplicações". Em: *Minicursos do VII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSEG 2007)*. Ed. por Luci Pirmez, Flávia Coimbra Delicato e Luiz Fernando Rust da Costa Carmo. Minicurso. Rio de Janeiro, Brasil: Sociedade Brasileira de Computação (SBC), 2007, pp. 54–102. DOI: 10.5753/sbc.10003.5.2.

[5] Souvik Bhattacharyya, Indradip Banerjee e Gautam Sanyal. "A Survey of Steganography and Steganalysis Technique in Image, Text, Audio and Video as Cover Carrier". Em: *Journal of Global Research in Computer Science* 2.4 (2011), pp. 1–16.

[6] Deepika Bansal e Rita Chhikara. "A Study on Steganography Techniques". Em: *International Journal of Engineering Research & Technology (IJERT)* 3.02 (2014).

[7] Supriadi Rustad et al. "Inverted LSB image steganography using adaptive pattern to improve imperceptibility". Em: *Journal of King Saud University - Computer and Information Sciences* 34.6, Part B (2022), pp. 3559–3568. DOI: 10.1016/j.jksuci.2020.12.017.

[8] Diego Zanchett et al. "Análise Comparativa de Métodos para Esteganografia Digital em Imagens". Em: *Anais do Computer on the Beach* 12 (2021), pp. 240–247.

[9] Osama F. AbdelWahab et al. "Hiding data in images using steganography techniques with compression algorithms". Em: *Telkomnika (Telecommunication Computing Electronics and Control)* 17.3 (2019), pp. 1168–1175. DOI: 10.12928/telkomnika.v17i3.12230.

[10] Luciano Agostini, Sergio Bampi e Grupo de Microeletrônica. "Projeto de uma Arquitetura de DCT 1D para a Compressão de Imagens JPEG". Em: *VII Workshop IBERCHIP*. 2001.

[11] Pratap Chandra Mandal. "A Study of Steganography Techniques Using Discrete Wavelet Transform". Em: *Journal of Global Research in Computer Science* 5.5 (2014), pp. 7–14.

[12] Moh Moh Zan e Nyein Aye. "A Modified High Capacity Image Steganography using Discrete Wavelet Transform". Em: *International Journal of Engineering Research & Technology (IJERT)* 2.8 (ago. de 2013), pp. –. DOI: 10.17577/IJERTV2IS80678.

[13] H. S. Majunatha Reddy e K. B. Raja. "High Capacity and Security Steganography Using Discrete Wavelet Transform". Em: *International Journal of Computer Science and Security (IJCSS)* 3.6 (2010), pp. 462–472.

[14] Matheus Eduardo Polachini. "Detecção de Esteganografia em Imagens Utilizando Aprendizado de Máquina". Em: (2019).

[15] Rafael Albuquerque et al. "EfficientNets aplicadas à esteganálise em imagens digitais". Em: *Revista de Engenharia e Pesquisa Aplicada* 6.2 (2021), pp. 362–372.

ACC – Algoritmo, Código e Conquista: Um Guia de Estudos para a Maratona de Programação SBC

João Pedro Schlindwein, Kleyton Hoffmann
Universidade do Oeste de Santa Catarina – Unoesc
jp.schlindwein@unoesc.edu.br, kleyton.hoffmann@unoesc.edu.br

***Abstract*—This paper presents a structured educational material for beginner competitive programming students, addressing the scarcity of Portuguese-language resources for this field in Brazil. The ACC methodology (Algorithm, Code, Conquest) organizes learning into three progressive stages: theoretical conceptualization, C++ implementation, and practical application in real contest problems. A systematic analysis of 150+ problems from the SBC Programming Marathon Regional Phase (2015–2024) identified the most frequent categories and guided the selection of didactically relevant problems. The material covers data structures, fundamental algorithms, and problem-solving strategies using C++ and the STL, complemented by video lectures available on YouTube. Validation with 11 participants from multiple institutions yielded over 81% of maximum ratings for clarity and methodology. The competitive performance of the Academic League of Competitive Programming (LAPC), achieving 5th place among 20 teams in the 2025 Regional Phase, provides further evidence of the material's effectiveness.**

***Index Terms*—Algorithms, Competitive Programming, C++, Data Structures, ICPC, SBC Programming Marathon.**

***Resumo*—Este artigo apresenta um material didático estruturado para estudantes iniciantes em programação competitiva, respondendo à carência de recursos em língua portuguesa nesse domínio no Brasil. A metodologia ACC (Algoritmo, Código, Conquista) organiza o aprendizado em três etapas progressivas: conceituação teórica, implementação em C++ e aplicação prática em problemas reais de competições. Uma análise sistemática de mais de 150 problemas da Fase Regional da Maratona de Programação da SBC (2015–2024) identificou as categorias mais frequentes e orientou a seleção de problemas com relevância didática. O material abrange estruturas de dados, algoritmos fundamentais e estratégias de resolução de problemas utilizando C++ e a STL, complementado por videoaulas disponibilizadas no YouTube. A validação com 11 participantes de múltiplas instituições obteve mais de 81% de avaliações máximas em clareza e metodologia. O desempenho competitivo da Liga Acadêmica de Programação Competitiva (LAPC), que alcançou o 5º lugar entre 20 equipes na Fase Regional de 2025, corrobora a efetividade do material.**

***Palavras chave*—Algoritmos, C++, Estruturas de Dados, ICPC, Maratona SBC de Programação, Programação Competitiva.**

I. Introdução

A programação competitiva constitui uma área de formação que integra teoria computacional e prática intensiva, exigindo dos participantes o domínio de algoritmos e estruturas de dados para a resolução de problemas sob restrições rigorosas de tempo e memória. O *International Collegiate Programming Contest* (ICPC), com mais de 50 mil participantes em 2022, é o principal evento global nesse cenário [1]. No Brasil, a Maratona de Programação, organizada pela Sociedade Brasileira de Computação (SBC) desde 1996, funciona como etapa qualificatória do ICPC e tem contribuído para a disseminação dessa cultura entre estudantes universitários [2].

Apesar da relevância da programação competitiva para a formação em Computação, a literatura evidencia uma lacuna no cenário educacional brasileiro: a escassez de materiais didáticos em português voltados ao público iniciante. Menezes et al. [3] propuseram o Método TOU (Treinamento Olímpico de Uberlândia), uma trilha adaptativa direcionada à Olimpíada Brasileira de Informática (OBI), com resultados positivos em competições nacionais. Mello [4], por sua vez, construiu um almanaque baseado em análise estatística dos problemas da etapa Nacional da Maratona SBC (2011–2021), voltado a competidores de desempenho médio a alto. Ambos os trabalhos representam avanços significativos; entretanto, nenhum deles aborda a preparação inicial para a Fase Regional da Maratona de Programação, deixando uma lacuna entre o público iniciante e os recursos disponíveis.

A criação da Liga Acadêmica de Programação Competitiva (LAPC) em 2024, na Universidade do Oeste de Santa Catarina (Unoesc), campus Joaçaba, evidenciou de forma concreta essa necessidade. Diante da expansão da equipe e da ausência de um material unificado em português, este trabalho propõe o desenvolvimento de um guia de estudos estruturado, fundamentado na metodologia ACC (Algoritmo, Código, Conquista). O material foi construído a partir de uma análise sistemática de mais de 150 problemas da Fase Regional da Maratona SBC no período de 2015 a 2024, utilizando a linguagem C++ e a STL [5], com base em obras de referência como Cormen et al. [6] e Halim e Halim [7]. O objetivo do artigo é documentar e validar essa progressão pedagógica acessível, que conduz o estudante desde os conceitos fundamentais até a resolução autônoma de problemas de competição.

II. Metodologia

A metodologia deste trabalho está organizada em três etapas: (1) coleta e análise de problemas de competições, (2) desenvolvimento do material didático segundo o *framework* ACC, e (3) validação com grupo focal e questionário estruturado.

A. Coleta e Análise de Problemas

A construção do material partiu de um levantamento sistemático dos problemas da Fase Regional da Maratona de Programação da SBC no período de 2015 a 2024, totalizando 10 edições da competição. Para cada prova, foram coletados os enunciados, as soluções oficiais e os metadados de desempenho das equipes. Os problemas foram classificados em macrocategorias algorítmicas, resultando na distribuição apresentada

na Tabela I. Destaca-se que um mesmo problema pode ser enquadrado em mais de uma categoria, de modo que a tabela reflete a soma de classificações, e não o total de questões.

TABELA I
CONTAGEM DE PROBLEMAS POR CATEGORIA MACRO (2015–2024)

Categoria Macro	Contagem
Ad-Hoc / Simulação / Lógica	28
Programação Dinâmica	21
Grafos / Árvores / Busca	21
Teoria dos Números / Matemática	16
Geometria Computacional	12
Estruturas de Dados	7
Guloso (*Greedy*)	7
Backtracking / Força Bruta	7
Busca Binária	4

A predominância das categorias Ad-Hoc (28), Programação Dinâmica (21) e Grafos (21) orientou a priorização temática do material. A seleção dos problemas que compõem o guia foi guiada por dois critérios objetivos: (a) taxa de acerto superior a 30%, garantindo dificuldade acessível ao público iniciante; e (b) relevância didática, priorizando problemas que exploram um conceito algorítmico central sem exigir múltiplas otimizações simultâneas.

B. Framework ACC: Algoritmo, Código, Conquista

O *framework* ACC foi concebido para estruturar cada unidade de estudo em três etapas progressivas:

- **Algoritmo**: conceituação do tópico, apresentação dos pré-requisitos e construção de uma compreensão intuitiva por meio de diagramas, fluxogramas e análise de complexidade;
- **Código**: tradução da teoria em implementação prática em C++, com codificação passo a passo, explicação das escolhas de estrutura e boas práticas de programação competitiva;
- **Conquista**: aplicação integral do conhecimento na resolução de um problema real da Maratona SBC, demonstrando desde a análise do enunciado até a submissão bem-sucedida em juiz online.

O material resultante abrange três eixos temáticos: (i) técnicas otimizadas de entrada e saída de dados; (ii) onze estruturas de dados da STL, incluindo `vector`, `queue`, `stack`, `set`, `map` e `priority_queue`; e (iii) algoritmos fundamentais como busca binária, *two pointers* e *prefix sum*, com referência a recursos como o CP-Algorithms [8].

Como produto complementar, foi desenvolvida uma série de videoaulas publicamente disponibilizada no YouTube, acessível através da playlist em https://www.youtube.com/playlist?list=PLdsGwHVisfxuz7rg57W2docazXUqqMrhp. Cada videoaula, com duração entre 20 e 30 minutos, segue o formato de *screencast* e aplica integralmente a metodologia ACC: na etapa *Algoritmo*, utiliza-se um quadro branco digital para construir diagramas e fluxogramas que fundamentam a compreensão intuitiva; na etapa *Código*, apresenta-se a implementação passo a passo em C++; e na etapa *Conquista*, resolve-se um problema real da Maratona SBC, percorrendo desde a análise do enunciado até a submissão no juiz online. Cada vídeo constitui uma unidade de estudo autocontida, permitindo ao estudante utilizar o material de forma flexível: seja para superar dificuldades pontuais, aprofundar conhecimentos ou como recurso de apoio em disciplinas de algoritmos e projetos de extensão. A playlist encontra-se em expansão contínua, com novos conteúdos sendo adicionados progressivamente.

C. Validação

A validação foi conduzida em duas frentes complementares. A primeira consistiu no acompanhamento do grupo focal composto pelos membros da LAPC, cujo desempenho em torneios internos e competições oficiais foi monitorado ao longo do período de adoção do material. A segunda frente utilizou um questionário estruturado com escala Likert de 1 a 5, avaliando quatro dimensões: clareza e compreensibilidade do conteúdo, eficácia da metodologia ACC, aplicabilidade dos conceitos aprendidos e capacidade de preparação para competições. O instrumento foi respondido por 11 participantes, dos quais 5 membros da LAPC (Unoesc) e 6 estudantes de outras universidades brasileiras, assegurando diversidade institucional na avaliação.

III. APLICAÇÃO DA METODOLOGIA ACC

Para demonstrar a operacionalização do *framework* ACC, esta seção apresenta a resolução de dois problemas da Maratona SBC pertencentes às categorias mais frequentes identificadas na análise: Programação Dinâmica e Busca Binária na Resposta. Cada resolução segue integralmente o ciclo Algoritmo–Código–Conquista, ilustrando como o material conduz o estudante da conceituação à solução completa.

A. Programação Dinâmica – Ginástica (2017)

Com 21 ocorrências no período analisado, a Programação Dinâmica (DP) constitui uma das categorias mais recorrentes na Fase Regional. O problema "Ginástica", da Primeira Fase de 2017, foi selecionado por apresentar taxa de acerto compatível com o público-alvo e por ilustrar de forma clara os dois princípios fundamentais da técnica: subestrutura ótima e superposição de subproblemas [6].

a.1) Algoritmo

O problema consiste em contar quantas sequências de T minutos são possíveis, no qual cada minuto possui um nível de dificuldade no intervalo $[M, N]$, e o nível deve variar exatamente em ± 1 a cada transição, sem extrapolar os limites do intervalo. O resultado é retornado módulo $10^9 + 7$.

Uma abordagem por força bruta geraria até 2^{T-1} sequências, atingindo $2^{49} \approx 560$ trilhões para $T = 50$, o que é inviável computacionalmente. A solução por DP define o estado $dp[i]$ como a quantidade de sequências válidas que terminam no nível $M + i$ no minuto atual. As transições respeitam as restrições de borda:

- Se $d = M$ (limite inferior): $dp'[0] = dp[1]$, pois o nível só pode ter vindo de $M + 1$;
- Se $d = N$ (limite superior): $dp'[k] = dp[k - 1]$, em que $k = N - M$, pois só pode ter vindo de $N - 1$;
- Caso contrário: $dp'[i] = (dp[i-1] + dp[i+1]) \mod (10^9 + 7)$, pois o nível pode ter vindo de $d - 1$ ou $d + 1$.

No primeiro minuto, cada nível possui exatamente uma sequência válida, de modo que $dp[i] = 1$ para todo i. A resposta final é a soma dos valores do vetor no último minuto.

a.2) Código

A implementação completa é apresentada no Código 1.

Código 1. Solução para o Problema Ginástica

```cpp
#include <bits/stdc++.h>
using namespace std;
const long long MOD = 1000000007;

int main() {
    int T, M, N;
    cin >> T >> M >> N;
    int size = N - M + 1;
    vector<long long> dp(size, 1);

    for (int t = 2; t <= T; ++t) {
        vector<long long> novo(size, 0);
        for (int i = 0; i < size; ++i) {
            int d = M + i;
            if (d == M)
                novo[i] = dp[i + 1];
            else if (d == N)
                novo[i] = dp[i - 1];
            else
                novo[i] = (dp[i-1] + dp[i+1]) % MOD;
        }
        dp = novo;
    }

    long long result = 0;
    for (int i = 0; i < size; ++i)
        result = (result + dp[i]) % MOD;
    cout << result << endl;
    return 0;
}
```

a.3) Conquista

Para a entrada $T = 3$, $M = 2$, $N = 5$ (níveis 2, 3, 4, 5), a Tabela II detalha a evolução do vetor dp a cada minuto.

TABELA II
TRANSIÇÕES DO VETOR DP PARA $T = 3$, $M = 2$, $N = 5$

Minuto	$dp[0]$ (d=2)	$dp[1]$ (d=3)	$dp[2]$ (d=4)	$dp[3]$ (d=5)
1	1	1	1	1
2	1	2	2	1
3	2	3	3	2

No minuto 2, $dp'[0] = dp[1] = 1$ (nível 2 só vem de 3), $dp'[1] = dp[0] + dp[2] = 1 + 1 = 2$ (nível 3 vem de 2 ou 4), e assim sucessivamente. A resposta final é $2 + 3 + 3 + 2 = 10$ sequências válidas. A complexidade temporal $O(T \cdot (N - M + 1))$ resulta em no máximo 5 milhões de operações para os limites do problema ($T \leq 50$, $N - M \leq 10^5$), executando dentro do limite de 1 segundo [9]. A complexidade espacial é $O(N - M + 1)$, pois apenas dois vetores são mantidos simultaneamente.

Este problema exemplifica como a DP transforma um problema de complexidade exponencial em uma solução eficiente por meio da identificação de estados e relações de recorrência, padrão aplicável a diversos outros problemas da Maratona SBC.

B. Busca Binária na Resposta – Comedores de Pipocas (2019)

A busca binária na resposta aplica o paradigma de redução logarítmica do espaço de busca a problemas de otimização com propriedade monotônica. Embora a busca binária apareça explicitamente em apenas 4 problemas do período analisado, essa técnica permeia diversas categorias e constitui ferramenta essencial em problemas de alocação de recursos e determinação de limiares [7]. O problema "Maratona Brasileira de Comedores de Pipocas" (2019) exemplifica essa abordagem.

Em síntese, dados N sacos de pipocas com quantidades variadas, C competidores em paralelo e uma taxa máxima T de consumo por segundo, determinar o tempo mínimo para consumir todas as pipocas. Cada competidor deve consumir uma sequência contígua de sacos, e cada saco deve ser atribuído a um único competidor.

b.1) Algoritmo

A enumeração direta das partições possui complexidade exponencial, inviável para $N \leq 10^5$. A estratégia central consiste em reformular o problema de otimização como problema de decisão: "dado um tempo X, é possível consumir todas as pipocas?" Essa formulação revela monotonicidade (se viável em X segundos, também o é em $X + 1$), permitindo aplicar busca binária.

O espaço de busca é delimitado por:

- Limite inferior: $\max\left(\left\lceil \frac{\max(P_i)}{T} \right\rceil, \left\lceil \frac{\sum P_i}{C \cdot T} \right\rceil\right)$, que corresponde ao maior entre o tempo para o saco mais pesado e o tempo com balanceamento perfeito;
- Limite superior: $\left\lceil \frac{\sum P_i}{T} \right\rceil$, o tempo caso apenas um competidor consumisse tudo.

Para cada candidato X, um algoritmo guloso verifica a viabilidade: calcula-se a capacidade $X \times T$ e percorrem-se os sacos sequencialmente, atribuindo-os ao competidor atual enquanto a soma não excede a capacidade. Quando excede, inicia-se um novo competidor. Se o total de competidores utilizados não ultrapassar C, o tempo X é viável.

b.2) Código

A implementação completa é apresentada no Código 2.

Código 2. Solução para Comedores de Pipocas

```cpp
#include <iostream>
#include <vector>
#include <algorithm>
#include <cmath>
using namespace std;

int N, C, T;
vector<int> sacos;

bool solve(int tempo) {
    long long cap = (long long)tempo * T;
    int comp = 1;
    long long soma = 0;
    for (int i = 0; i < N; i++) {
        if (sacos[i] > cap) return false;
        if (soma + sacos[i] <= cap) {
            soma += sacos[i];
        } else {
            comp++;
            soma = sacos[i];
            if (comp > C) return false;
        }
    }
    return true;
}

int buscaBinaria() {
    int maxS = *max_element(sacos.begin(),
                            sacos.end());
    long long somaT = 0;
```

```
    for (int p : sacos) somaT += p;
    int esq = max((int)ceil((double)maxS / T),
        (int)ceil((double)somaT / ((long long)C * T)));
    int dir = (int)ceil((double)somaT / T);
    int resp = dir;
    while (esq <= dir) {
        int meio = esq + (dir - esq) / 2;
        if (solve(meio)) {
            resp = meio;
            dir = meio - 1;
        } else {
            esq = meio + 1;
        }
    }
    return resp;
}

int main() {
    ios_base::sync_with_stdio(false);
    cin.tie(NULL);
    cin >> N >> C >> T;
    sacos.resize(N);
    for (int i = 0; i < N; i++) cin >> sacos[i];
    cout << buscaBinaria() << "\n";
    return 0;
}
```

b.3) Conquista

Para a entrada $N = 5$, $C = 3$, $T = 4$ e sacos $[5, 8, 3, 10, 7]$, a soma total é 33 pipocas. Os limites do espaço de busca são: inferior $\max(\lceil 10/4 \rceil, \lceil 33/12 \rceil) = \max(3, 3) = 3$ e superior $\lceil 33/4 \rceil = 9$. A busca binária testa $meio = 6$ (viável), $meio = 4$ (viável), $meio = 3$ (inviável), convergindo para a resposta ótima de 4 segundos. Com capacidade individual de $4 \times 4 = 16$ pipocas, a distribuição gulosa resulta em: competidor 1 consome $[5, 8]$ (13 pipocas), competidor 2 consome $[3, 10]$ (13 pipocas) e competidor 3 consome $[7]$ (7 pipocas), todos dentro da capacidade.

A complexidade total é $O(N \log(\text{tempo\_max}))$. Com $N \leq 10^5$ e espaço de busca de até 10^9, são realizadas no máximo $10^5 \times 30 = 3 \times 10^6$ operações, executando confortavelmente dentro dos limites temporais típicos de 1 a 2 segundos.

Os dois exemplos ilustram a versatilidade do *framework* ACC em categorias distintas (contagem combinatória com DP e otimização com busca binária na resposta), demonstrando que o ciclo Algoritmo–Código–Conquista oferece uma estrutura replicável para o ensino de diferentes tópicos algorítmicos. A abordagem é extensível a todo o conteúdo do material, incluindo técnicas de *two pointers*, *prefix sum* e estruturas de dados da STL.

IV. Resultados e Discussão

A validação do material didático foi conduzida em três frentes: acompanhamento do grupo focal da LAPC, monitoramento de desempenho em competições oficiais e avaliação quantitativa por questionário estruturado.

A. Desempenho em Competições

A LAPC, fundada em 2024 na Unoesc (Joaçaba), adotou o material desenvolvido como recurso principal de preparação. Cabe ressaltar que a Maratona de Programação da SBC é uma competição por equipes de três integrantes, com provas de cinco horas contendo entre 12 e 15 problemas, na qual a pontuação é determinada pelo número de problemas resolvidos e, em caso de empate, pelo tempo corrigido. Durante o período de utilização do material, os seguintes resultados foram registrados:

- **Fase Regional da Maratona SBC (2025):** a LAPC alcançou o 5º lugar entre 20 equipes, posicionando-se como a segunda melhor universidade da sede. Trata-se de um resultado expressivo para uma liga com apenas um ano de existência.
- **Maratona Feminina de Programação:** uma integrante classificou-se para a segunda fase nacional, representando a primeira participação da instituição nessa modalidade.

Os torneios internos realizados ao longo do ano revelaram evolução consistente na capacidade dos membros de identificar categorias de problemas, selecionar estruturas de dados adequadas e implementar soluções dentro dos limites de tempo, habilidades diretamente trabalhadas pelo material.

B. Avaliação Quantitativa

O questionário foi aplicado a 11 participantes: 5 membros da LAPC e 6 estudantes de outras universidades brasileiras, garantindo diversidade institucional. As quatro dimensões avaliadas por escala Likert (1–5) e seus resultados são apresentados na Figura 1 e na Tabela III.

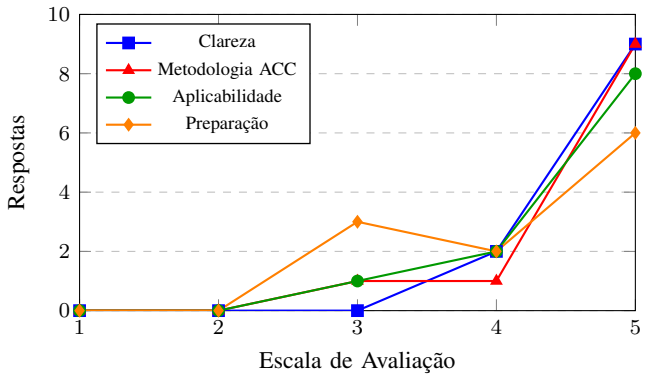


Fig. 1. Distribuição das avaliações do material didático por dimensão.

TABELA III
Resultados da Avaliação por Dimensão (Escala Likert 1–5)

Dimensão	Nota 5	Nota 4	Média
Clareza	81,8%	18,2%	4,82
Metodologia ACC	81,8%	9,1%	4,73
Aplicabilidade	72,7%	18,2%	4,64
Preparação	54,5%	18,2%	4,27

A dimensão de clareza obteve a maior média (4,82), com 100% das respostas concentradas nas notas 4 e 5, indicando que a segmentação proposta pelo *framework* ACC favorece a compreensão do conteúdo. A metodologia ACC recebeu avaliação igualmente elevada (4,73), sugerindo que o ciclo Algoritmo–Código–Conquista é percebido como eficaz pelos estudantes. A aplicabilidade (4,64) confirma que os conceitos ensinados são transferíveis para a resolução de problemas em competições. A dimensão de preparação apresentou a menor média (4,27), com 27,3% de nota 3, resultado coerente com o escopo deliberadamente introdutório do material, que não abrange tópicos avançados como grafos complexos ou otimizações de programação dinâmica.

C. Discussão

A convergência entre os resultados do questionário e o desempenho competitivo sustenta a validade do material proposto. O 5º lugar da LAPC na Fase Regional, alcançado em apenas sua segunda participação, sugere que a progressão pedagógica do *framework* ACC contribui para a formação efetiva de competidores iniciantes. É relevante observar que as categorias mais abordadas no material (Ad-Hoc, Programação Dinâmica e Busca Binária) correspondem às classes de problemas tipicamente resolvidos por equipes de nível intermediário na Fase Regional, o que sustenta a pertinência da seleção temática realizada.

Do ponto de vista qualitativo, os membros da LAPC reportaram que a estrutura do *framework* ACC facilitou a transição entre o estudo teórico e a aplicação em problemas reais, reduzindo a barreira inicial enfrentada por estudantes que não possuíam experiência prévia em competições. A organização em unidades autocontidas permitiu que cada integrante avançasse no próprio ritmo, utilizando as videoaulas como recurso complementar para revisão e aprofundamento.

Contudo, reconhece-se que fatores complementares, como o engajamento na liga, a prática contínua em plataformas online como Codeforces e Beecrowd e o uso de recursos adicionais, podem ter influenciado os resultados observados. A disponibilização das videoaulas no YouTube amplia o potencial de impacto do material para além da LAPC, possibilitando validações futuras com públicos mais amplos e diversificados.

V. Conclusão

Este trabalho apresentou o desenvolvimento e a validação de um guia de estudos para programação competitiva em língua portuguesa, fundamentado na metodologia ACC (Algoritmo, Código, Conquista). A partir da análise sistemática de mais de 150 problemas da Fase Regional da Maratona SBC ao longo de dez edições (2015–2024), foram identificadas as categorias algorítmicas mais frequentes, com destaque para Ad-Hoc/Simulação (28 ocorrências), Programação Dinâmica (21) e Grafos (21), o que orientou a construção de um material progressivo. O guia abrange desde técnicas otimizadas de entrada e saída de dados, passando por onze estruturas de dados da STL, até algoritmos fundamentais como busca binária, *two pointers* e *prefix sum*, complementado por uma série de videoaulas disponibilizadas publicamente no YouTube.

A metodologia ACC, ao organizar o aprendizado no ciclo Algoritmo–Código–Conquista, demonstrou ser eficaz na construção progressiva do conhecimento. A validação quantitativa com 11 participantes de múltiplas instituições revelou médias superiores a 4,27 em todas as dimensões avaliadas, com destaque para clareza (4,82) e eficácia da metodologia (4,73). O desempenho da LAPC na Maratona SBC 2025, que alcançou o 5º lugar entre 20 equipes em sua segunda participação, e a classificação de uma integrante para a fase nacional da Maratona Feminina constituem evidências complementares da efetividade da abordagem na formação de competidores iniciantes.

Como limitações, reconhece-se que a amostra de 11 respondentes, embora representativa de diferentes contextos institucionais, restringe a generalização estatística dos resultados. A ausência de grupo de controle impede a atribuição causal exclusiva dos resultados ao material desenvolvido.

Como trabalhos futuros, propõe-se a expansão do material em três direções: (1) estruturas de dados avançadas, como *Segment Tree*, *Union-Find* e Árvore de Fenwick, que são recorrentes em problemas da Final Brasileira; (2) algoritmos de grafos complexos, incluindo fluxo máximo, componentes fortemente conexas e caminhos mínimos com variações; e (3) técnicas de otimização avançada, como programação dinâmica com *Convex Hull Trick* e aritmética modular. Adicionalmente, planeja-se a adaptação do conteúdo para a linguagem Python, ampliando a acessibilidade para estudantes com diferentes perfis de formação, e a condução de validações com amostras maiores e comparações controladas com outros materiais didáticos disponíveis.

Ao oferecer um recurso didático acessível e validado em português, este trabalho contribui para a democratização da programação competitiva no Brasil, reduzindo barreiras de entrada e fortalecendo a preparação de estudantes para competições nacionais e internacionais como o ICPC.

VI. Agradecimentos

Os autores agradecem aos membros da Liga Acadêmica de Programação Competitiva (LAPC) da Unoesc, campus Joaçaba, pela participação ativa na validação do material e pelo engajamento contínuo nos treinamentos. Agradecem também aos estudantes de outras universidades que contribuíram com a avaliação do questionário, e à Universidade do Oeste de Santa Catarina pelo apoio institucional ao desenvolvimento deste trabalho.

Referências

[1] ICPC. *The International Collegiate Programming Contest*. [Online]. Disponível em: https://icpc.global. Acesso em: 10 mar. 2025. 2025.

[2] Sociedade Brasileira de Computação. *Maratona de Programação*. [Online]. Disponível em: https://maratona.sbc.org.br. Acesso em: 11 mar. 2025. 2023.

[3] C. Menezes, G. Menezes, S. Santos e J. Pereira. "Uberlândia nas Competições de Programação de Computadores após a Aplicação do Método TOU". Em: *Anais do XXXV Simpósio Brasileiro de Informática na Educação (SBIE)*. Rio de Janeiro, RJ, Brasil, 2024, pp. 2627–2636. DOI: 10.5753/sbie.2024.245017.

[4] R. G. de Mello. *Uma Abordagem de Ciência de Dados na Construção de um Almanaque para Programação Competitiva*. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação), Universidade do Estado de Santa Catarina. Joinville, SC, Brasil, 2022.

[5] B. Stroustrup. *The C++ Programming Language*. Fourth. Upper Saddle River, NJ: Addison-Wesley, 2013.

[6] T. H. Cormen, C. E. Leiserson, R. L. Rivest e C. Stein. *Introduction to Algorithms*. Fourth. Cambridge, MA: MIT Press, 2022.

[7] S. Halim e F. Halim. *Competitive Programming 4: The Lower Bound of Programming Contests in the 2020s*. Singapore: Lulu Press, 2020.

[8] CP-Algorithms. *CP-Algorithms*. [Online]. Disponível em: https://cp-algorithms.com. Acesso em: 13 mar. 2025. 2025.

[9] Sociedade Brasileira de Computação. *Informações sobre a Prova: XXII Maratona de Programação, Primeira Fase 2017*. [Online]. Disponível em: https://maratona.sbc.org.br/hist/2017/prim-fase17/info_contest.pdf. Acesso em: 8 set. 2025. 2017.

AUTORES



João Pedro Schlindwein é graduado em Engenharia de Computação pela Universidade do Oeste de Santa Catarina (Unoesc), campus Joaçaba. Um dos fundadores da Liga Acadêmica de Programação Competitiva (LAPC), da qual foi presidente durante a graduação, possui interesse nas áreas de algoritmos, estruturas de dados e programação competitiva. Atua no desenvolvimento de materiais didáticos sobre competições de programação e trabalha como desenvolvedor de software.



Kleyton Hoffmann possui graduação em Engenharia Elétrica pela Universidade do Oeste de Santa Catarina (2011), mestrado e doutorado em Engenharia Elétrica pela Universidade Federal de Santa Catarina (UFSC), no Grupo de Concepção e Análise de Dispositivos Eletromagnéticos (GRUCAD). É professor em tempo integral da Unoesc, atuando nos cursos de Engenharia Elétrica e Engenharia de Computação. Pesquisa temas relacionados a sistemas embarcados, processamento digital de sinais, inteligência artificial e robótica, sendo coordenador da equipe de robótica da instituição. É vice-líder do grupo de pesquisa em Aplicações Tecnológicas Industriais (GATI) da Unoesc Joaçaba e membro dos Núcleos Docentes Estruturantes dos cursos de Engenharia Elétrica e Engenharia de Computação.

Sistema em Tempo Real para Detecção de Gestos Manuais da Língua Brasileira de Sinais com Aprendizado de Máquina

Gabriel Rech Garcia de Moura, Kleyton Hoffmann
Universidade do Oeste de Santa Catarina - Unoesc
gabrielrechmoura@gmail.com.br, kleyton.hoffmann@unoesc.edu.br

***Abstract*—This work presents a computational system for real-time recognition of the Brazilian Sign Language (Libras). The architecture uses the MediaPipe library for spatial extraction of 21 hand landmarks and a hybrid neural network (1D CNN and Bi-LSTM) for spatio-temporal classification. A custom dataset was built for training, comprising 6,000 videos of 30 gestures recorded by 10 volunteers. The model achieved a 90% accuracy on the test set. The average inference latency was only 40.89 ms operating exclusively on a CPU, proving the system's feasibility on conventional hardware. The main limitation observed is the 2D camera occlusion of visually similar gestures, such as the letters 'T' and 'F'. Integrated into a graphical user interface, the solution validates the effectiveness of this approach as an assistive technology for communicational accessibility.**

***Index Terms*—computer vision, Libras, MediaPipe, neural networks.**

***Resumo*—Este trabalho apresenta um sistema computacional para reconhecimento da Língua Brasileira de Sinais (Libras) em tempo real. A arquitetura utiliza a biblioteca MediaPipe para extração espacial de 21 *landmarks* manuais e uma rede neural híbrida (CNN 1D e Bi-LSTM) para a classificação espaço-temporal. Para o treinamento, construiu-se uma base de dados própria com 6.000 vídeos de 30 gestos, gravados por 10 voluntários. O modelo alcançou 90% de acurácia no conjunto de teste. A latência média de inferência foi de apenas 40,89 ms operando exclusivamente em CPU, atestando a viabilidade do sistema em *hardwares* convencionais. A principal limitação observada reside na oclusão em câmeras 2D para gestos visualmente similares, como as letras 'T' e 'F'. Integrada a uma interface gráfica, a solução valida a eficácia dessa abordagem como tecnologia assistiva para acessibilidade comunicacional.**

***Palavras chave*—Libras, MediaPipe, redes neurais, visão computacional.**

I. Introdução

A Língua Brasileira de Sinais (Libras) é o principal meio de comunicação utilizado por grande parte da comunidade surda no Brasil, sendo reconhecida oficialmente como meio legal de comunicação e expressão, com estrutura linguística própria [1, 2]. Com mais de 10 milhões de brasileiros apresentando algum grau de surdez [3, 4], evidencia-se a importância de desenvolver tecnologias assistivas que promovam a acessibilidade comunicacional e ampliem a inclusão social.

No campo da tecnologia, os avanços em visão computacional e aprendizado de máquina têm possibilitado interpretar sinais visuais complexos. Diversos estudos vêm explorando o uso da biblioteca MediaPipe em conjunto com redes neurais (como CNN e LSTM) para o reconhecimento automático de gestos [5, 6, 7, 8, 9]. No entanto, muitas dessas abordagens apresentam limitações: ou dependem de sensores corporais, ou focam apenas no reconhecimento de sinais estáticos [10, 11], o que restringe sua aplicabilidade em contextos de comunicação dinâmica.

Diante disso, este trabalho apresenta o desenvolvimento de um software de reconhecimento de sinais manuais da Libras para funcionamento em tempo real. A proposta baseia-se na utilização da biblioteca MediaPipe para a captura dos pontos-chave das mãos via webcam, combinada com uma rede híbrida de CNN 1D e `Bidirectional(LSTM)` para a classificação temporal dos gestos. Para treinar o modelo, foi construída uma base de dados própria contendo 30 gestos (o alfabeto completo e quatro palavras), coletados a partir de 10 voluntários distintos.

O objetivo principal deste estudo é avaliar o desempenho e a viabilidade do sistema operando em ambiente desktop, analisando métricas rigorosas de acurácia e latência para validar a eficácia da abordagem em tempo real. O artigo está estruturado da seguinte forma: a Seção II apresenta a fundamentação teórica; a Seção III descreve a metodologia e a implementação; a Seção IV detalha os resultados obtidos, incluindo o desempenho de inferência; e a Seção V traz as considerações finais.

II. Fundamentação Teórica

Esta seção apresenta os trabalhos correlatos e as tecnologias baseadas em visão computacional e aprendizado de máquina que fundamentam o desenvolvimento do sistema proposto.

A. Trabalhos Correlatos

Diversos estudos exploram o uso de técnicas de visão computacional para o reconhecimento automático da Libras. Trabalhos como o de Santos et al. [10] apresentaram bons resultados na identificação de sinais, porém com foco em imagens estáticas. Como a Libras depende fortemente da dinâmica gestual, abordagens temporais tornam-se essenciais.

Nesse contexto, Agrawal et al. [5] e Cherrate et al. [6] propuseram sistemas utilizando a biblioteca MediaPipe aliada a redes recorrentes LSTM para o reconhecimento sequencial. Avançando na captura de características espaço-temporais, Stefano et al. [8] e Guerra et al. [12] demonstraram a eficácia do uso de Redes Neurais Convolucionais (CNN). O presente trabalho alinha-se a esses esforços de processamento dinâmico, adotando uma abordagem que une a leveza do rastreamento por pontos-chave com o poder de classificação de redes neurais híbridas.

B. Visão Computacional e Redes Neurais

No contexto deste trabalho, a visão computacional permite identificar e analisar os gestos manuais de forma não invasiva. Para o processamento eficiente desses dados, destacam-se as Redes Neurais Convolucionais (CNN), que são eficazes na extração de padrões espaciais a cada instante de tempo. Por sua vez, as redes recorrentes do tipo LSTM (*Long Short-Term Memory*) são projetadas para capturar relações temporais entre os quadros, permitindo interpretar a fluidez do movimento.

A combinação dessas arquiteturas (CNN e LSTM) tem se mostrado altamente eficaz em sistemas de reconhecimento de sinais em tempo real, como demonstrado por Sravani et al. [7]. Tal abordagem híbrida representa uma alternativa robusta e eficiente, sendo perfeitamente compatível com o objetivo de desenvolver uma solução funcional para ambientes desktop sem a necessidade de hardware de alto custo.

C. MediaPipe Hands

O rastreamento espacial é realizado via MediaPipe, uma biblioteca do Google que oferece soluções de visão computacional otimizadas para tempo real. Utiliza-se especificamente o módulo *MediaPipe Hands* [13], que detecta e rastreia 21 pontos-chave (*landmarks*) tridimensionais em cada mão. A Figura 1 ilustra os pontos extraídos por esta ferramenta.

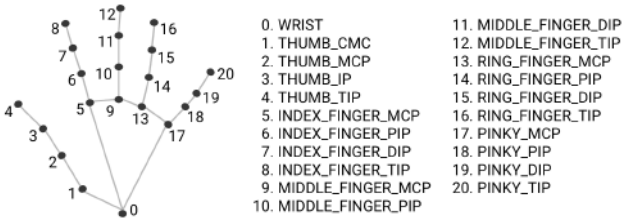


Fig. 1. Landmarks detectados pelo MediaPipe Hands.

Essa detecção opera com alta precisão sem exigir sensores externos ou marcações físicas. Os *landmarks* representam as posições relativas das articulações e atuam como os dados numéricos de entrada para a rede neural, substituindo o processamento pesado de matrizes de pixels e garantindo a baixa latência do sistema.

III. Materiais e Métodos

Esta seção detalha o *pipeline* de desenvolvimento do sistema de reconhecimento em tempo real, abrangendo desde a concepção da base de dados até o treinamento da arquitetura híbrida e a medição de desempenho da inferência. O sistema foi integralmente desenvolvido na linguagem Python, utilizando OpenCV para manipulação de vídeo, MediaPipe para visão computacional tridimensional e a API Keras/TensorFlow para a modelagem da rede neural.

A. Coleta de Dados e Estruturação

Para garantir a generalização do modelo perante diferentes características biométricas e condições de luminosidade, construiu-se uma base de dados própria contendo 30 gestos da Libras: as 26 letras do alfabeto manual e quatro palavras de interação (`OI`, `SIM`, `NÃO` e `TCHAU`).

A coleta foi realizada via *webcam* padrão (1280×720 pixels) com a participação de 10 voluntários distintos. Cada voluntário gravou 20 amostras para cada um dos 30 gestos selecionados, resultando em um *dataset* robusto de 6.000 vídeos originais (10 voluntários $\times$ 20 amostras $\times$ 30 gestos). Todas as capturas foram realizadas a uma taxa constante de 30 quadros por segundo (fps). Para garantir dados suficientes de cada classe, os gestos estáticos foram gravados em vídeos de 2 segundos, enquanto os gestos dinâmicos, que exigem maior tempo de execução, abrangeram 4 segundos de gravação. Destaca-se que o autor da pesquisa não compôs a base de treinamento, visando atuar como um usuário inédito durante a etapa de validação em tempo real.

A Figura 2 apresenta exemplos de quadros extraídos durante o processo de gravação, ilustrando a variabilidade de cenários e sujeitos presentes na base de dados.



Fig. 2. Exemplos de quadros extraídos dos vídeos coletados, evidenciando a diversidade de voluntários e ambientes.

B. Pré-processamento e Extração Espacial

A conversão dos vídeos brutos em dados processáveis ocorreu por meio do módulo estático do *MediaPipe Hands*. Para cada quadro (*frame*), o algoritmo detectou a presença da mão e extraiu as coordenadas tridimensionais normalizadas (x, y, z) de 21 *landmarks* articulares. Esses dados foram concatenados em um vetor unidimensional de 63 características. Nos quadros em que a mão sofreu oclusão ou falha de detecção, o vetor foi preenchido com zeros, garantindo a integridade da série temporal.

Sequencialmente, aplicou-se um janelamento temporal. Os dados extraídos foram segmentados em blocos contínuos de 30 *frames*, sem sobreposição. Dessa forma, cada amostra de entrada para a rede neural assumiu o formato matricial $(30, 63)$. Os rótulos nominais das classes foram convertidos para valores numéricos (*LabelEncoder*) e transformados para o formato *one-hot encoding*, adequando-os à função de custo da rede.

C. Arquitetura CNN 1D e Bi-LSTM

Para a classificação dos padrões espaço-temporais, implementou-se uma rede neural híbrida, cuja topologia é

ilustrada na Figura 3.

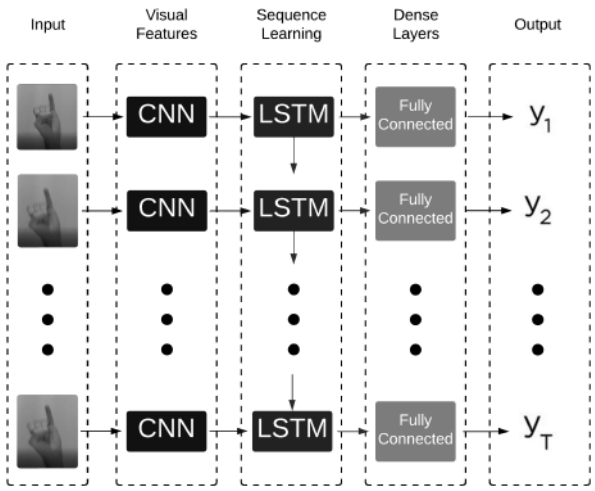


Fig. 3. Arquitetura do modelo híbrido CNN 1D e Bi-LSTM implementado.

O fluxo de processamento inicia-se com uma camada `Conv1D`, que opera sobre a janela temporal extraindo as correlações espaciais instantâneas (ex: a curvatura dos dedos e a distância entre falanges). O mapa de características resultante é estabilizado por `BatchNormalization` e tem sua dimensionalidade reduzida via `MaxPooling1D`.

A saída convolucional é então submetida a uma camada `Bidirectional(LSTM)`. Esta estrutura processa a evolução temporal do movimento nos sentidos direto e reverso, capacidade crucial para contornar a alta similaridade interclasse (ex: a sutil diferença de movimento entre os sinais 'A' e 'SIM'). Camadas de *Dropout* (30%) foram inseridas após as etapas extratoras para mitigar o *overfitting*. A classificação culmina em camadas densas, finalizando com a função de ativação *softmax*.

D. Treinamento e Inferência

A base de dados segmentada foi dividida de forma estratificada: 20% destinados ao conjunto de teste isolado e os 80% restantes particionados para treino e validação. O modelo foi compilado com o otimizador *Adam*, adotando uma taxa de aprendizado conservadora de 0.0001 para evitar oscilações na função de perda (*categorical crossentropy*).

Para mitigar o desbalanceamento natural ocorrido após a exclusão de janelas temporais inválidas, aplicou-se a técnica de balanceamento de pesos (`class_weight='balanced'`). O treinamento ocorreu por até 50 épocas em lotes (*batches*) de 32 amostras, auxiliado por uma rotina de *early stopping* (paciência de 10 épocas) para restaurar os pesos de menor erro de validação.

Finalmente, a inferência em tempo real foi estruturada em um *loop* de captura contínua via OpenCV e interface gráfica via Tkinter. O sistema acumula as coordenadas de rastreamento em um *buffer* dinâmico; ao atingir o limite de 30 quadros, a matriz é submetida ao modelo salvo (`.h5`), retornando a classe predita e o índice de confiança, com a latência de execução rigorosamente monitorada para validar a viabilidade operacional em hardware de uso geral (CPU).

IV. Resultados e Discussões

Após o treinamento da arquitetura híbrida, o desempenho do modelo foi avaliado utilizando os 20% de dados separados exclusivamente para teste. A Figura 4 apresenta o comportamento da acurácia e da função de perda (*loss*) ao longo das épocas de treinamento.

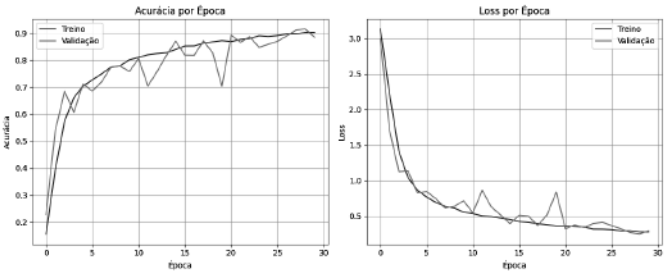


Fig. 4. Acurácia (esquerda) e perda (direita) para os conjuntos de treino e validação.

As curvas demonstram uma convergência rápida e estável. O paralelismo entre as métricas de treino e validação, aliado ao acionamento do *early stopping* próximo à época 30, atesta que o modelo foi capaz de generalizar o aprendizado sem apresentar sobreajuste (*overfitting*).

No conjunto de testes inéditos, o classificador alcançou uma acurácia global e um *F1-score* médio (*macro avg*) de 90%. Para compreender o comportamento da rede perante as 30 classes gestuais, gerou-se a matriz de confusão apresentada na Figura 5.

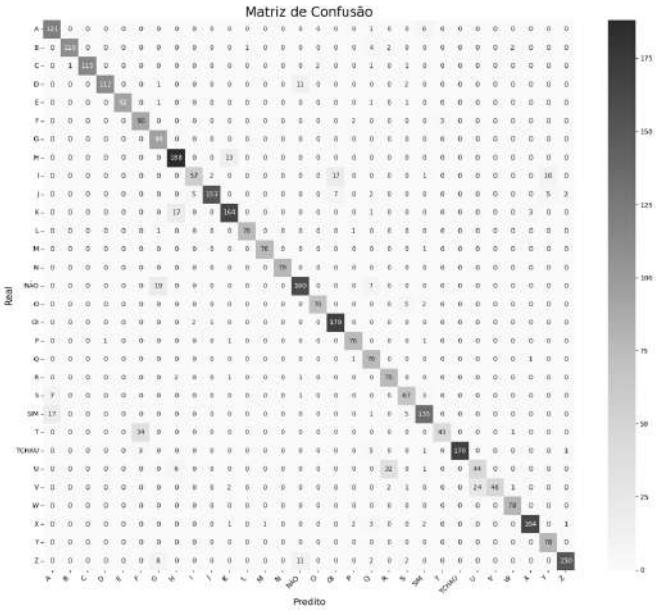


Fig. 5. Matriz de confusão obtida no conjunto de teste para as 30 classes da Libras.

A análise da matriz revela que a esmagadora maioria das predições concentra-se na diagonal principal. Os erros de classificação detectados não são aleatórios, mas de natureza morfológica, ocorrendo sistematicamente entre sinais que compartilham configurações espaciais quase idênticas.

As confusões mais expressivas ocorreram entre as letras 'T' e 'F' (34 ocorrências) e na cadeia de similaridade entre 'V', 'U'

e ’R’ (’U’ classificado como ’R’ 32 vezes, e ’V’ como ’U’ 24 vezes). Sinais com dinâmicas bidirecionais similares também apresentaram trocas naturais, como a sobreposição entre ’I’ e ’OI’ (17) e o par recíproco ’H’ e ’K’ (completando 30 trocas mútuas).

Esses erros mais frequentes entre alguns gestos não significam que o modelo falhou em aprender, mas sim que existe uma limitação na própria captura de vídeo por uma câmera comum (2D). Na Libras, a única diferença visual entre as letras ’T’ e ’F’, por exemplo, é a posição do dedo polegar (que fica por dentro ou por fora do dedo indicador). Como os vídeos foram gravados de frente, a falta de profundidade real faz com que um dedo esconda o outro (oclusão). Embora o MediaPipe tente calcular a profundidade (coordenada z), essa estimativa perde precisão quando a mão está muito fechada. Com isso, os dados numéricos gerados para a rede neural ficam praticamente idênticos para os dois gestos, causando a confusão.

Apesar dessas sobreposições isoladas resultarem em *F1-scores* locais mais baixos para classes como ’U’ (0.58) e ’T’ (0.69), a vasta maioria do vocabulário obteve métricas de precisão e revocação superiores a 90%. Gestos como ’C’, ’E’, ’M’, ’N’ e ’TCHAU’ atingiram índices quase perfeitos, demonstrando a robustez da arquitetura híbrida na classificação correta dos sinais.

Por fim, atestou-se a viabilidade do sistema para operação contínua através da medição da latência de inferência em tempo real. Os testes de desempenho foram conduzidos em um computador portátil (*notebook*). Devido à ausência de suporte nativo do *framework* TensorFlow para a aceleração na GPU integrada deste equipamento, a execução do modelo ocorreu estritamente sobre a CPU. Mesmo operando diante dessa limitação de *hardware*, o sistema registrou um tempo médio de resposta de 40,89 ms (com desvio padrão de 4,95 ms) por predição, oscilando entre 33,52 ms e 73,44 ms — desconsiderando o custo de alocação inicial (*warm-up*).

Essa baixa latência demonstra que, em regime de execução contínua, o desempenho do modelo é não apenas rápido, mas altamente estável. Tais resultados garantem uma experiência fluida e validam a viabilidade do sistema para operação em tempo real, inclusive em ambientes com recursos computacionais modestos.

A integração bem-sucedida de todas essas métricas — captura, abstração espacial, janelamento e classificação ultrarrápida — culminou em uma aplicação de desktop plenamente funcional. A Figura 6 ilustra a interface gráfica (GUI) desenvolvida, operando em tempo real.

O sistema processa o fluxo visual isolando a predição textual em um painel lateral, garantindo uma experiência de usuário clara e livre de poluição visual sobre a execução do sinal, validando o objetivo principal da pesquisa.

V. Conclusão

Este trabalho apresentou o desenvolvimento de um sistema computacional para o reconhecimento em tempo real de 30 gestos da Libras. A combinação da extração de pontos-chave (MediaPipe) com uma rede neural híbrida (CNN 1D e Bi-LSTM) mostrou-se uma abordagem altamente eficaz. Treinado com uma base própria de 6.000 vídeos, o modelo alcançou

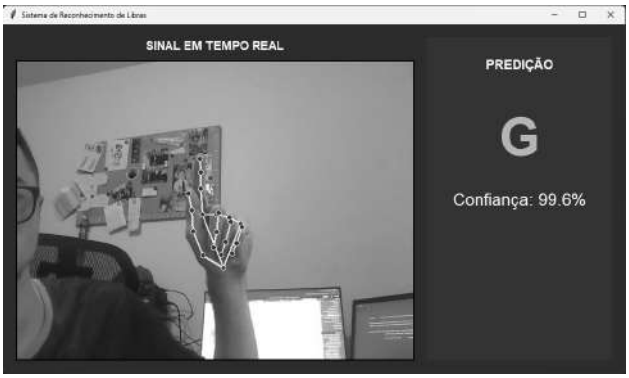


Fig. 6. Interface gráfica do sistema operando em tempo real, exibindo o rastreamento espacial no painel de vídeo e a respectiva predição da rede neural.

uma acurácia de 90% e um tempo médio de resposta de 40,89 ms, rodando exclusivamente na CPU.

Esses resultados comprovam que o sistema é rápido e viável para uso em computadores comuns. A análise dos dados mostrou que a maior dificuldade do modelo ainda é diferenciar gestos visualmente muito parecidos (como ’T’ e ’F’). No entanto, o uso da rede bidirecional ajudou a reduzir esse problema, resultando em uma aplicação final estável e integrada a uma interface gráfica amigável para o usuário.

Como trabalhos futuros, sugere-se expandir a base de dados para incluir mais voluntários e novos sinais, aumentando o vocabulário do sistema. Além disso, a evolução da rede para reconhecer frases contínuas, a inclusão de expressões faciais e a otimização do código para rodar em aplicativos de celular são os próximos passos para transformar este projeto em uma ferramenta de acessibilidade ainda mais completa.

Como perspectivas para trabalhos futuros, pretende-se ampliar o banco de dados para incluir novos sinais e uma maior quantidade de voluntários, diversificando características como faixa etária e biotipo. Além disso, planeja-se submeter o sistema a diferentes condições de iluminação para validar a robustez do modelo em cenários variados. A evolução da arquitetura para o reconhecimento de frases contínuas, a inclusão de expressões faciais e a portabilidade da solução para dispositivos móveis — aproveitando a baixa latência já demonstrada em hardware convencional — são passos fundamentais para transformar este projeto em uma ferramenta de acessibilidade ainda mais completa e versátil.

Referências

[1] Sílvia Andreis-Witkoski. *Introdução à libras: língua, história e cultura*. UTFPR Editora, 2015. ISBN: 9788570141439.

[2] Ronice Muller de Quadros, Aline Lemos Pizzio e Patrícia Luiza Ferreira Rezende. *Língua Brasileira de Sinais I*. Universidade Federal de Santa Catarina, 2009.

[3] Ministério dos Direitos Humanos e da Cidadania. *Brasil tem 18,6 milhões de pessoas com deficiência, indica pesquisa divulgada pelo IBGE e MDHC*. URL: https://www.gov.br/mdh/pt-br/assuntos/noticias/2023/julho/brasil-tem-18-6-milhoes-de-pessoas-com-deficiencia-indica-pesquisa-divulgada-pelo-ibge-e-mdhc (acesso em 29/03/2025).

[4] IBGE e MDHC. *Mais de 10 milhões de brasileiros apresentam algum grau de surdez*. URL: https://jornal.usp.br/atualidades/mais-de-10-milhoes-de-brasileiros-apresentam-algum-grau-de-surdez/#:~:text=Essa%20porcentagem%20significa%20que%20mais,%C3%A0%20dem%C3%AAncia%20se%20n%C3%A3o%20detectada. (acesso em 29/03/2025).

[5] Sachin Agrawal, Agnishrota Chakraborty e M. Rajalakshmi. "Real-Time Hand Gesture Recognition System Using MediaPipe and LSTM". Em: *International Journal of Research Publication and Reviews* 3 (2022), pp. 2509–2515. ISSN: 2582-7421. URL: https://ijrpr.com/uploads/V3ISSUE4/IJRPR3693.pdf.

[6] Meryem Cherrate, My Abdelouahed Sabri, Ali Yahyaouy e Abdellah Aarab. "Recognizing Sign Language Gestures Using a Hybrid Spatio-Temporal Deep Learning Model". Em: *Procedia Computer Science* 191 (2021), pp. 232–239.

[7] S. S. Sravani, K. V. Rao et al. "Hand Gesture Recognition Using MediaPipe and LSTM". Em: *2022 6th International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. IEEE, 2022, pp. 154–159.

[8] Gabriel dos S. L. Stefano, Wesley L. Passos, Jonathan N. Gois, Gabriel M. Araujo e Amaro A. de Lima. "Um sistema de reconhecimento de sinais em Libras usando CNN e LSTM". Em: *XXXIX Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT 2021)*. Fortaleza, CE, Brasil, 2021. URL: https://biblioteca.sbrt.org.br/articlefile/2860.pdf.

[9] Muneeb Ur Rehman, Fawad Ahmed, Muhammad Attique Khan, Usman Tariq, Faisal Abdulaziz Alfouzan, Nouf M. Alzahrani e Jawad Ahmad. "Dynamic Hand Gesture Recognition Using 3D-CNN and LSTM Networks". Em: *Computers, Materials & Continua* 70 (2022), pp. 4675–4690. ISSN: 1546-2226. DOI: https://doi.org/10.32604/cmc.2022.019586. URL: http://www.techscience.com/cmc/v70n3/44942.

[10] Adriel Vieira Santos, Iago Franco Bacurau, Jayne de Morais Silva e Talles Brito Viana. "UM MÓDULO DE SOFTWARE PARA O RECONHECIMENTO DE SINAIS ESTÁTICOS DA LÍNGUA BRASILEIRA DE SINAIS". Em: *Acta Kariri Pesq. e Des. Crato/CE* 2 (2017), pp. 84–90. URL: http://actakariri.crato.ifce.edu.br/index.php/actakariri/article/view/30/28.

[11] R. Sharma et al. "Deep Learning-based Real-time Sign Language Recognition System for Assistive Communication". Em: *2021 International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 2021.

[12] Rúbia Reis Guerra, Raoni Salgado de Amorim, Bruno Brandoli Machado e Amaro A. de Lima. "Desenvolvimento de uma Base de Dados de Sinais de Libras para Aprendizado de Máquina: Estudo de Caso com CNN 3D". Em: *Anais do XV Simpósio Brasileiro de Automação Inteligente (SBAI 2019)*. Ouro Preto, MG, Brasil, 2019. DOI: http://dx.doi.org/10.17648/sbai-2019-111451. URL: https://www.researchgate.net/publication/338202250_Desenvolvimento_de_uma_Base_de_Dados_de_Sinais_de_Libras_para_Aprendizado_de_Maquina_Estudo_de_Caso_com_CNN_3D.

[13] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang e Matthias Grundmann. "MediaPipe Hands: On-device Real-time Hand Tracking". Em: *CVPR Workshop on Computer Vision for Augmented and Virtual Reality*. 2020.

Percepções de Estudantes de Tecnologia da Informação sobre a utilização de forma ética do ChatGPT no Ensino

Patrícia Alves de Sousa, Samuel Sousa Teles, Luis Felipe de Freitas Gomes
Igor Ramsés Temóteo dos Santos, Rubens Abraão da Silva Sousa, Alex Lacerda Ramos
Instituto Federal de Educação, Ciência e Tecnologia do Ceará
patricia.sousa@darmlabs.ifce.edu.br, samuel.teles@darmlabs.ifce.edu.br, luis.gomes@darmlabs.ifce.edu.br,
igor.ramses@darmlabs.ifce.edu.br, rubens.silva@darmlabs.ifce.edu.br, alex.lacerda@ifce.edu.br

***Resumo*—Com o avanço da Inteligência Artificial (IA) Generativa, compreender como essa ferramenta se integra à ética em sala de aula tornou-se um desafio. É necessário entender como os alunos utilizam ferramentas como o ChatGPT de forma ética. O objetivo deste estudo é compreender como estudantes de graduação e técnicos de TI veem seu uso do ChatGPT de forma ética. Este estudo foi realizado por meio de um formulário com 10 perguntas em escala Likert com estudantes do IFCE, EEEP e UFC divididos em 5 áreas. O formulário mostrou que o ChatGPT é utilizado para apoiar o aprendizado na área de TI, mas ainda há preocupações com a autonomia.**

***Palavras-chave*—ChatGPT, Educação, Ética, Inteligência Artificial, Tecnologia da Informação.**

***Abstract*—With the advance of Generative Artificial Intelligence (AI), understanding how this tool integrates with classroom ethics has become a challenge. The aim of this study is to understand how I.T. undergraduate and technical students view their use of ChatGPT ethically. This study was conducted using a form with 10 questions on a Likert scale with students from IFCE, EEEP and UFC divided into 5 areas. The form showed that ChatGPT is used to support learning in the IT field, but there are still concerns about autonomy.**

***Palavras-chave*—Artificial Intelligence, ChatGPT, Education, Ethics, Information Technology.**

I. Introdução

Atualmente, presencia-se o intenso avanço no desenvolvimento de ferramentas tecnológicas para a facilitação de atividades humanas, sobretudo para o processamento de dados em larga escala e a aprendizagem de máquina (ML) a partir desses dados. A Inteligência Artificial tem sido integrada a vários serviços através da interpretação de dados, automação de sistemas e otimização de processos [1].

Em meio a considerável variedade de artefatos de IA disponíveis para uso, destaca-se o ChatGPT da OpenAI. Desde o seu lançamento de forma gratuita, em novembro de 2022, houve ampla disseminação de um chatbot de geração de texto humanizado e capaz de realizar com eficiência atividades que vão desde responder perguntas até traduzir idiomas e gerar conteúdo [2]. Compreendendo o crescimento do uso de IA na resolução de atividades diversas, inclusive educacionais, a aplicação rapidamente começou a ser utilizada em meio acadêmico.

A ferramenta se demonstrou promissora na interpretação e refinamento de textos, podendo até mesmo perceber a dificuldade do usuário e sugerir soluções, bem como passo a passo para desempenhar tarefas e aumentar a produtividade, se tornando muito útil na personalização de experiências de aprendizagem [3].

Apesar dos avanços da aplicabilidade do chatbot no contexto educacional ter se revelado eficiente, há muitas preocupações sobre seu uso de forma ética dentro de cursos de técnicos e de graduação, sendo necessário compreender como os estudantes utilizam ferramentas como o ChatGPT no ensino [4].

Saylam et al. [5] apontam em seu estudo a preocupação e seu uso para a substituição da realização de atividades curriculares que deveriam ser executadas pelos próprios estudantes, tornando essa uma das preocupações éticas mais discutidas. Entretanto, nota-se também a utilização do ChatGPT como um recurso de alto potencial pedagógico na prática de distinção de necessidades de aprendizado, a partir das demandas informadas.

Diante disso, compreender como estudantes da área de Tecnologia da Informação (T.I) utilizam o ChatGPT para a área acadêmica se faz necessário. Assim, este trabalho tem como objetivo analisar esse uso por meio da aplicação de um formulário eletrônico baseado em escala Likert estruturado em cinco áreas: i) Conhecimento e Familiaridade, ii) Percepção sobre uso pedagógico, iii) Compreensão sobre o Uso Ético e Responsável, iv) Impacto na Autonomia e nos Papéis e v) Perspectiva geral.

O trabalho está dividido nas seguintes seções: i) introdução: que apresenta conceitos iniciais e a problemática que deu embasamento para o estudo; ii) trabalhos relacionados: apresenta trabalhos que buscaram compreender o uso de IAs generativas na educação; iii) método: apresenta a abordagem adotada para aplicação do formulário; iv) resultados e discussões, apresenta os resultados e discute com a literatura os achados; e v) conclusão: traz a síntese do trabalhos com bases nos achados.

II. Trabalhos Relacionados

Esta seção reúne e analisa estudos recentes que investigam o uso do ChatGPT na educação, abrangendo desde o

ensino básico até o ensino superior. O objetivo é compreender como a ferramenta tem sido integrada em diferentes contextos pedagógicos, destacando suas potencialidades, limitações e implicações éticas.

O estudo de Cotton et al. [2] investiga o uso do ChatGPT no ensino superior, com foco no contexto acadêmico. Os autores percebem que, embora o modelo permita aprendizagem sob medida, respostas rápidas e acessibilidade, ele também gera preocupações sobre plágio e desigualdade no acesso. O artigo expõe maneiras de usar a IA de forma ética, criando regras claras nas instituições, capacitando professores e combinando avaliações de máquinas e pessoas para diminuir os riscos de desonestidade.

Zhang e Tur [6] executaram uma análise completa sobre o uso do ChatGPT na condição K-12, tendo em vista, no contexto voltado para a educação básica. Eles encontraram 13 artigos e fizeram uma análise SWOT. Entre os pontos positivos, estão a criação de conteúdo personalizado, o suporte no planejamento das aulas e o incentivo ao aprendizado em grupo. No entanto, o estudo mostra que a qualidade das respostas pode ser fraca e que é preciso treinar os professores para usar a ferramenta de forma eficaz e ética. O estudo também destaca a importância da criação conjunta de experiências de aprendizado com IA, com professores e alunos trabalhando juntos em atividades relevantes.

Complementarmente, Zeb et al. [7] discutem o papel do uso do ChatGPT no ensino superior, enfatizando as possibilidades de transformação pedagógica e os riscos associados ao seu uso indiscriminado. O artigo explora tanto as reações entusiásticas quanto céticas ao surgimento da tecnologia e propõe que as universidades adotem posturas proativas, com políticas claras de uso, ações de conscientização sobre integridade acadêmica e adaptações nos formatos de avaliação. Além disso, os autores destacam a importância de equilibrar inovação tecnológica com princípios pedagógicos sólidos, evitando tanto a tecnofilia exagerada quanto o receio infundado frente à inovação.

Rajan e Niranjan [8] definem o uso do ChatGPT como uma "espada de dois gumes" no ambiente educacional, especialmente no ensino superior. O estudo evidencia que, quando usado de forma orientada e integrada a atividades reflexivas, o ChatGPT pode estimular significativamente o pensamento crítico e a criatividade dos estudantes. No entanto, os autores alertam que a utilização demasiada ou sem mediação pode acabar levando a uma dependência cognitiva e à perda de autonomia intelectual. Assim, é reforçado a necessidade de práticas pedagógicas que incluam diretrizes éticas e o acompanhamento docente para garantir que a ferramenta seja um apoio e não um substituto do raciocínio próprio.

Al-Kfairy [9] realizou uma revisão de estudos empíricos que utilizaram o Modelo de Aceitação de Tecnologia (TAM) para entender como docentes estão adotando o ChatGPT em suas práticas pedagógicas. Os resultados mostram que os professores reconhecem o potencial da ferramenta por sua facilidade de uso e aplicabilidade no planejamento de aulas. Porém, existem barreiras como a falta de treinamento específico e a falta de políticas institucionais claras sobre o uso da IA ainda dificultam uma adoção mais ampla. O estudo conclui que, além do acesso à tecnologia, é essencial oferecer formação continuada e estabelecer normas de uso seguras e bem definidas.

Em síntese, os estudos analisados evidenciam o potencial e os desafios do uso do ChatGPT na educação. No entanto, há uma lacuna quanto à percepção de alunos de cursos técnicos e superiores em tecnologia sobre o uso ético da ferramenta. Este artigo busca contribuir nesse sentido ao investigar, por meio de um questionário em escala Likert, o conhecimento e a percepção dos estudantes do ensino médio e superior de cursos da área de tecnologia, vinculados ao Instituto Federal do Ceará (IFCE), à Universidade Federal do Ceará (UFC) e à Escola Estadual de Educação Profissional (EEEP) , sobre o uso ético do ChatGPT como aliado ético e pedagógico no processo de aprendizagem acadêmica.

III. MÉTODO

Esta seção apresenta os procedimentos adotados para a realização deste trabalho. A metodologia adotada consiste em três etapas: i) Delineamento da pesquisa, a qual apresenta a abordagem utilizada para análise e compreensão do uso do ChatGPT de forma ética nos cursos da área de Tecnologia da Informação (TI); ii) Instrumento de coleta de dados, que apresenta a criação e estruturação do formulário e sua aplicação junto aos estudantes; e iii) Análise dos dados, que mostra o formato da análise dos dados.

A pesquisa adotou uma abordagem quantitativa de caráter descritivo, utilizando uma escala Likert de cinco pontos aplicada em cinco áreas conforme a Tabela I, com o objetivo de compreender a percepção dos participantes sobre o uso da IA generativa. O estudo contou com a participação de 120 estudantes: 23 alunos do ensino médio técnico do IFCE, 30 da Escola Estadual de Educação Profissional (EEEP) do Ceará, 7 do curso técnico subsequente, 58 de graduação do IFCE e 2 da Universidade Federal do Ceará (UFC).

Tabela I
DESCRIÇÃO DAS ÁREAS E OBJETIVOS DE COMPREENDIMENTO

Áreas	**Objetivo**
Conhecimento e Familiaridade	Compreender quanto o participante está familiarizado com o uso do ChatGPT ou outras IAs generativas, considerando sua frequência e compreensão das funcionalidades.
Percepção sobre uso pedagógico	Compreender como o participante percebe o potencial do ChatGPT na melhoria da qualidade do ensino e sua utilidade como apoio à aprendizagem.
Uso Ético e Responsável	Compreender o nível de segurança sobre práticas éticas, incluindo prevenção de plágio e identificação de usos inadequados ou antiéticos.
Impacto na Autonomia	Refletir sobre como o uso do ChatGPT pode influenciar o desenvolvimento da autonomia, do pensamento crítico e do papel ativo do aluno.
Perspectiva geral	Analisar a visão geral sobre o uso do ChatGPT na educação, no equilíbrio entre obter respostas prontas e estimular a criatividade e responsabilidade.

O processo de coleta de dados seguiu as seguintes etapas: (1) convite aos estudantes do IFCE, EEEP e UFC; (2) pre-

paração do ambiente de laboratório para discentes do IFCE para acesso ao Google Forms; (3) o instrumento de coleta consistiu em um formulário contendo 10 perguntas, conforme LINK (https://doi.org/10.5281/zenodo.15604581), distribuídas em cinco áreas temáticas, com duas questões destinadas a cada área

A análise foi realizada por meio de estatística descritiva, considerando a frequência absoluta das respostas, e os resultados foram apresentados em tabelas. Para participar do estudo, os respondentes precisaram aceitar o Termo de Consentimento Livre e Esclarecido (TCLE), e todos os dados foram anonimizados, a fim de garantir que os participantes não fossem identificados.

IV. RESULTADOS E DISCUSSÕES

Esta seção apresenta os resultados obtidos e discute com a literatura as implicações e evidências empíricas encontradas no contexto do ensino superior e médio. Apresenta-se uma análise descritiva das respostas por meio de tabelas demonstrando a distribuição por frequência dos resultados obtidos por cada pergunta. A seguir a Tabela II apresenta a distribuição de respostas para a pergunta 1 do questionário.

Tabela II
EU ME SINTO FAMILIARIZADO COM AS FUNCIONALIDADES E CAPACIDADES EM USAR O CHATGPT?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	62	51,7%
Concordo	36	30%
Não concordo, nem discordo	17	14,2%
Discordo	4	3,3%
Discordo Totalmente	1	0,8%

De acordo com a Tabela II, há uma concordância majoritária sobre a familiaridade com as funcionalidades do ChatGPT: 62 participantes (51,7%) concordam totalmente, e 36 participantes (30%) concordam. Isso indica que a maioria dos respondentes se sente familiarizada com o uso da ferramenta. Elhassan et al. [10] evidencia em uma pesquisa realizada com estudantes de medicina na Alfaisal University (Arábia Saudita), que mais de 90% dos estudantes relataram conhecer a ferramenta, cuja porcentagem é similar à encontrada de 81,7% de familiaridade, somando os resultados, o que indica uma crescente familiaridade com as funcionalidades do ChatGPT no contexto educacional. A seguir a Tabela III apresenta a distribuição de respostas para a pergunta 2 do questionário.

Tabela III
EU UTILIZO O CHATGPT OU OUTRA IA GENERATIVA COM FREQUÊNCIA?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	39	32,5%
Concordo	48	40%
Não concordo, nem discordo	22	18,3%
Discordo	7	5,8%
Discordo Totalmente	4	3,3%

De acordo com a Tabela III, observa-se uma concordância majoritária em relação à frequência de uso do ChatGPT ou de outras IAs generativas: 39 participantes (32,5%) concordam totalmente e 48 participantes (40%) concordam. Isso indica que 72,5% dos respondentes utilizam essas ferramentas com frequência, revelando não apenas familiaridade, mas também engajamento ativo com tecnologias de inteligência artificial. Por outro lado, uma minoria expressiva de 27,5% dos participantes demonstrou pouca ou nenhuma frequência no uso. Krause et al. [11] demonstra que a maior parte dos estudantes (42,45%) usa o ChatGPT de forma moderada, algumas vezes por mês, enquanto um número considerável (34,30%) já o usa várias vezes por semana, indicando que a ferramenta está se tornando um recurso regular nos estudos. Esses dados reforçam a crescente presença da IA generativa no cotidiano dos respondentes, especialmente em contextos educacionais. A seguir a Tabela IV apresenta a distribuição de respostas para a pergunta 3 do questionário.

Tabela IV
O CHATGPT OU OUTRA IA GENERATIVA PODE MELHORAR A QUALIDADE DE ENSINO?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	40	33,3%
Concordo	56	46,7%
Não concordo, nem discordo	16	13,3%
Discordo	4	3,3%
Discordo Totalmente	4	3,3%

De acordo com a Tabela IV, a grande maioria dos respondentes acredita que o uso do ChatGPT ou de outras IAs generativas pode melhorar a qualidade do ensino. Isso é evidenciado pelos 40 participantes (33,3%) que concordam totalmente e pelos 56 participantes (46,7%) que concordam, totalizando 76,6% de concordância, argumento esse assegurado também pela pequena parcela dos respondentes que demonstra neutralidade (13,3%) ou discordância (6,6% no total), o que reforça a percepção de que essas ferramentas são vistas, majoritariamente, como aliadas no processo educacional, resultado que também é reforçado por Farrokhnia et al [12] que demonstra que o ChatGPT pode aumentar o acesso à informação, facilitar o aprendizado personalizado e diminuir a carga de trabalho de ensino. A seguir a Tabela V apresenta a distribuição de respostas para a pergunta 4 do questionário.

Tabela V
O CHATGPT PODE SER UTILIZADO COMO FERRAMENTA PARA AUXILIAR O APRENDIZADO?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	58	48,3%
Concordo	52	43,3%
Não concordo, nem discordo	5	4,2%
Discordo	1	0,8%
Discordo Totalmente	4	3,3%

De acordo com a Tabela V, a imensa maioria dos participantes acredita que o ChatGPT pode ser utilizado como uma ferramenta para auxiliar o aprendizado. Isso é evidenciado pelos 58 participantes (48,3%) que concordam totalmente e 52 participantes (43,3%) que concordam. Isso indica que 91,6%

dos respondentes reconhecem o potencial da ferramenta como apoio educacional, revelando uma ampla aceitação de seu uso no processo de ensino-aprendizagem. Por outro lado, uma minoria de 9 (8,3%) dos participantes demonstrou neutralidade ou discordância quanto a essa possibilidade. Rejeb et al. [13] apresenta que o ChatGPT é uma ferramenta educacional essencial que aprimora as habilidades de escrita dos alunos e promove ambientes de aprendizagem dinâmicos e interativos. A seguir a Tabela VI apresenta a distribuição de respostas para a pergunta 5 do questionário.

Tabela VI
EU SINTO SEGURO SOBRE AS MELHORES PRÁTICAS DE COMO USAR O CHATGPT OU OUTRA IA GENERATIVA DE FORMA ÉTICA, EVITANDO PLÁGIO?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	22	18,3%
Concordo	60	50%
Não concordo, nem discordo	20	16,7%
Discordo	13	10,8%
Discordo Totalmente	1	0,8%

Com base nos dados da Tabela VI, observa-se que a maior parte dos participantes demonstra confiança quanto ao uso responsável do ChatGPT e de outras IAs generativas, sobretudo no que se refere à prevenção de plágio. Isso se reflete nos 22 participantes (18,3%) que afirmaram estar totalmente de acordo, além dos 60 (50%) que também expressaram concordância. Assim, é possível afirmar que 68,3% dos respondentes se sentem preparados para empregar essas ferramentas de forma ética. Em contrapartida, 34 participantes (28,3%) adotaram uma posição neutra ou demonstraram algum grau de discordância, o que sugere a necessidade de mais orientações e discussões sobre práticas seguras e responsáveis nesse contexto. Michel-Villarreal et al. [14] apresenta que o ChatGPT oferece benefícios e oportunidades potenciais para o ensino superior, mas exige políticas, diretrizes e pesquisas empíricas claras para enfrentar os desafios e garantir a integridade acadêmica. A seguir a Tabela VII apresenta a distribuição de respostas para a pergunta 6 do questionário.

Tabela VII
EU SEI IDENTIFICAR O USO INADEQUADO E/OU ANTIÉTICO DO CHATGPT OU IA GENERATIVA?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	31	25,8%
Concordo	57	47,5%
Não concordo, nem discordo	22	18,3%
Discordo	7	5,8%
Discordo Totalmente	3	2,5%

De acordo com a Tabela VII, observa-se uma concordância significativa em relação à capacidade dos respondentes de identificar o uso inadequado e/ou antiético do ChatGPT ou de outras IAs generativas: 31 participantes (25,8%) concordam totalmente e 57 participantes (47,5%) concordam. Isso indica que 73,3% dos respondentes afirmam saber reconhecer práticas inadequadas envolvendo o uso dessas ferramentas. Por outro lado, uma parcela de 26,7% demonstra neutralidade ou discordância, o que sugere que ainda há espaço para fortalecer a formação ética e crítica no uso da inteligência artificial. Chen et al. [15] revela que estudantes demonstram consciência crítica: reconhecem possíveis falhas nos resultados de ferramentas como ChatGPT e fazem esforços para verificação cruzada, indicando percepção sobre usos éticos e antiéticos. A seguir a Tabela VIII apresenta a distribuição de respostas para a pergunta 7 do questionário.

Tabela VIII
EU SINTO QUE O CHATGPT OU OUTRA IA GENERATIVA INFLUENCIA NEGATIVAMENTE NO DESENVOLVIMENTO E AUTONOMIA?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	10	8,3%
Concordo	30	25%
Não concordo, nem discordo	55	45,8%
Discordo	20	16,7%
Discordo Totalmente	5	4,2%

A Tabela VIII revela que há uma divisão nas opiniões sobre o impacto do ChatGPT e outras IAs generativas no desenvolvimento e na autonomia dos usuários. Apenas uma parte dos participantes acredita que essas ferramentas exercem uma influência negativa, somando 33,3% de concordância. A maior parte dos respondentes, no entanto, manteve-se neutra, com 55 pessoas (45,8%) não se posicionando claramente. Esses dados sugerem ausência de consenso entre os participantes sobre os impactos da IA no desenvolvimento da autonomia. Essa tendência de neutralidade também pode estar relacionada ao que é apontado por Deng et al. [16] que demonstra que o uso excessivo de ferramentas de IA, como o ChatGPT, pode gerar dependência cognitiva, o que compromete a autonomia dos estudantes. A seguir a Tabela IX apresenta a distribuição de respostas para a pergunta 8 do questionário.

Tabela IX
EU SINTO QUE O CHATGPT O OUTRA IA GENERATIVA PODE DIMINUIR O DESENVOLVIMENTO DO PENSAMENTO CRÍTICO E DA AUTONOMIA?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	23	19,2%
Concordo	42	35%
Não concordo, nem discordo	44	36,7%
Discordo	9	7,5%
Discordo Totalmente	2	1,7%

De acordo com a Tabela IX, observa-se que 54,2% dos participantes concordam que o uso do ChatGPT ou de outras IAs generativas pode diminuir o desenvolvimento do pensamento crítico e da autonomia. Contudo, uma parcela considerável dos respondentes (44 participantes, 36,7%) manteve-se neutra, sugerindo a necessidade de debates mais aprofundados sobre o tema. Farooqi et al. [17] destacam que a percepção dos impactos da IA ainda é dividida entre educadores e alunos. A seguir a Tabela X apresenta a distribuição de respostas para a pergunta 9 do questionário.

De acordo com a Tabela X, observa-se que 62,5% de concordância que veem o ChatGPT ou outra IA generativa

Tabela X
EU VEJO O CHATGPT OU OUTRA IA GENERATIVA COMO UMA FERRAMENTA PARA OBTER RESPOSTAS RÁPIDAS DO QUE ESTIMULAR A CRIATIVIDADE E APROFUNDAMENTO?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	26	21,7%
Concordo	49	40,8%
Não concordo, nem discordo	24	20%
Discordo	18	15%
Discordo Totalmente	3	2,5%

mais como uma ferramenta para obter respostas rápidas do que para estimular a criatividade. Habib et al. [18] apresenta que a IA tem o potencial de dar suporte ao pensamento criativo, mas também tem impactos negativos na confiança criativa. A seguir a Tabela XI apresenta a distribuição de respostas para a pergunta 10 do questionário.

Tabela XI
EU VEJO QUE AO USAR O CHATGPT OU OUTRA IA GENERATIVA A VERIFICAÇÃO DA PRECISÃO DAS RESPOSTAS É DE RESPONSABILIDADE DO USUÁRIO?

Resposta	**Qtd**	**% Total**
Concordo Totalmente	37	30,8%
Concordo	58	48,3%
Não concordo, nem discordo	17	14,2%
Discordo	8	6,7%
Discordo Totalmente	0	0%

De acordo com a Tabela XI, 79,1% dos respondentes compreendem que cabe ao próprio usuário a responsabilidade de verificar a precisão das respostas. Esse posicionamento é respaldado por Liu et al. [19], que demonstram que apenas cerca de 51,5% das frases geradas por motores de busca generativos são totalmente verificáveis.

Diante dos resultados apresentados, observa-se ampla familiaridade e aceitação do uso do ChatGPT no contexto educacional. Entretanto, persistem preocupações relacionadas aos possíveis impactos na autonomia dos estudantes, no desenvolvimento do pensamento crítico e no uso ético e responsável dessas tecnologias.

V. CONCLUSÃO

O presente estudo evidencia que o ChatGPT vem se mostrando como uma ferramenta conhecida e utilizada de forma ampla por estudantes do ensino técnico e superior de cursos da área de tecnologia. Os dados indicam que mais de 80% dos participantes estão familiarizados com as funcionalidades da ferramenta e relatam utilizá-la com frequência. A percepção predominante é de que o ChatGPT contribui positivamente para o processo de ensino, tanto na melhoria da qualidade de ensino quanto como auxílio aos estudos.

Em relação à ética no uso, a maioria dos participantes afirmam se sentir segura quanto às práticas para evitar o plágio e identificar uso inadequado e/ou antiético — embora cerca de 20% afirmem incerteza, sinalizando a necessidade de reforço na formação ética sobre o uso da IA.

A questão da autonomia gerou percepções mais heterogêneas. Quase metade dos participantes acredita que o uso da IA pode impactar negativamente o desenvolvimento do pensamento crítico, evidenciando que, embora amplamente reconhecida, existem preocupações quanto à sua influência na construção da autonomia intelectual. Além disso, a maioria reconhece que a responsabilidade pela verificação da precisão das respostas é do próprio usuário, o que sugere uma postura crítica em relação ao conteúdo.

Esta pesquisa apresenta limitações relevantes, como o foco exclusivo na área de Tecnologia da Informação (T.I.) e o período específico de coleta de dados. Recomenda-se, em trabalhos futuros, a ampliação do estudo para incluir cursos de outras áreas e a realização de pesquisas longitudinais para acompanhar a evolução destas percepções. Conclui-se que o ChatGPT é um aliado no aprendizado, mas é essencial que as instituições implantem políticas claras e estratégias que estimulem o pensamento crítico e a autonomia.

REFERÊNCIAS

[1] S. Choudari, R. Sanwal, N. Sharma, S. Shastri, D. A. P. Singh, and G. Deepa, "Data science collaboration in human ai: Decision optimization using human-centered automation," in *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*. IEEE, 11 2024, pp. 1–6.

[2] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of chatgpt," *Innovations in Education and Teaching International*, vol. 61, pp. 228–239, 3 2024.

[3] S. Begum, M. Vimochana, V. Arunadevi, A. C. A. Dhivya, V. A. Vuyyuru, and M. Faizal, "Using ai for personalized english writing feedback on educational platforms," in *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*. IEEE, 12 2024, pp. 1–6.

[4] E. Zabala and H. S. Narman, "Development and evaluation of an ai-enhanced python programming education system," in *2024 IEEE 15th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. IEEE, 10 2024, pp. 787–792.

[5] S. Saylam, N. Duman, Y. Yildirim, and K. Satsevich, "Empowering education with ai: Addressing ethical concerns," *London Journal of Social Sciences*, pp. 39–48, 9 2023.

[6] P. Zhang and G. Tur, "A systematic review of chatgpt use in k-12 education," *European Journal of Education*, vol. 59, 6 2024.

[7] A. Zeb, R. Ullah, and R. Karim, "Exploring the role of chatgpt in higher education: opportunities, challenges and ethical considerations," *The International Journal of Information and Learning Technology*, vol. 41, pp. 99–111, 1 2024.

[8] S. Rajan and L. Niranjan, "The double-edged sword of chatgpt: fostering and hindering creativity in postgraduate academics in bengaluru," *International Journal of Educational Management*, vol. 39, pp. 317–337, 2 2025.

[9] M. Al-kfairy, "Factors impacting the adoption and acceptance of chatgpt in educational settings: A narrative review of empirical studies," *Applied System Innovation*, vol. 7, p. 110, 11 2024.

[10] S. E. Elhassan, M. R. Sajid, A. M. Syed, S. A. Fathima, B. S. Khan, and H. Tamim, "Assessing familiarity, usage patterns, and attitudes of medical students toward chatgpt and other chat-based ai apps in medical education: Cross-sectional questionnaire study," *JMIR Medical Education*, vol. 11, 2025.

[11] S. Krause, B. H. Panchal, and N. Ubhe, "Evolution of learning: Assessing the transformative impact of generative ai on higher education," *Frontiers of Digital Education*, vol. 2, p. 21, 6 2025.

[12] M. Farrokhnia, S. K. Banihashem, O. Noroozi, and A. Wals, "A swot analysis of chatgpt: Implications for educational practice and research," *Innovations in Education and Teaching International*, vol. 61, pp. 460–474, 5 2024.

[13] A. Rejeb, K. Rejeb, A. Appolloni, H. Treiblmaier, and M. Iranmanesh, "Exploring the impact of chatgpt on education: A web mining and machine learning approach," *The International Journal of Management Education*, vol. 22, p. 100932, 3 2024.

[14] R. Michel-Villarreal, E. Vilalta-Perdomo, D. E. Salinas-Navarro, R. Thierry-Aguilera, and F. S. Gerardou, "Challenges and opportunities of generative ai for higher education as explained by chatgpt," *Education Sciences*, vol. 13, p. 856, 8 2023.

[15] K. Chen, A. C. Tallant, and I. Selig, "Exploring generative ai literacy in higher education: student adoption, interaction, evaluation and ethical perceptions," *Information and Learning Sciences*, vol. 126, pp. 132–148, 2 2025.

[16] R. Deng, M. Jiang, X. Yu, Y. Lu, and S. Liu, "Does chatgpt enhance student learning? a systematic review and meta-analysis of experimental studies," *Computers & Education*, vol. 227, p. 105224, 4 2025.

[17] M. T. K. Farooqi, I. Amanat, and S. M. Awan, "Ethical considerations and challenges in the integration of artificial intelligence in education: A systematic review," *Journal of Excellence in Management Sciences*, vol. 3, pp. 35–50, 9 2024.

[18] S. Habib, T. Vogel, X. Anli, and E. Thorne, "How does generative artificial intelligence impact student creativity?" *Journal of Creativity*, vol. 34, p. 100072, 4 2024.

[19] N. F. Liu, T. Zhang, and P. Liang, "Evaluating verifiability in generative search engines," 4 2023. [Online]. Available: http://arxiv.org/abs/2304.09848

AUTORES



Patrícia Alves de Sousa Graduanda em Análise e Desenvolvimento de Sistemas pelo IFCE (Campus Canindé). Atualmente, desenvolve atividades de análise de requisitos no Laboratório DARM e atua como pesquisadora de Iniciação Científica e gerente de projetos no Laboratório de Aprendizagem de Máquina e Projetos Inovadores (LAMPI/IFCE).



Luis Felipe de Freitas Gomes Graduando em Análise e Desenvolvimento de Sistemas pelo Instituto Federal do Ceará. Atua na área de Testes e Qualidade de Software, com experiência em gestão da qualidade e liderança técnica em projetos de Pesquisa, Desenvolvimento e Inovação no DARM Labs, em parceria com empresas nacionais e multinacionais. Tem interesse em Engenharia de Software, Testes de Software e Garantia da Qualidade, com foco na confiabilidade e melhoria contínua de sistemas.



Samuel Sousa Teles Mestrando em Ciência da Computação pela Universidade Estadual do Ceará, é graduado em Redes de Computadores pelo Instituto Federal do Ceará (2025). Possui experiência em Engenharia de Software, com ênfase em Qualidade de Software e Detecção de Bugs, atuando principalmente nos temas de desenvolvimento seguro e detecção de bugs.



Igor Ramsés Temóteo dos Santos Mestrando em Ciência da Computação pela Universidade Estadual do Ceará, é graduado em Análise e Desenvolvimento de Sistemas pelo Instituto Federal do Ceará (2025). Possui experiência em Engenharia de Software, com ênfase em Qualidade de Software e Detecção de Bugs, atuando principalmente nos temas de desenvolvimento seguro e detecção de bugs.



Rubens Abraão da Silva Sousa Mestrando em Ciência da Computação (UECE) e Graduado em ADS (IFCE). Atua como Líder de Qualidade de Software e Requisitos no DARM Labs e pesquisador no LAMPI, com foco em Engenharia de Software e Testes.



Alex Lacerda Ramos atua como professor no Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE). É Doutor e Mestre em Informática Aplicada pela Universidade de Fortaleza (UNIFOR), com bolsa PROSUP-CAPES. Possui especialização em Docência pelo IFCE e pela Universidade Católica Dom Bosco (UCDB). É graduado em Análise e Desenvolvimento de Sistemas pelo Instituto Federal do Piauí (IFPI). Tem experiência na área de Ciência da Computação, com ênfase em Desenvolvimento de Sistemas de Informação e Redes de Computadores. Suas áreas de interesse incluem Internet das Coisas (IoT), computação em nuvem, inteligência artificial, aprendizado de máquina e deep learning.

Collaborative Tool to Ease the Collection of Datasets for training Computer Vision Models for Retail Establishments

Marcelo Henrique de Oliveira Pereira Vilela & Renzo Paranaiba Mesquita
Instituto Nacional de Telecomunicações - Inatel
marcelo.henrique@gec.inatel.br, renzo@inatel.br

***Abstract*— The use of computer vision in retail has gained prominence due to its potential to automate product detection and improve operational efficiency. However, the development of robust machine learning models requires access to high-quality datasets, which are often scattered or difficult to find. This study presents the development of a website that serves as a centralized repository of dataset links for retail product recognition. The platform aims to facilitate access to curated datasets, supporting researchers and developers in training and improving computer vision models. The project includes dataset categorization, search functionalities, and community contributions to ensure continuous updates. The results demonstrate that the proposed solution simplifies dataset discovery, contributing to the advancement of AI applications in the retail sector.**

***Index Terms*— Computer vision, retail, dataset repository, machine learning, product recognition.**

I. Introduction

Computer Vision is a branch of Artificial Intelligence that enables machines to interpret and understand visual information extracted from images and videos. In recent years, advances in deep learning techniques have significantly improved the performance of computer vision systems, enabling their application in several domains such as healthcare, autonomous vehicles, manufacturing, and retail environments [1].

In the retail sector, computer vision technologies have been increasingly adopted to automate processes such as product recognition, inventory monitoring, and customer behavior analysis. These applications rely on machine learning models capable of detecting and classifying products from images captured in real-world environments [2].

Despite the rapid evolution of model architectures and training techniques, one of the main challenges in developing computer vision systems is the acquisition of suitable datasets. High-quality annotated datasets are essential for training robust models, however, such datasets are often difficult to find, fragmented across different platforms, or not specifically designed for certain domains such as retail [3].

To address this issue, this work proposes a collaborative platform designed to facilitate the discovery and collection of datasets focused on retail products. The platform acts as a centralized repository of dataset links, allowing users to search datasets using filtering mechanisms, access learning resources for model training, and contribute new datasets through a collaborative workflow based on GitHub.

The remainder of this paper is organized as follows: Section II presents the background concepts related to computer vision and collaborative dataset platforms. Section III describes the materials and methods used in the development of the proposed system. Section IV presents the obtained results and usability evaluation. Finally, Section V concludes the paper and discusses possible future work.

II. Background

A. Computer Vision

Computer Vision is a field of artificial intelligence that enables machines to interpret visual information extracted from images and videos. Instead of merely enhancing images, computer vision systems aim to recognize patterns, understand context, and support decision-making based on visual data. Techniques such as object detection, segmentation, and image classification are widely used, and modern approaches rely heavily on deep learning to improve performance across domains including healthcare, autonomous vehicles, and surveillance systems [4].

Several architectures have been developed to perform visual recognition tasks efficiently. Among them, You Only Look Once (YOLO) stands out for its real-time object detection capability, using a single-stage detection pipeline that combines high speed and accuracy. Since its first version, the model has evolved through multiple iterations such as YOLOv3, YOLOv5, YOLOv7, YOLOv8 and YOLO-NAS, incorporating improvements such as attention mechanisms and modular architectures [5]. In this work, YOLO was selected due to its efficiency, ease of implementation, and strong performance in real-time detection tasks.

Other well-known detection frameworks include Faster R-CNN, which uses region proposal networks for high accuracy, Single Shot MultiBox Detector (SSD), which balances speed and accuracy, and DETR, a transformer-based architecture that applies attention mechanisms to object detection tasks [6].

Training a computer vision model typically follows a structured pipeline composed of dataset collection and annotation, data preprocessing (such as normalization and augmentation), model training using labeled data, validation and testing to evaluate performance, and finally deployment for real-world inference [4].

B. Related Works

Several platforms have been created to facilitate access to datasets for machine learning and computer vision research. Roboflow Universe provides a large collection of annotated datasets for object detection across multiple domains [7]. The UCI Machine Learning Repository hosts datasets such as the Online Retail dataset, which supports studies on transaction patterns and customer behavior [8]. Similarly, Kaggle provides a collaborative environment where users can share datasets and machine learning projects, including those related to retail scenarios [9]. Google Dataset Search aggregates datasets from multiple repositories, enabling users to locate data across different domains through metadata indexing [10].

Although these platforms provide valuable resources, most of them are designed as general-purpose repositories rather than domain-specific solutions. As a result, developers working with retail-focused computer vision systems often face difficulties locating curated datasets suitable for product recognition tasks. The platform proposed in this work aims to address this limitation by providing a collaborative repository specifically focused on retail product datasets.

III. Materials and Methods

A. GitHub

GitHub is an online platform for hosting and collaboratively developing software projects based on the Git distributed version control system. Since its launch in 2008, it has become the largest social coding platform, combining code management with collaborative and social features for developers [11]. By providing public repositories, version tracking, and community interaction, GitHub enables developers to collaborate efficiently on software projects worldwide. As of 2021, the platform hosts more than 73 million developers and numerous open-source and industrial projects, making it a central component of the modern software development ecosystem [12].

Collaboration on GitHub is mainly supported through the fork-and-pull request model, where contributors create independent copies of repositories and propose modifications that are reviewed before integration [13]. This workflow facilitates external collaboration while maintaining project quality control. Additional features such as issue tracking systems allow teams to manage bugs, feature requests, and development tasks within the same environment [14].

GitHub also provides automation and integration tools that support modern development workflows. Pull requests enable structured discussions around code changes [13], while GitHub Actions allows the automation of tasks such as building, testing, and deployment in response to repository events [15]. These continuous integration mechanisms improve development efficiency, although they require careful configuration and monitoring [16]. Additionally, GitHub Copilot provides AI-assisted code suggestions within development environments, which can increase productivity while still requiring human oversight to ensure code quality [17].

In this project, several GitHub features were used to support both system development and collaborative dataset contribution. GitHub Pages was employed to host the platform frontend as a static website [18]. The pull request workflow enables contributors to submit new dataset entries while maintaining a structured review process. In addition, GitHub Actions automatically validates submitted files to ensure compliance with the required format. Together, these mechanisms enable the creation of an open and low-cost collaborative platform for storing and curating datasets for retail computer vision applications.

B. Method

The block diagram presented in Figure 1 illustrates the main functionalities and user flows of the collaborative platform developed to facilitate the collection and use of datasets for training computer vision models in retail.

Once the user accesses the website, it is possible to follow three main options that represent the core functionalities of the platform:

- Find Dataset: Search for datasets already registered inside the platform to use in computer vision applications.
- Learn YOLO: Access learning resources to understand how to train models.
- Contribute: Submit a new dataset to the platform.

Each of the options will be better explained in the following subtopics.

b.1. Path 1 – Finding a Dataset

- **Choose the best filter for what you need:** Users searching for datasets can filter by specific labels (e.g., Beans, Milk, Cofee) to narrow down results to those relevant to their problem.
- **Choose the best dataset for your problem:** After filtering, users review the available options and select the dataset that best fits their use case.
- **Click the dataset link:** Each dataset entry includes a direct link for access.
- **Redirected to external dataset page:** Clicking the link takes the user to the dataset's original hosting location (e.g., Google Drive, Roboflow, or another GitHub repository).
- **Download the dataset:** Users can then download the dataset for use in their own training project.

b.2. Path 2 – Learning YOLO

- **Training Tips:** If the user selects "Training Tips:", the platform provides access to a curated section with helpful training resources.
- **Select YOLO version:** Users can choose the version of YOLO (e.g., v5, v8) that matches their training needs.
- **Redirected to YOLOvN documentation:** After selecting a version, users are redirected to the official documentation or trusted third-party guides.
- **Follow the YOLOvN guide:** These resources support users, especially beginners, in training their own object detection models.

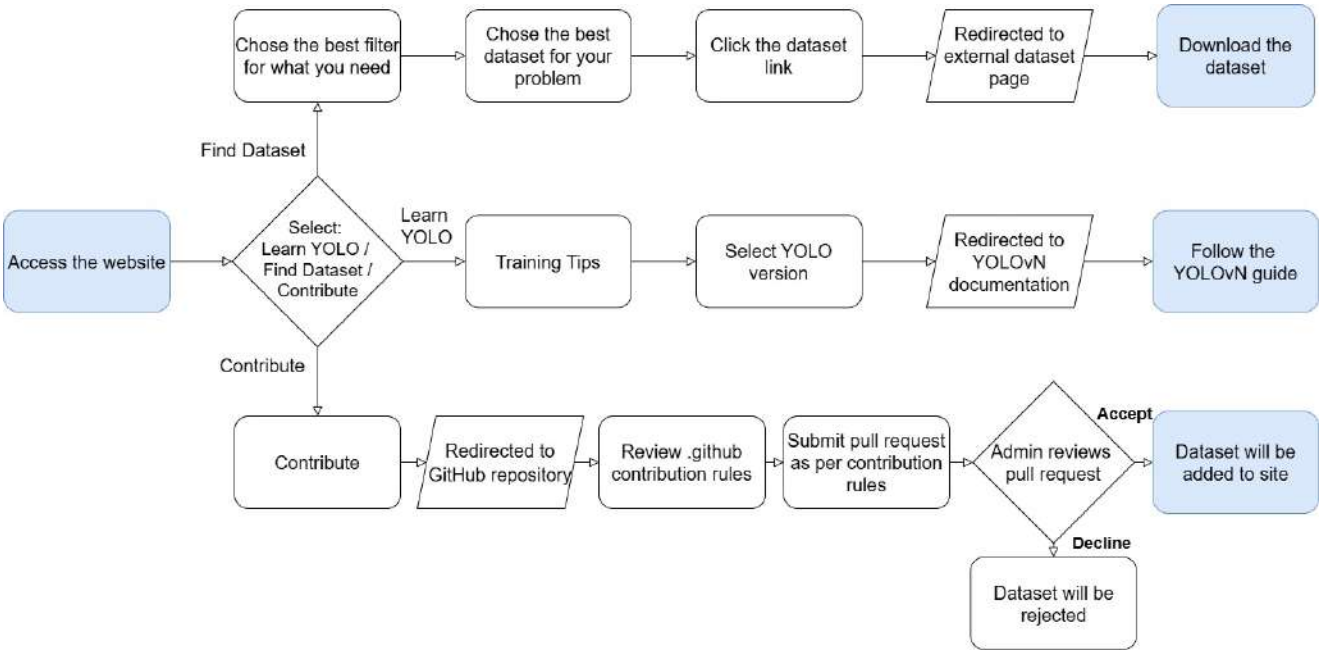


Fig. 1. Block diagram of the proposed solution

b.3. Path 3 – Contributing a Dataset

- **Contribute:** Users interested in contributing datasets can select the "Contribute" option on the homepage.
- **Redirected to GitHub repository:** This action redirects the user to the project's GitHub repository.
- **Review `.github` contribution rules:** Contributors are instructed to read the contribution guidelines available in the `.github` folder.
- **Submit pull request as per contribution rules:** A pull request is submitted including the dataset link and required metadata, following the project's standards.
- **Admin reviews pull request:** Before the manual review, an automated pipeline using GitHub Actions performs an initial validation to ensure the contribution follows the required format. Then, a project administrator evaluates the submission to verify its overall validity and structure.
- **Decision – Accept or Decline:**
  - **Accept:** If approved, the dataset is merged and the website is automatically updated via GitHub Pages.
  - **Decline:** If the contribution does not meet the requirements, the pull request is rejected and the dataset is not included.

IV. Results

A. Proposed Platform

As illustrated in Figure 2, the platform's interface was designed with a strong focus on user-friendliness. To find a dataset, users simply access the platform, use the filter options above the dataset table to select a label that matches their specific needs, and then browse the filtered results. From the listed options, the user can choose the dataset that best addresses their use case or the requirements of their client, making the retrieval process practical and intuitive.

For contributors, the process is equally accessible: by clicking the "Contribute" button, users are redirected to the project's GitHub repository. There, following the contribution guidelines available in the .github folder, contributors can open a pull request and insert the dataset link into the collaborative .txt database included in the repository.

Additionally, the platform includes a "Training Tips" section located in the upper-right corner. This feature offers a curated list of resources and official documentation for YOLO, supporting users, especially those unfamiliar with model training in preparing their own computer vision models effectively.

B. Usability Research

To evaluate the usability and clarity of the proposed platform, we conducted a small-scale user study involving 10 undergraduate students from Inatel. Participants were asked to complete a set of tasks and respond to a short questionnaire aimed at assessing the accessibility, clarity, and perceived value of the application.

The first question investigated how frequently the participants use GitHub. Results revealed that 70% use GitHub regularly, 20% frequently, and 10% rarely, suggesting that most of the target audience is familiar with the platform. The second question asked whether the tutorial provided for the contribution process was clear. All participants answered affirmatively.

When asked about the level of difficulty in following the contribution steps, 90% rated the process as "easy" and 10% as "medium," as depicted in Figure 3. The result reinforced the intuitive design of the platform. Additionally, all participants reported encountering no technical issues when accessing or interacting with the available datasets.

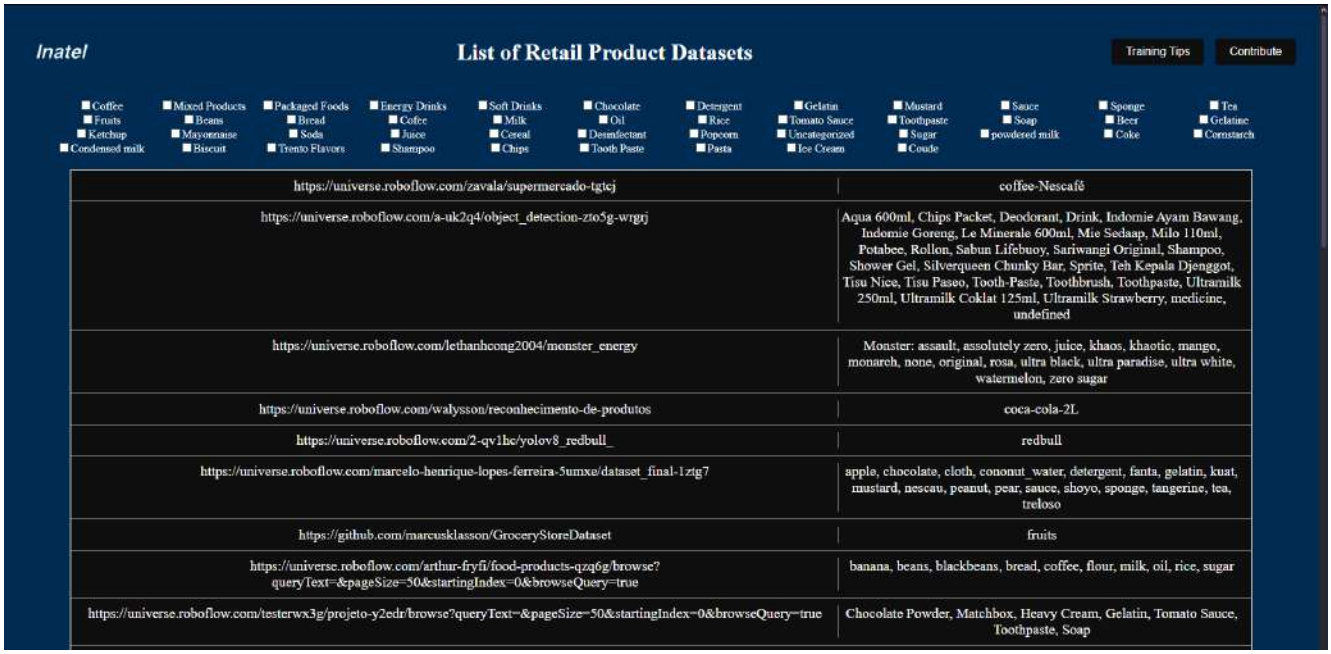


Fig. 2. The application's website: Retail-Product-Datasets

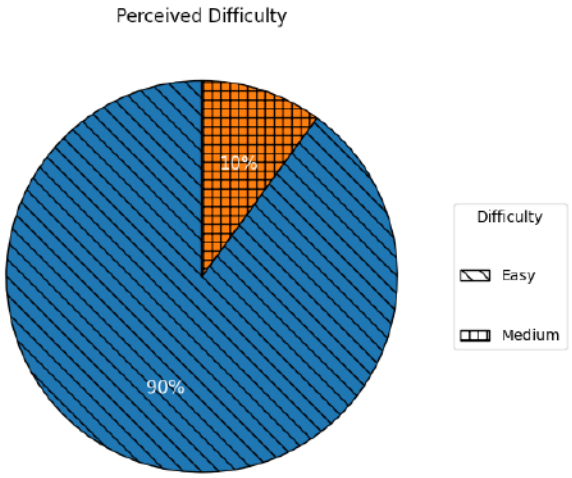


Fig. 3. Collaborate difficulty

A key question asked whether the platform could effectively contribute to retail establishments seeking to train new computer vision models. All participants answered "yes" and provided insightful justifications, such as:

- The platform facilitates access to pre-organized and labeled datasets, saving time during data preparation.
- It addresses a major challenge in computer vision model development: the availability of good and sufficiently large datasets.
- The system organizes datasets by product categories, making it easier to find suitable training data for retail use cases.
- It enables faster prototyping and improves the accuracy of real-world product recognition systems.

Finally, when asked for suggestions to improve the platform, most participants stated that the process was already clear and objective. A few suggested that a short video tutorial might help users unfamiliar with GitHub. Regarding additional features, no new functionality was requested, which indicates that the current scope of the platform was considered sufficient by the majority of users.

Overall, the feedback supports the effectiveness of the tool in facilitating collaboration and improving dataset accessibility for retail-focused computer vision applications.

V. Conclusions

The development of computer vision applications relies fundamentally on the availability of high-quality, well-annotated datasets. No matter how advanced a model architecture or training technique may be, a model cannot achieve meaningful results without data that accurately represents the target domain.

Datasets are what enable models to learn, generalize, and perform effectively in real-world scenarios. However, acquiring or creating suitable datasets remains one of the greatest challenges for practitioners, particularly in specialized areas such as retail. Public repositories often provide general-purpose data that, although useful, seldom meets the specific requirements of complex applications. This highlights the critical need for solutions that facilitate the curation of domain-specific datasets, capable of supporting the development of reliable computer vision models.

In summary, the purpose of this project is to address one of the main challenges in developing computer vision applications for retail: the difficulty of finding or creating domain-specific datasets. While public repositories and popular dataset

platforms often provide general-purpose data, they rarely offer datasets that match the specific requirements of retail environments, such as labeled images of packaged goods as seen on store shelves.

The proposed platform simplifies both the search for relevant datasets and the contribution of new ones by providing an accessible interface, filtering tools, and clear contribution workflows. By supporting the creation and sharing of specialized datasets, the project aims to facilitate the development of computer vision models that can be applied to real-world retail use cases, such as product recognition and inventory management.

In this project, we achieved the goal of developing a user-friendly, free, and collaborative platform for storing retail datasets. We implemented this solution using a popular platform, GitHub, which allowed us to create an easily accessible environment where users can search for datasets using filters, learn how to apply them in computer vision, and contribute by adding new datasets. Each contributed dataset is verified by an automated GitHub Actions workflow and reviewed by an administrator before being added to the platform.

The development of the proposed solution involved several technical challenges, such as:

- Identifying a free and collaborative platform that facilitates the storage and sharing of datasets.
- Designing a contribution process that enables users to collaborate easily while automating as much of the validation and integration as possible.
- Creating a structure that helps users clearly understand the purpose of each dataset and efficiently filter datasets using general tags.

Future improvements for the proposed platform could include:

- Implementing a GitHub Actions workflow capable of automatically verifying whether a contributed dataset link already exists in the database, in order to prevent duplicates and improve data integrity.
- Enhancing the user interface by applying a more refined design to the website and improving the structure of dataset tables, which are currently built with simple `div` elements due to the project's primary focus on functionality rather than aesthetics.
- Refining the tagging system to ensure a consistent and standardized set of tags for filtering and identifying datasets.
- Registering a custom domain for the platform to make access easier and improve its professional appearance. This could be done by purchasing a domain from a provider such as Google Domains, Namecheap, or GoDaddy and linking it to the GitHub Pages hosting the application.

References

[1] Suman Paneru and Idris Jeelani. "Computer vision applications in construction: Current state, opportunities & challenges". en. In: *Autom. Constr.* 132.103940 (Dec. 2021), p. 103940.

[2] Erkan Demir, Kazim Yildiz, Onder Demir, and Eyup Emre Ulku. "Computer Vision Based Intelligent 3D CNC Machines". In: *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*. Istanbul, Turkey: IEEE, Oct. 2020.

[3] Florian Scheidegger, Roxana Istrate, Giovanni Mariani, Luca Benini, Costas Bekas, and Cristiano Malossi. "Efficient image dataset classification difficulty estimation for predicting deep-learning accuracy". en. In: *Vis. Comput.* 37.6 (June 2021), pp. 1593–1610.

[4] Shreya M Shelke, Indrayani S Pathak, Aniket P Sangai, Dipali V Lunge, Kalyani A Shahale, and Harsha R Vyawahare. "A review paper on computer vision". en. In: *International Journal of Advanced Research in Science, Communication and Technology* 8901 (Mar. 2023), pp. 673–677.

[5] Juan Terven, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS". en. In: *Mach. Learn. Knowl. Extr.* 5.4 (Nov. 2023), pp. 1680–1716.

[6] Ciprian Orhei, Silviu Vert, Muguras Mocofan, and Radu Vasiu. "End-to-End Computer Vision Framework: An open-source platform for research and education". en. In: *Sensors (Basel)* 21.11 (May 2021), p. 3691.

[7] Floriana Ciaglia, Francesco Saverio Zuppichini, Paul Guerrie, Mark McQuade, and Jacob Solawetz. "Roboflow 100: A Rich, Multi-Domain Object Detection Benchmark". In: (2022).

[8] UCI Machine Learning Repository. *Online Retail Dataset*. 2025. URL: https://archive.ics.uci.edu/dataset/352/online+retail (visited on 04/16/2025).

[9] Kaggle. *Find Open Datasets and Machine Learning Projects*. 2025. URL: https://www.kaggle.com/datasets (visited on 04/16/2025).

[10] Google. *Dataset Search*. 2025. URL: https://datasetsearch.research.google.com/ (visited on 04/16/2025).

[11] Laura Dabbish, Colleen Stuart, Jason Tsay, and James Herbsleb. "Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository". In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW)* 1 (2012), pp. 1277–1286.

[12] Eirini Kalliamvakou, Georgios Gousios, Kelly Blincoe, Leif Singer, Daniel M. German, and Daniela Damian. “An In-depth Study of the Promises and Perils of Mining GitHub”. In: *Empirical Software Engineering* 21.5 (2016), pp. 2035–2071.

[13] Yue Yu, Gang Yin, Tao Wang, Cheng Yang, and Huaimin Wang. “Determinants of Pull-based Development in the Context of Continuous Integration”. In: *Science China Information Sciences* 59.8 (2016), pp. 1–19.

[14] Zhifang Liao, D. He, Z. Chen, X. Fan, Yan Zhang, and Shanshan Liu. “Exploring the Characteristics of Issue-related Behaviors in GitHub Using Visualization Techniques”. In: *IEEE Access* 6 (2018), pp. 24003–24015.

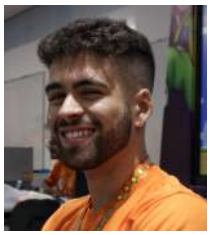
[15] Mairieli Wessel, Joseph Vargovich, Marco A. Gerosa, and Christoph Treude. “GitHub Actions: The Impact on the Pull Request Process”. In: *Empirical Software Engineering* 28.131 (2023).

[16] Jadson Santos, Daniel Alencar da Costa, Shane McIntosh, and Uirá Kulesza. “On the Need to Monitor Continuous Integration Practices”. In: *arXiv preprint arXiv:2409.05101* (2024).

[17] Beiqi Zhang, Peng Liang, Xiyu Zhou, Aakash Ahmad, and Muhammad Waseem. “Practices and Challenges of Using GitHub Copilot: An Empirical Study”. In: *International Conference on Software Engineering and Knowledge Engineering (SEKE)* (2023).

[18] Prayudi Utomo and Falahah. “Building Serverless Website on GitHub Pages”. In: *IOP Conf. Ser. Mater. Sci. Eng.* 879.1 (July 2020), p. 012077.

AUTHORS



Marcelo Henrique de Oliveira Pereira Vilela is a Computer Engineering student at the National Institute of Telecommunications (Inatel), currently in his tenth semester in the first semester of 2026. Throughout his academic career, he has participated in several scientific initiation projects in the areas of Artificial Intelligence and Robotics, resulting in the publication of scientific articles related to these topics. Currently, he works at Von Braun Labs, where he initially performed activities in the area of Quality Assurance (QA), conducting software testing and system validation. He currently works as a backend and artificial intelligence developer, participating in the development of solutions based on data processing, machine learning, and distributed systems.



Renzo Paranaíba Mesquita is a Computer Engineer (INATEL), Specialist and Master in Computing (UNIFEI) and Ph. D. in Production Engineering (UNIFEI). He works as a professor and coordinator of the Computer Engineering and Software Engineering courses at the National Institute of Telecommunications (INATEL). He is interested in the areas of Software Development, Data Analysis and Computational Simulation.

Desenvolvimento de uma plataforma para otimização da orientação psicopedagógica realizada pelo NOE e Mind++ Lab do Instituto Nacional de Telecomunicações - Inatel

Guilherme Augusto Barucke Marcondes, Ludimila Rocha da Silveira,
Tiago Augusto Costa Carvalho & Thomas Ceccon Guimarães
Instituto Nacional de Telecomunicações - Inatel
guilherme@inatel.br, ludimila@inatel.br, tiago.augusto@gec.inatel.br, thomas.ceccon@gec.inatel.br

***Abstract*—This paper presents the development of an web platform designed to optimize psychopedagogical guidance offered by the Educational Guidance Center (NOE) and the Mind++ Lab at the National Institute of Telecommunications (Inatel). This initiative aims to enhance support for students' academic and professional development through digital tools focused on behavioral training and soft skills development. The platform was architected using PHP for the back-end and HTML and CSS for interface structuring and styling, adhering to Inatel's technological standards to ensure compatibility and integration with the institutional infrastructure. The development environment included Visual Studio Code, GitHub for version control, and XAMPP for local testing. The application provides informative content, downloadable materials, and academic support resources covering self-knowledge, study planning, time management, and performance improvement. This platform is expected to strengthen psychopedagogical guidance, significantly contributing to students' academic and professional success, aligning with contemporary demands for technology-mediated education.**

***Index Terms*—psychopedagogical guidance, soft skills, web development, Higher education,educational platform.**

***Resumo*—Este artigo apresenta o desenvolvimento de uma plataforma Web, concebida para otimizar a orientação psicopedagógica oferecida pelo Núcleo de Orientação Educacional (NOE) e pelo Mind++ Lab do Instituto Nacional de Telecomunicações (Inatel). A iniciativa visa aprimorar o suporte ao desenvolvimento acadêmico e profissional dos estudantes, por meio de ferramentas digitais focadas no treinamento comportamental e no desenvolvimento de soft skills. A plataforma foi arquitetada utilizando PHP para o back-end e HTML e CSS para a estruturação e estilização da interface, seguindo os padrões tecnológicos do Inatel para garantir compatibilidade e integração com a infraestrutura institucional. O ambiente de desenvolvimento incluiu Visual Studio Code, GitHub para controle de versão e XAMPP para testes locais. A aplicação disponibiliza conteúdos informativos, materiais para download e recursos de apoio acadêmico que abordam autoconhecimento, planejamento de estudos, gestão do tempo e melhoria de desempenho. Espera-se que esta plataforma fortaleça a orientação psicopedagógica, contribuindo significativamente para o sucesso acadêmico e profissional dos estudantes, alinhando-se às demandas contemporâneas por educação mediada por tecnologia.**

***Palavras chave*—orientação psicopedagógica, soft skills, desenvolvimento web ,ensino superior,plataforma educacional.**

I. Introdução

O cenário educacional contemporâneo, especialmente no ensino superior, exige uma abordagem holística para o desenvolvimento do estudante, que transcenda a mera formação técnica. A aquisição de competências comportamentais e socioemocionais, amplamente conhecidas como soft skills, tornou-se um diferencial crítico para o sucesso acadêmico e a inserção no mercado de trabalho [1]. Habilidades como comunicação eficaz, gestão do tempo, organização pessoal, autoconhecimento e inteligência emocional são cada vez mais valorizadas e consideradas essenciais para a adaptabilidade e resiliência em ambientes dinâmicos [2].

No Instituto Nacional de Telecomunicações (Inatel), o Núcleo de Orientação Educacional (NOE), em colaboração com o Mind++ Lab, desempenha um papel fundamental no suporte psicopedagógico aos alunos. Este suporte abrange acompanhamento individual e coletivo, visando o desenvolvimento dessas competências, em consonância com diretrizes internacionais sobre habilidades socioemocionais [3]. A crescente demanda por esses serviços e a necessidade de ampliar o alcance das ações do Mind++ Lab evidenciam a urgência de implementar soluções tecnológicas que otimizem e qualifiquem esses processos, tornando-os mais acessíveis e eficientes [4].

Nesse contexto, surgiu a proposta de desenvolvimento de uma plataforma web dedicada à orientação psicopedagógica e psicológica. O objetivo principal é democratizar o acesso dos estudantes a conteúdos formativos, ferramentas de apoio acadêmico e recursos de treinamento comportamental. A plataforma permite que os alunos acessem materiais de forma autônoma, em qualquer local e horário, promovendo maior flexibilidade e autonomia em seu percurso de desenvolvimento pessoal e acadêmico. Tal abordagem alinha-se aos princípios de experiência do usuário e usabilidade, essenciais para a efetividade de ambientes digitais de aprendizagem [5]. A plataforma desenvolvida pode ser acessada em: https://inatel.br/mind-new/.

A escolha das tecnologias para o desenvolvimento da plataforma como: HTML e CSS para a estruturação e estilização da interface, e PHP para a implementação das funcionalidades de back-end, foi pautada pelas diretrizes institucionais do Inatel. Essa conformidade garante alinhamento com a infraestrutura existente e as boas práticas já consolidadas na instituição, facilitando a manutenção e a escalabilidade do sistema [6]. O

desenvolvimento seguiu princípios de engenharia de software, assegurando modularidade e robustez. O ambiente de desenvolvimento utilizou XAMPP como servidor local e GitHub para controle de versão, garantindo um processo organizado e colaborativo.

Este artigo detalha o processo de concepção, desenvolvimento e implementação da plataforma, apresentando suas funcionalidades e a metodologia adotada. Discute-se, ainda, como a solução contribui para o fortalecimento do suporte psicopedagógico no ensino superior, promovendo o desenvolvimento integral dos estudantes e aprimorando a comunicação institucional do Mind++ Lab. A plataforma representa um avanço significativo na digitalização e otimização dos serviços de apoio ao estudante, consolidando um ambiente virtual que atende às necessidades emergentes da educação superior.

II. Metodologia

O desenvolvimento da plataforma Web do Mind++ Lab adotou uma abordagem incremental e centrada no usuário, alinhada aos princípios da engenharia de software [5]. Esta metodologia permitiu a construção do sistema em etapas progressivas, facilitando a incorporação de *feedback* e a realização de ajustes contínuos, o que é crucial para sistemas educacionais que demandam alta usabilidade e relevância pedagógica [7].

A. Levantamento de requisitos e planejamento

A fase inicial consistiu no levantamento detalhado das demandas junto ao Mind++ Lab, por meio de reuniões e entrevistas com a equipe e potenciais usuários. As principais necessidades identificadas incluíram:

- **Centralização de informações**
  Criar um ambiente digital unificado para todos os recursos e serviços do Mind++ Lab.
- **Divulgação estruturada**
  Organizar e apresentar as ações coletivas (workshops, palestras) de forma clara e acessível.
- **Apresentação das frentes de atuação**
  Detalhar os serviços oferecidos pelo laboratório (plantão psicológico, orientação profissional, etc.).
- **Disponibilização de conteúdos de apoio:**
  Oferecer materiais educativos para autoconhecimento, planejamento de estudos e desenvolvimento de *soft skills*.

Com base nesses requisitos, foi elaborado o planejamento estrutural da aplicação, definindo o fluxo de navegação e a arquitetura da informação. A plataforma foi concebida com um menu principal intuitivo, composto pelas abas ”Home”, ”Eventos”e ”Sobre”, visando otimizar a clareza organizacional e a experiência do usuário, em conformidade com heurísticas de usabilidade [5].

B. Tecnologias empregadas

A escolha das tecnologias foi guiada pela compatibilidade com a infraestrutura existente do Inatel e pela robustez para o desenvolvimento Web:

HTML (*HyperText Markup Language*): Utilizado para a estruturação semântica do conteúdo das páginas, garantindo acessibilidade e organização.

CSS (*Cascading Style Sheets*): Empregado para a estilização visual, padronização da interface e responsividade, assegurando a adaptação a diferentes dispositivos e tamanhos de tela.

PHP (*Hypertext Preprocessor*): Linguagem de programação *server-side* utilizada para a organização dinâmica das páginas, integração de conteúdos e funcionalidades interativas.

XAMPP: Servidor local utilizado para o ambiente de desenvolvimento e testes, simulando o funcionamento completo da aplicação em um ambiente controlado.

Visual Studio Code: Ambiente de Desenvolvimento Integrado (*IDE*) para a codificação e organização do projeto.

GitHub: Plataforma para controle de versão e gerenciamento colaborativo do código-fonte, garantindo rastreabilidade e segurança no desenvolvimento.

C. Processo de Desenvolvimento e testes

O desenvolvimento seguiu um ciclo iterativo, com foco na modularidade e facilidade de manutenção. Foram realizados testes funcionais contínuos para validar a navegação, a integridade dos links e a correta exibição dos conteúdos. Além disso, verificações de compatibilidade e responsividade foram conduzidas para assegurar a adaptação adequada da interface em diversos dispositivos, um aspecto crucial para plataformas educacionais que buscam ampla acessibilidade [8]. A validação da usabilidade foi um processo contínuo, buscando garantir que a plataforma atendesse às necessidades dos usuários de forma organizada, acessível e funcional, promovendo uma experiência de aprendizado engajadora [9].

III. Resultados e Discussão

O desenvolvimento da plataforma web do Mind++ Lab culminou em uma aplicação institucional estruturada e organizada, que centraliza as informações do laboratório de forma clara e acessível. A implementação do menu principal com as abas ”Home”, ”Eventos”e ”Sobre”proporcionou uma divisão lógica do conteúdo, contribuindo significativamente para a melhoria da usabilidade e da comunicação institucional [10]. A organização por seções permite que o usuário encontre rapidamente as informações desejadas, reduzindo a complexidade da navegação e aumentando a eficiência da interação com a plataforma.

A aba ”Home”serve como página inicial e porta de entrada para o usuário, apresentando uma introdução institucional do Mind++ Lab, sua missão, objetivos e propósito educacional. Nesta seção, são destacadas as ações voltadas ao desenvolvimento de soft skills, como comunicação, inteligência emocional, organização acadêmica e planejamento de carreira, que são competências cada vez mais demandadas no mercado de trabalho [1]. A estrutura visual foi planejada para ser clara e objetiva, facilitando a compreensão da finalidade do laboratório. A integração com a plataforma Moodle, que oferece acesso a conteúdos complementares e materiais estruturados, fortalece o acompanhamento oferecido pelo laboratório, promovendo maior autonomia no processo de aprendizagem e desenvolvimento dos estudantes. Conforme apresentado na Figura 1.



Fig. 1. Página inicial da plataforma Mind++ Lab.

A aba ”Eventos”foi projetada para centralizar e divulgar todas as atividades promovidas pelo Mind++ Lab ao longo do semestre, incluindo workshops, palestras e encontros temáticos. A centralização dessas informações em um único ambiente digital facilita o acesso por parte dos alunos e contribui para um maior engajamento nas atividades propostas. Com base na Figura 2. Do ponto de vista organizacional, essa estrutura reduz a dispersão de informações e melhora a comunicação institucional, um fator crítico para a eficácia de programas de apoio estudantil [4].



Fig. 2. Página de divulgação de eventos do Mind++ Lab.

A aba ”Sobre”concentra as informações detalhadas do Mind++ Lab, incluindo sua missão, frentes de atuação, modalidades de atendimento e equipe responsável. Esta seção detalha o ”Plantão Mind++ Lab”, evidenciando os horários e formatos de atendimentos disponíveis, e ilustra as diversas frentes de atuação do laboratório, como Orientação Profissional e Orientação Psicológica e Psicopedagógica. A clareza e a abrangência das informações apresentadas nesta seção são cruciais para que os estudantes compreendam a totalidade do suporte oferecido, promovendo a utilização efetiva dos recursos disponíveis. A organização estrutural adotada contribui para uma navegação intuitiva, melhora a experiência do usuário e fortalece a comunicação institucional do Mind++ Lab, evidenciando a efetividade da solução desenvolvida na promoção do desenvolvimento integral dos estudantes.



Fig. 3. Seção de Plantão do Mind++ Lab.

IV. Conclusão

O desenvolvimento da plataforma web do Mind++ Lab representa um avanço significativo na organização e digitalização das informações e serviços oferecidos pelo laboratório.

A estruturação da aplicação, com as abas ”Home”, ”Eventos”e ”Sobre”, proporcionou uma navegação intuitiva e funcional, facilitando o acesso dos alunos aos serviços psicológicos e psicopedagógicos essenciais para seu desenvolvimento acadêmico e pessoal ao longo da graduação.

A unificação em um ambiente digital das diversas frentes de atuação do Mind++ Lab, que incluem plantão psicológico, orientação psicopedagógica, workshops temáticos, orientação profissional e adaptação acadêmica, ampliou a visibilidade das ações do setor e fortaleceu a comunicação com a comunidade acadêmica.

A disponibilização de materiais em formato PDF, ferramentas de organização acadêmica e o direcionamento à plataforma Moodle com conteúdo exclusivo contribuem diretamente para o desenvolvimento contínuo das competências socioemocionais e profissionais dos estudantes, alinhando-se às necessidades do mercado de trabalho e às demandas por uma formação integral [1, 2].

Do ponto de vista técnico, a escolha e a aplicação de HTML, CSS e PHP permitiram a construção de uma aplicação estruturada, organizada e de fácil manutenção. O uso de ferramentas como XAMPP para testes locais e GitHub para versionamento garantiu maior controle e segurança durante todo o processo de desenvolvimento. A adesão aos padrões tecnológicos institucionais do Inatel assegurou a compatibilidade com a infraestrutura existente, facilitando a integração e a sustentabilidade da plataforma a longo prazo.

Em suma, esta versão da plataforma consolidou uma ferramenta funcional, acessível e alinhada às necessidades do Mind++ Lab, com o potencial de otimizar significativamente a orientação psicopedagógica no ensino superior e de promover o desenvolvimento integral dos estudantes do Inatel.

Agradecimentos

Gostaríamos de expressar nossa gratidão a Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMING, ao Instituto Nacional de Telecomunicações - INATEL, e ao Mind++ Lab pelo apoio técnico e recursos que viabilizaram a continuidade e aprimoramento deste projeto.

Referências

[1] WORLD ECONOMIC FORUM. *The Future of Jobs Report 2020*. Rel. técn. World Economic Forum, 2020. URL: https://www3.weforum.org/docs/WEFFutureofJobs2020.pdf.

[2] CENTRO UNIVERSITÁRIO CENTRAL PAULISTA (UNICEP). "Desenvolvimento de Soft Skills na Universidade: Como os alunos podem se preparar?" Em: *UNICEP* (2025). URL: https://www.unicep.edu.br/post/desenvolvimento-de-soft-skills-na-universidade.

[3] OECD. *Skills for Social Progress: The Power of Social and Emotional Skills*. Paris: OECD Publishing, 2015. URL: https://www.oecd.org/education/skills-for-social-progress-9789264226159-en.htm.

[4] Jakob NIELSEN. "Ten Usability Heuristics for User Interface Design". Em: *Nielsen Norman Group* (1994). URL: https://www.nngroup.com/articles/ten-usability-heuristics/.

[5] Roger S. PRESSMAN e Bruce R. MAXIM. *Software Engineering: A Practitioner's Approach*. 8ª ed. New York: McGraw-Hill Education, 2016.

[6] M. R. et al. LIMA. "O uso das tecnologias virtuais para mediação pedagógica e desenvolvimento de Hard e Soft Skills em cursos da área da saúde: relato de experiência". Em: *Tecnologias Educacionais* (2020). URL: https://downloads.editoracientifica.com.br/articles/200700767.pdf.

[7] A. LOURENCO, J. S. NAVARRO-LOLI e S. DOMINGUEZ-LARA. "Technological Innovation in Engineering Education: A Psychopedagogical Approach for Sustainable Development". Em: *Sustainability* 17.14 (2025), p. 6429. URL: https://www.mdpi.com/2071-1050/17/14/6429.

[8] F. Q. et al. SIQUEIRA. "Unindo Forças: Estratégias Psicopedagógicas para a Adaptação e Permanência no Ensino Superior". Em: *CONEDU*. 2025. URL: https://editorarealize.com.br/artigo/visualizar/111307.

[9] A. M. et al. ARRUDA. "O uso de metodologias ativas na educação e formação em saúde: implicações para o desenvolvimento de hard e soft skills". Em: *Revista DCS* (2026). URL: https://revista.faculdadedcs.com.br/index.php/revistadcs/article/view/226.

[10] T. E. C. SANTOS e P. S. TEIXEIRA. "A atuação da Psicopedagogia em Instituições de Nível Superior". Em: *Caderno Pedagógico* (2025). URL: https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/3281.

Autores



Guilherme Augusto Barucke Marcondes Engenheiro Elétrico com Ênfase Em Eletrônica e Telecom, Mestre em Telecomunicações pelo Instituto Nacional de Telecomunicações, e Doutorado em Engenharia de Produção pela Universidade Federal de Itajubá. Especialista em administração pela FEA/USP, formação como conselheiro pelo IEL/MG (2025) e é certificado como PMP pelo PMI. Professor titular, Pró-Diretor de Graduação, Vice-Diretor e membro do Conselho Diretor do INATEL. Revisor de itens do INEP/MEC para ENADE (Engenharia de Computação). Foi membro do Conselho Diretor da Fundação Instituto Nacional de Telecomunicações de 2005 a 2018. Foi Coordenador dos Cursos de Engenharia de Computação (2010 a 2022) e de Engenharia de Software (2018-2022) do Inatel. Atuou como professor convidado do mestrado profissional da UnB em Telecomunicações. Experiência de mais de 20 anos em projetos de PD na área de tecnologia da informação e nas áreas de gerência de projetos e de negócios. Experiência na Gestão de Projetos Internacionais.



Ludimila Rocha da Silveira
Psicóloga especialista em Educação Corporativa. Atua com o desenvolvimento de pessoas e treinamento comportamental para performance profissional. Coordenadora do Núcleo de Orientação Educacional e do Mind++ Lab, o laboratório de treinamento comportamental e aprimoramento das habilidades profissionais no principal Centro de Excelência em Ensino e Pesquisa na área de Engenharias - o Instituto Nacional de Telecomunicações - INATEL.



Tiago Augusto Costa Carvalho
Graduando em Engenharia da Computação (10º período) pelo Instituto Nacional de Telecomunicações (INATEL);
Técnico em Eletrônica com ênfase em Automação; Atuante em projetos de desenvolvimento de software, sistemas embarcados e automação.



Thomas Ceccon Guimarães
Graduando no 10° período de Engenharia da Computação no Instituto Nacional de Telecomunicações (INATEL).

Checklist-Based Code Review Process: An Experience Report

Pedro Alyson T. dos Santos, Samuel Sousa Teles, Rubens Abraão da Silva Sousa,
Manoel Filho, Alex Lacerda Ramos
Instituto Federal de Ciência e Tecnologia do Ceará (IFCE)
pedro.alyson09@aluno.ifce.edu.br, samuel.teles@darmlabs.ifce.edu.br, rubens.silva@darmlabs.ifce.edu.br,
manoel.filho@darmlabs.ifce.edu.br, alex.lacerda@ifce.edu.br

***Abstract*—In the context of a real RD&I project, code review often suffers from a lack of clear criteria, resulting in superficial analysis and rework. The motivation for this study was to mitigate these problems through a structured checklist to standardize validation and increase software quality. The method adopted was an experience report (2024-2025) with a mixed approach, analyzing quantitative metrics from SonarQube and qualitative feedback from eight developers. The results showed significant reductions of 60.1% in the average number of bugs and 80.44% in code smells, as well as continuous improvement in test coverage and reduction of duplicate code. It is concluded that the checklist is effective in structuring the process and reducing defects, serving as a learning tool, as long as it is kept in continuous evolution and combined with human critical thinking.**

***Index Terms*—Code Review, Checklist, Software Quality, Experience Report**

I. Introduction

In software development, code review is a fundamental practice for reducing and preventing problems such as deviations from coding standards, architectural flaws, maintenance difficulties, and inconsistencies between implementation and requirements. Recent studies indicate that well-structured code reviews significantly improve code quality, readability, and maintainability [1].

However, challenges persist when there are no clear criteria, efficient communication between team members, or well-defined objectives for the review. This lack of alignment can lead to inconsistent reviews that focus on superficial aspects, such as coding style, at the expense of critical issues, such as functional defects or architecture violations. In addition, fragmented communication can result in rework, delays, and conflicts among team members, especially in agile or distributed environments, reducing the strategic impact of the review on the development cycle [2, 3].

Poorly structured reviews can have significant consequences, including increased rework, longer delivery times, and the introduction of flaws into production, compromising the reliability and maintainability of the system over time. Recent research demonstrates that the adoption of structured review practices, including checklists, allows reviewers to detect defects and code smells more effectively, reducing effort and rework [4]. In addition, structured checklists increase the efficiency of the review process by helping reviewers focus on critical aspects [5].

Given this, there is a need for a mechanism that assists developers during code review, promoting uniformity, traceability, and continuous evolution of review criteria. In this context, the adoption of a structured checklist proves to be a viable alternative. When incorporated into the workflow, the checklist can become a requirement for the approval of changes to the code, contributing to the continuous maturation of the process and the improvement of software quality.

This paper presents an experience report on a real project linked to a Research, Development, and Innovation (RD&I) laboratory. The study investigates how the implementation of a checklist in the code review process impacted software quality, using a mixed approach:

- **Quantitative analysis:** collection of quantitative data using metrics for bugs, code smells, code duplication, and test coverage obtained from SonarQube[1]. The analysis period covered June 2024 to June 2025, ensuring consistency and robustness in the assessment.
- **Qualitative analysis:** based on developers' perceptions, allowing an understanding of the team's practical experience and complementing the interpretation of quantitative data.

The study is presented in the following sections: Introduction, which presents the context and objective of this work; Related Works, discussing previous studies that highlight contributions, advances, and gaps underpinning the relevance of this work; Method, detailing the step-by-step process used to construct the research; Results, presenting the data obtained and its implications; and Conclusion, summarizing the contributions of the work, highlighting limitations, and suggesting future research directions.

II. Related Works

[6] demonstrate through a case study the relationship between coverage, participation, and expertise in code review and long-term software quality, based on post-release defects. The results show that modern practices share a significant link with software quality. However, the study presents data obtained from open source and domain-specific projects, which may reduce the diversity of the study.

In another approach, [7] presents research on the effect of the review rate on the efficiency of defect removal in software quality, using the Personal Software Process (PSP) approach. The results prove that the PSP review rate is a significant factor affecting the effectiveness of defect removal. However, the

[1]https://www.sonarsource.com/products/sonarqube/

study is developed in an educational environment and may not reflect real professional development practices and pressures.

More recently, [8] demonstrates that explicit code review strategies, such as checklists and guided lists, improve effectiveness and efficiency and reduce the cognitive load on reviewers. However, the study does not consider the effects of learning or adaptation by reviewers over time.

Furthermore, [9] experimentally investigated the effectiveness of checklist-based reading (CBR) compared to ad hoc reading in code inspections. The results of the experiment indicate that CBR is significantly superior to ad hoc reading in terms of effectiveness, efficiency, effort, and number of false positives. However, the study presents the analysis in printed code, which does not represent the real context of modern code review tools.

[10] propose a new taxonomy aimed at categorizing design discussions that emerge during the code review process. Through an empirical study focusing on TypeScript projects, the authors investigate how the qualitative content of reviews is related to future software maintenance. Although the study does not directly address the use of checklists, it provides a solid basis for understanding the types of feedback that can be standardized or prioritized through explicit strategies, such as checklists. This perspective reinforces the importance of tools that aid communication and consistency in reviewing structural aspects of code.

[11] presents a method to support code reviewers in the task of understanding and evaluating changes submitted in code reviews, classifying files in a patch that need specific attention. The study confirmed that the use of modern text embedding techniques can improve the prediction of which files need attention and optimize the effort of reviewers. However, although reordering files increases comments and recovery of critical points, the real impact on the final quality of the software is not fully evaluated.

In summary, despite the advances presented by these studies, there is a lack of reports analyzing the use of checklists in real environments, with evaluation by multiple code quality metrics. Thus, this work proposes to explore the implementation of a checklist-based code review process through a report on experience in a real development context, with the aim of contributing to the standardization and improvement of software quality.

III. Method

The study was conducted in a single project, with the same team of eight developers (four seniors and four juniors) throughout the period (June 2024 to June 2025), with no significant changes in the team or processes. Thus, the main methodological change observed was the introduction of the checklist integrated with SonarQube and Jenkins[2].

In addition, the study presents a collection of quantitative data, through the extraction of metrics of identified bugs, code smell, code duplication, and test coverage in SonarQube, during the aforementioned period. In addition to the metrics extracted from SonarQube, records from the Github[3] tracking system (pull requests and issues) were analyzed, allowing for consistent comparisons between the period before and after the adoption of the checklist.

The experiment aims to reduce problems and increase maintenance in the source code, i.e., code smell, undue duplication, bugs, and increased test coverage.

A. Definition of the Checklist

The items were defined based on recurring problems identified in the previous process, such as lack of test coverage, standardization failures, low maintainability, and duplication of database migrations. Based on these critical points, the team systematized minimum criteria for objective validation in code review.

Thus, the checklist was separated into topics that were problematic for the developed code, namely: the architecture and code writing standard, which would be to maintain the structure and integrity of the development process, from implementation in the code to the opening of the PR; unit tests, verifying that the tests cover the solution in the code, that they have the correct reference to the task, and that they cover the proposed problem acceptance criteria; the implemented solution, which would be the verification to ensure quality and functionality in the resolution; in the case of the application back-end, there are database migration criteria, with the correct version numbering order and no fixed IDs in these files; Maintainability, which would be the requirements to keep the code reusable and modular, making maintenance easy; and another serious issue is approval in the Jenkins pipeline, which would be a way to run the application and tests to ensure total integrity and not break the main branch.

With this in mind, the checklist used for the experiment was designed and its items separated as follows:

1) **Code Standardization and Writing**
   a) The commit message follows the standard established by the project.
   b) There is no commented code without justification.
   c) There are no debug logs in the code.
   d) There is no use of $\star$ in imports.
   e) There is no duplicate code.
2) **Unit Tests**
   a) The code is completely covered by unit tests.
   b) The tests refer to the corresponding task.
   c) The tests ensure coverage of the task acceptance criteria.
3) **Implemented Solution**
   a) The solution is of high quality and considered efficient.
   b) The code does not compromise system performance.
   c) The code maintains proper integration with other systems.
   d) The code is protected against failures and exceptions.
   e) The code does not negatively interfere with existing functionalities.
   f) The code meets security requirements.
4) **Database Migrations**
   a) Migrations follow the correct numbering order.

[2]https://www.jenkins.io//
[3]https://github.com/

b) There are no fixed IDs in migrations.

5) **Maintainability**
   a) The code is modular and reusable.
   b) The implementation is located in the appropriate class.
6) **Others**
   a) The pipeline in Jenkins was successfully approved.

B. Execution

Next, Figure 1 shows the process of creating a *Pull Request* (PR) used in performing code review.



Fig. 1. Pull Request (PR) creation flow.

As shown in Figure 1, the code review begins with the creation of the PR after the solution development is complete, then assigns two reviewers to perform the code checks, and immediately communicates with them. Next, Figure 2 shows the code review execution process with the help of the checklist.



Fig. 2. Code Review Workflow with Checklist.

As shown in Figure 2, each reviewer requested in the PR (Pull Request) must verify each modification made by the developer and, at the end, check the items on the checklist that can be considered in relation to the code. On the other hand, if not all items on the list have been checked, the reviewer requests modifications from the owner of the pull request, so that they can only approve when all items on the checklist have been checked. In addition to manual validation by reviewers, the process was integrated with Jenkins, which automated test execution and ensured the integrity of the main branch. This combination reduced the need for manual functional testing and increased the reliability of the process.

IV. Results

The period analyzed covers data from the front-end application of a real project linked to a Research, Development, and Innovation (RD&I) laboratory, collected through SonarQube between June 2024 and June 2025. The checklist was implemented in the code review process starting in September 2024. Next, in the section on improvements observed, Figure 3 shows the total number of bugs extracted during this period.

A. Quantitative Analysis



Fig. 3. Total Bugs.

According to Figure 3, from June to October 2024, there was a fluctuating trend in the number of issues classified as "bugs" by SonarQube, a category used by the tool to represent potential defects identified through static analysis, but with little deviation from the average of 261 occurrences per month. There was also a significant reduction in these occurrences, from 258 in October to 104 in the following months, representing an improvement of approximately 60.1% compared to the average for the period prior to the application of the checklist. These data suggest that the use of the checklist in the code review process may have contributed to the prevention of potential problems arising from incorrect or poorly designed implementations.. Next, Figure 4 shows the total number of lines considered code smell in the period from June 2024 to June 2025.

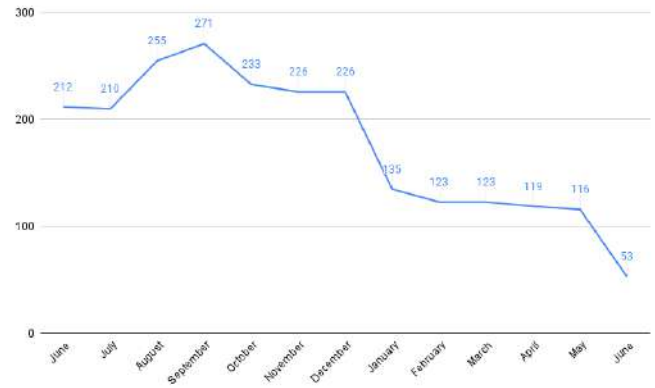


Fig. 4. Total Lines Considered Code Smell.

Figure 4 shows that there was a reduction in the number of poorly designed lines of code. In September, the highest number was approximately 271 lines, while in December, this number fell to around 226 lines. This represents an improvement of approximately 16.61%. In addition, in January and June, there was a significant reduction, with the numbers falling to 135 in January and 53 lines in June, representing a final reduction of 80.44% of lines considered code smell compared to September, when the checklist was first used. This clearly shows that the use of the checklist, in addition

to preventing code smell, may have influenced the progressive decrease, consequently representing a relative increase in software quality. Next, Figure 5 shows the total number of duplicated lines in the period from June 2024 to June 2025.

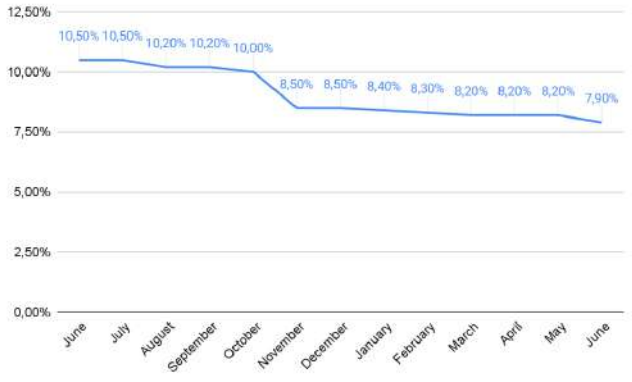


Fig. 5. Total Duplicate Lines.

According to Figure 5, there was a decrease in code duplication, with an average of 10.30% of duplicate lines between June and October. Soon after, there was a 1.8% drop in November compared to previous months. In addition, June showed 7.90% of duplicate lines, representing a 2.60% decrease compared to the months with the highest percentage of duplicate lines, which were June and July, with 10.5% in both. Thus, with the use of the checklist that highlights the requirement to eliminate code duplication for code review approval, a gradual decrease can be seen. Next, Figure 6 shows the percentage of test coverage from October 2024 to June 2025. It should be noted that values prior to October 2024 were not included in SonarQube, which is why they were not considered in the figure or in the data analysis.

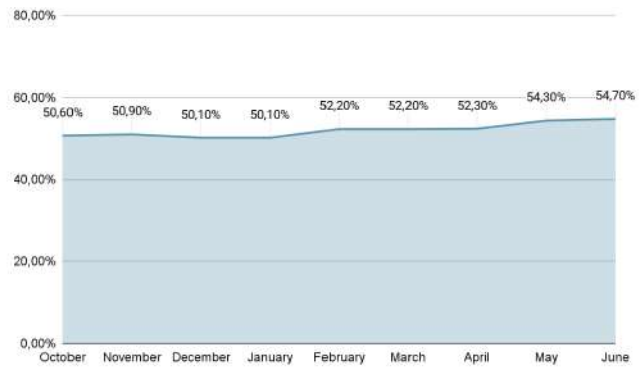


Fig. 6. Total Lines with Test Coverage.

Figure 6 shows a continuous improvement in unit test coverage over the period analyzed, starting at 50.6% in October 2024 and reaching an average of 54.7% by the end of June 2025. This steady evolution indicates that the total number of lines of code with test coverage increased in the months considered. This result shows that the project's coding is aligned with the adopted checklist, which recommends mandatory testing for each function developed.

B. Qualitative Analysis

In order to add value to the development team's perspectives on the code review process, using the checklist as a support for code review, some developers were asked to express their opinions at the end of the experiment. The following is the opinion of a member of the development team:

> *"The downside of using the checklist in the review is the use of a fixed scope. If a problem is not listed, then it will be passed as approved. Therefore, in my opinion, there is a flaw."*

From this perspective, it is necessary to constantly review the items on the checklist. It is essential to hold regular meetings to discuss possible improvements, expanding its scope. This ensures greater security for the reviewer and reduces errors by incorporating new items that cover previously unforeseen scenarios.

Thus, it is observed that the checklist requires periodic meetings to review items in order to receive feedback from reviewers and other developers, keeping it constantly aligned with the needs of the project and, thus, always ensuring quality in the code review process.

On the other hand, the checklist can also be used as a reminder for essential checks, assisting both the reviewer and the developer during the process. Below is a quote from another developer:

> *"One of the ways I use the checklist is as a reminder. That way, I don't lose my train of thought."*

As a result, reviewers often get lost during code reviews when their objectives are scattered and decentralized. However, with the use of a checklist, the objectives become centralized and clear, with each item explaining what is expected of the code so that it is best suited to the project. This way, reviewers can feel more confident when performing the process because, even though they have the freedom to critically analyze the code, the standards established in the project they are involved in ensure that they do not lose their train of thought, since the checklist helps direct their attention to critical points.

In addition, when used correctly, the checklist can contribute to reducing review time. This is because, when implementing a solution that is already aligned with the checklist criteria, requests for changes to the code are reduced. This view is reinforced by another quote from a developer:

> *"When I use the checklist for programming, I am able to develop with the requirements for approval in mind."*

The use of the checklist not only guides the review, but also serves as a guide during development, allowing the programmer to anticipate potential problems and align their implementation with the approval criteria. Thus, when developers integrate the checklist into their routine, they are able to produce code that better meets project expectations, reducing future revisions and increasing the efficiency of the development process.

However, in more complex situations, the use of the checklist alone may not be sufficient. In such cases, a more in-depth review may be necessary. One team member raises the following question:

"But can the checklist be used when a more thorough check is required? For example, if the implementation requires greater attention to detail at the external level, is this situation covered by any item on the checklist?"

This observation reinforces the idea that the checklist should be seen as a support tool, not a substitute for critical analysis. In more complex contexts, it can be supplemented by additional reviews that consider nuances and specificities that go beyond the previously established criteria.

This observation reinforces the idea that the checklist should be seen as a support tool, not a substitute for critical analysis. While the checklist organizes the review process and reminds reviewers of essential checks, it cannot cover all possible situations that may arise during development. In more complex contexts, such as integrations with external systems, performance checks, or specific project requirements, simply following the checklist may not be sufficient.

In such cases, it should be supplemented by additional reviews, using the reviewer's technical judgment to identify nuances and specificities that go beyond the previously established criteria. In addition, the combined use of the checklist with critical analysis contributes to code standardization, increases process efficiency, and reduces the need for rework, while strengthening the reviewer's ability to deal with unforeseen situations. To further improve its effectiveness, the checklist can be updated periodically based on team feedback, making it a flexible tool that can be adapted to project changes and demands.

V. Conclusion

The experience of incorporating a checklist into code reviews proved to be not only a technical control practice, but also a tool for learning and maturing the development team. However, the results show limitations that were considered: the checklist does not replace the reviewer's critical sense, and the effectiveness of the checklist is related to a specific context, and its results may vary in other scenarios.

The results reveal a considerable drop in bugs in the analyzed period from 266 to 104 (a 60.9% drop), a reduction in code smells from 271 to 53 lines (an 80.44% improvement), a reduction in code duplication from 10.5% to 7.9%, and an increase in test coverage from 50.6% to 54.7%. These data reveal that the checklist helped structure and qualify the review process, raising code quality and promoting practical and measurable improvements.

In addition, feedback from developers revealed that there was perceived value in using the checklist in practice, but that it is not static and therefore requires continuous review to keep pace with the evolution of the project and best practices.

The lessons learned highlight the value of the checklist as a support for developers, the need for continuous refinement, the cadence of deliveries for predictability, and the importance of communication between teams. The experience gained can be replicated in other contexts, adjusting the checklist items according to the critical issues identified, although the process requires adaptations for different domains.

Based on the experience reported in this study, it can be concluded that the use of checklists in the code review process had a concrete positive impact on code quality. In addition to contributing to standardization, clarity, and consistency in reviews. Future research may explore continuous review and adjustment as the team and project mature, as well as exploring the use of checklists as a form of review in other software development contexts.

References

[1] Dr. D. I. De Silva, W. A. C Pabasara, S. V Sangkavi, Wijerathne L.G.A.T.D, Wijesundara W.M.K.H, and Reezan S.A. "The Effectiveness of Code Reviews on Improving Software Quality: An Empirical Study". In: *International Journal of Recent Technology and Engineering (IJRTE)* 12 (2 July 2023), pp. 1–10. ISSN: 22773878. DOI: 10.35940/ijrte.B7666.0712223. URL: https://www.ijrte.org/portfolio-item/B76660712223/.

[2] Alberto Bacchelli and Christian Bird. "Expectations, outcomes, and challenges of modern code review". In: *2013 35th International Conference on Software Engineering (ICSE)*. IEEE, May 2013, pp. 712–721. ISBN: 978-1-4673-3076-3. DOI: 10.1109/ICSE.2013.6606617.

[3] Oleksii Kononenko, Olga Baysal, and Michael W. Godfrey. "Code review quality". In: *Proceedings of the 38th International Conference on Software Engineering*. ACM, May 2016, pp. 1028–1038. ISBN: 9781450339001. DOI: 10.1145/2884781.2884840.

[4] Xiaofeng Han, Amjed Tahir, Peng Liang, Steve Counsell, Kelly Blincoe, Bing Li, and Yajing Luo. "Code Smells Detection via Modern Code Review: A Study of the OpenStack and Qt Communities". In: (May 2022). URL: http://arxiv.org/abs/2205.07535.

[5] Friederike Hillemann, Joseph Burant, Antica Culina, and Stefan Vriend. *Code review in practice: A checklist for computational reproducibility and collaborative research in ecology and evolution*. Apr. 2025. DOI: 10.32942/X26S6P.

[6] Shane McIntosh, Yasutaka Kamei, Bram Adams, and Ahmed E. Hassan. "An empirical study of the impact of modern code review practices on software quality". In: *Empirical Software Engineering* 21 (5 Oct. 2016), pp. 2146–2189. ISSN: 1382-3256. DOI: 10.1007/s10664-015-9381-9.

[7] C.F. Kemerer and M.C. Paulk. "The Impact of Design and Code Reviews on Software Quality: An Empirical Study Based on PSP Data". In: *IEEE Transactions on Software Engineering* 35 (4 July 2009), pp. 534–550. ISSN: 0098-5589. DOI: 10.1109/TSE.2009.27.

[8] Pavlína Wurzel Gonçalves, Enrico Fregnan, Tobias Baum, Kurt Schneider, and Alberto Bacchelli. "Do Explicit Review Strategies Improve Code Review Performance?" In: *Proceedings of the 17th International Conference on Mining Software Repositories*. ACM, June 2020, pp. 606–610. ISBN: 9781450375177. DOI: 10.1145/3379597.3387509.

[9] Rufus Olalere Oladele and H. O. Adedayo. “On Empirical Comparison of Checklist-based Reading and Adhoc Reading for Code Inspection”. In: *International Journal of Computer Applications* 87 (1 Feb. 2014), pp. 35–39. ISSN: 09758887. DOI: 10.5120/15174-3251.

[10] Ratnadira Widyasari, Ting Zhang, Abir Bouraffa, Walid Maalej, and David Lo. “Explaining Explanations: An Empirical Study of Explanations in Code Reviews”. In: *ACM Transactions on Software Engineering and Methodology* 34 (6 July 2025), pp. 1–30. ISSN: 1049-331X. DOI: 10.1145/3708518.

[11] Doriane Olewicki, Sarra Habchi, and Bram Adams. “An Empirical Study on Code Review Activity Prediction and Its Impact in Practice”. In: *Proceedings of the ACM on Software Engineering* 1 (FSE July 2024), pp. 2238–2260. ISSN: 2994-970X. DOI: 10.1145/3660806.

AUTORES



Pedro Alyson Teixeira dos Santos is a graduate student in Systems Analysis and Development at the Federal Institute of Ceará. He has experience in the field of Software Engineering, with an emphasis on Software Development, working mainly in the following areas: automation and secure development.



Samuel Sousa Teles holds a bachelor’s degree in Computer Networks from the Federal Institute of Ceará (2025) and is pursuing a master’s degree in Computer Science at the State University of Ceará. He has experience in Software Engineering, with an emphasis on Software Quality and Bug Detection, working mainly on the following topics: secure development and bug detection.



Rubens Abraão da Silva Sousa holds a bachelor’s degree in Systems Analysis and Development from the Federal Institute of Ceará (2025) and is pursuing a master’s degree in Computer Science at the State University of Ceará. He is currently the technical leader for requirements and testing in an IT bill. He has experience in Software Engineering, with an emphasis on Software Quality, Defect and Vulnerability Detection, and Artificial Intelligence. He works mainly in the following areas: secure development, API vulnerability detection, and bug prediction.



Manoel Lopes Filho holds a bachelor’s degree in Teleinformatics Engineering from the Federal University of Ceará (2010) and a professional master’s degree in Applied Computing from the State University of Ceará (2014). He is currently an EBTT professor at the Federal Institute of Ceará – Canindé Campus and is pursuing a PhD in Computing at the Federal University of Ceará. He has experience in the field of Computer Science, with an emphasis on Computer Systems.



Alex Lacerda Ramos works as a professor at the Federal Institute of Education, Science, and Technology of Ceará (IFCE). He holds a PhD and a master’s degree in Applied Computer Science from the University of Fortaleza (UNIFOR), with a PROSUP-CAPES scholarship. He has a specialization in Teaching from IFCE and Dom Bosco Catholic University (UCDB). He holds a bachelor’s degree in Systems Analysis and Development from the Federal Institute of Piauí (IFPI). He has experience in the field of Computer Science, with an emphasis on Information Systems Development and Computer Networks. His areas of interest include the Internet of Things (IoT), cloud computing, artificial intelligence, machine learning, and deep learning.

Comparing Package-Based and Sprint-Based Delivery Structures: An Experience Report in an RD&I Environment

Luis Felipe de Freitas Gomes, Leina Soares Lima, Rubens Abraão da Silva Sousa
Manoel Lopes Filho, Alex Lacerda Ramos, Ronaldo Tadeu Pontes Milfont
Instituto Federal de Educação, Ciência e Tecnologia do Ceará
luis.gomes@darmlabs.ifce.edu.br, leina.lima@darmlabs.ifce.edu.br, rubens.silva@darmlabs.ifce.edu.br,
manoel.filho@ifce.edu.br, alex.ramos@ifce.edu.br, ronaldotpmilfont@gmail.com

***Abstract*—The way agile teams structure delivery increments may influence defect accumulation, release cadence, and deadline predictability. This study compares two delivery structures applied to the same web-based system in a Research, Development, and Innovation (RD&I) laboratory: (i) consolidated feature packages across three sprints and (ii) single-feature delivery per sprint. The analysis was based on objective data extracted from Jira over a 12-month period, including bugs per sprint, bugs per functionality, release frequency, lead time, and deadline compliance. Results show that the sprint-based model reduced the average number of bugs per delivered functionality by approximately 52% and doubled the release frequency when compared to the package-based model, while maintaining full deadline compliance. The findings indicate that delivery structure directly affects defect distribution and release dynamics. This experience report provides empirical evidence to support decision-making regarding increment organization in agile environments.**

***Index Terms*—Agile development, empirical study, release management, sprint delivery.**

***Resumo*—A forma como equipes ágeis estruturam seus incrementos de entrega pode influenciar o acúmulo de defeitos, a cadência de releases e a previsibilidade de prazos. Este estudo compara duas estruturas de entrega aplicadas ao mesmo sistema web em um laboratório de Pesquisa, Desenvolvimento e Inovação (PD&I): (i) pacotes consolidados ao longo de três sprints e (ii) entrega de uma funcionalidade por sprint. A análise foi baseada em dados objetivos extraídos do Jira durante 12 meses, incluindo bugs por sprint, bugs por funcionalidade, frequência de releases, lead time e cumprimento de prazo. Os resultados indicam que o modelo por sprint reduziu em aproximadamente 52% a média de bugs por funcionalidade entregue e dobrou a frequência de releases em comparação ao modelo em pacotes, mantendo cumprimento integral dos prazos. Os achados evidenciam que a estrutura de entrega influencia diretamente a dinâmica de validação e distribuição de defeitos.**

***Palavras chave*—Desenvolvimento ágil, entrega por sprint, estudo empírico, gestão de releases.**

I. Introduction

The growing technical and organizational complexity of software systems has intensified the challenge of delivering value frequently while maintaining quality and predictability. Empirical evidence indicates that even in organizations adopting agile methods, delivery performance varies significantly depending on planning discipline, coordination mechanisms, and the way increments are structured for release [1, 2]. Thus, it is not merely the adoption of Scrum that determines project success, but also how functionalities are organized and distributed across iterative cycles.

In agile environments, incremental delivery aims to reduce risk by shortening feedback loops and enabling early validation of functionality. However, studies conducted in industrial contexts show that factors such as inter-team dependencies, scope volatility, and coordination complexity can compromise delivery flow and deadline compliance [3, 4]. As a result, many teams adapt Scrum in practice, modifying how functionalities are consolidated and released, which does not always align with the theoretical principle of delivering a potentially shippable increment at the end of each sprint.

Research, Development, and Innovation (RD&I) environments introduce additional challenges when compared to traditional industrial software development. These contexts are typically characterized by high scope volatility, exploratory requirements, and increased technical uncertainty [3, 2]. Functionalities are often not fully defined at the beginning of development cycles and evolve continuously based on stakeholder feedback. This dynamic nature intensifies the difficulty of maintaining delivery predictability, controlling defect accumulation, and structuring delivery increments effectively [1].

Despite the maturity of agile research, challenges related to delivery organization remain. Inadequate release structuring strategies may lead to the accumulation of changes, concentration of defects, delayed value perception by customers, and reduced predictability [1]. A recurring dilemma concerns whether teams should group multiple functionalities into larger packages released after several cycles or deliver individual functionalities at the end of each sprint. This decision directly affects workload distribution, exposure to integration risks, and stakeholder engagement.

Recent investigations have examined release planning practices and coordination mechanisms in industrial agile environments [3, 4]. Although these studies provide relevant evidence on factors influencing delivery performance, they do not perform direct comparisons between different increment organization strategies within the same project context. Therefore, an empirical gap remains regarding the comparative evaluation of multi-sprint bundled deliveries versus single-feature

sprint deliveries, particularly in Research, Development, and Innovation (RD&I) environments.

Given this gap, this paper presents a comparative experience report of two agile delivery strategies applied within the same RD&I project conducted in partnership with a higher education institution. The first strategy organizes deliveries into feature packages consolidated over multiple sprints, whereas the second adopts single-feature deliveries at each sprint. Based on quantitative and qualitative data extracted from Jira, including number of defects, tasks, user stories, and deadline compliance, this study analyzes how delivery structuring influences defect accumulation, predictability, and perceived value.

This paper is structured as follows: Section II presents the related work and theoretical background on release planning and increment organization in agile methods; Section III describes the adopted method and the empirical study context; Section IV presents the results and comparative analysis of the investigated strategies; and finally, Section V discusses the conclusions, practical implications, and directions for future work.

II. Related Work

This section presents related studies representing the current state of the art in the topic addressed in this work.

Bi et al. [5] empirically investigated how *release notes* are produced and used in real-world projects, aiming to understand their role in communication between teams and end users. The study analyzed artifacts and organizational practices to identify patterns in the production and consumption of release notes. The results revealed significant variations in content quality and structure. However, the focus is limited to the communication artifact itself and does not examine how the underlying increment structure influences operational delivery metrics in production.

Sampaio et al. [6] proposed COPLAM, a method to support continuous planning at the team level, with the goal of reducing inconsistencies between backlog refinement and sprint planning decisions. The method was evaluated through practical application and structured observation in an organizational context. The authors report improvements in planning continuity and work organization. Nevertheless, the study does not assess post-release outcomes nor measure effects on production metrics such as defect rates, release frequency, or deadline compliance.

The work presented in [7] aimed to incorporate explicit uncertainty treatment into release planning. The proposed approach was evaluated by comparing plans generated with and without uncertainty analysis. The results indicate increased robustness of the generated plans. However, the evaluation focuses on planning quality and does not include longitudinal assessment of real delivery performance, nor does it compare alternative increment consolidation structures in operational environments.

Pasuksmit et al. [8] examined iterative planning predictability by proposing models to anticipate documentation changes in work items. The study employed data analysis techniques to identify factors affecting planning reliability. The findings suggest potential for anticipating inconsistencies during execution. However, the research concentrates on artifact stability and does not analyze how different increment organization strategies affect defect rates, release cadence, or lead time in production settings.

Mahdavi-Hezavehi et al. [9] investigated heuristics and metrics associated with the use of *feature toggles*, aiming to support incremental releases and mitigate integration risks. The study was based on empirical analysis of repositories and evaluation of structural metrics. The authors provide guidelines for managing toggles over time. Nonetheless, the research does not compare delivery structures nor quantitatively assess their impact on production indicators such as defects per functionality or schedule predictability.

Uludağ et al. [10] conducted a systematic literature review to map practices, challenges, and trends in large-scale agile development. The method followed a structured protocol for search, selection, and evidence classification. The results highlight coordination issues, dependencies, and release planning challenges. However, as a secondary and aggregative study, it does not provide controlled empirical comparison of different delivery structures within the same organizational context.

Noei et al. [11] analyzed release-wise refactoring patterns using large-scale mining techniques to identify recurring structural changes across successive versions. The results show variations in refactoring types and intensity throughout the software lifecycle. However, the study does not control for delivery organization models nor explicitly relate release structure to defect metrics or predictability within a single project environment.

Pérez-Piqueras et al. [12] introduced ProjectION, an intelligent tool to support release planning through multi-objective optimization. The evaluation included experimental comparison between manually generated and algorithmically generated plans. The findings indicate improved solution quality under optimization criteria. However, the study evaluates planning effectiveness rather than operational performance after execution and does not empirically compare distinct increment organization strategies within the same system.

In contrast to the studies above, this work conducts a longitudinal comparison of two delivery structures applied to the same web-based system and team, using operational metrics extracted directly from Jira. By maintaining organizational context and functional equivalence across periods, the study isolates the effect of increment structure on defects per functionality, release frequency, lead time, and deadline compliance. Therefore, the main contribution lies in providing empirical evidence focused specifically on the unit of increment consolidation and its observable effects in production.

III. Method

The study was conducted on a single web-based system developed in a Research, Development, and Innovation (RD&I) laboratory over two consecutive six-month periods. During the first period, a package-based delivery strategy was adopted, while in the second period a sprint-based delivery strategy was implemented. The functionalities developed belonged to the same system and domain and were considered equivalent

in relative complexity, based on story point estimations conducted during sprint planning. In total, approximately ten main functionalities were analyzed across both periods.

The team remained the same throughout both phases and consisted of one Product Owner, one Product Engineer, three requirements analysts, developers, and four testers. No team members were replaced during the study. The team's experience evolved naturally over time due to project continuity, but roles and responsibilities remained unchanged. Sprints had a fixed and uniform duration throughout the entire observation period.

All activities were managed using Jira, which was used to record user stories, technical tasks, defects, story point estimates, and release versions. Quantitative data were collected regarding deadline compliance, number of bugs per sprint, number of bugs per functionality, lead time, and release frequency in production.

A. Package-Based Delivery Strategy

In the package-based delivery strategy, planning considered a cycle composed of three consecutive sprints to consolidate functionalities. The cycle scope included multiple features, improvements, and accumulated corrections. The process began with elicitation and scope definition conducted by the Product Owner and the requirements team, resulting in user stories and technical tasks registered in Jira.

Figure 1 presents an overview of the delivery flow adopted in this package-based model.

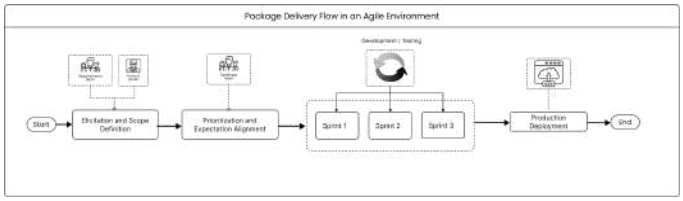


Fig. 1. Package-Based Delivery Flow in an Agile Environment.

As illustrated in Figure 1, the workflow is organized around a three-sprint cycle in which functionalities are gradually implemented and integrated before the final release. Development activities occur across the sprints, while validation is continuously performed in a staging environment by the testing team.

During Sprint Planning, tasks were estimated using Planning Poker based on story points. Capacity was calculated considering 15 story points per developer per sprint. Tasks were distributed across the three sprints and monitored through daily meetings.

At the end of the third sprint, a consolidated testing phase was conducted covering all functionalities included in the package. After validation, the complete set of changes was released to production as a single version. Bugs identified during the cycle were recorded in Jira and prioritized for correction either within the ongoing cycle or in subsequent sprints.

B. Sprint-Based Delivery Strategy

In contrast to the package-based delivery strategy described previously, the sprint-based model organizes work so that each sprint results in the delivery of a single fully completed functionality. In this approach, each iteration aims to complete the entire cycle of definition, development, validation, and production release.

The process began with detailed scope definition involving the client and the Product Owner, followed by technical analysis conducted by the requirements team and the Product Engineer. Functionalities were decomposed into user stories and technical tasks before the sprint started.

Figure 2 presents an overview of the delivery flow adopted in the sprint-based model.

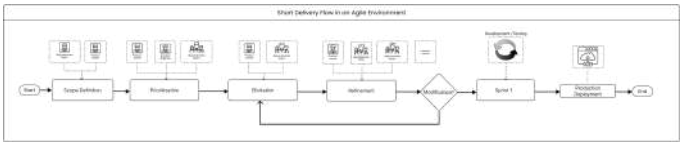


Fig. 2. Sprint-Based Delivery Flow in an Agile Environment.

As illustrated in Figure 2, all activities required for delivering the functionality-including planning, development, validation, and release preparation-are performed within the same sprint cycle, allowing each iteration to result in a complete production delivery.

Refinement was performed prior to Sprint Planning with participation from requirements, development, and testing teams. Estimation was also conducted using Planning Poker. Capacity was redefined to 10 story points per developer per sprint, reserving approximately 5 points to address unforeseen issues such as bug fixes and emergent adjustments.

During the sprint, the functionality was implemented and continuously validated in the staging environment. Identified defects were recorded and addressed within the same cycle whenever feasible. At the end of the sprint, the validated functionality was released to production, completing the cycle of definition, development, validation, and delivery.

To summarize the structural differences between the investigated delivery strategies, Table I presents a comparison of the main characteristics of the package-based and sprint-based models.

TABLE I
COMPARISON OF DELIVERY MODEL CHARACTERISTICS

Characteristic	Package-Based	Sprint-Based
Primary delivery unit	Package of features	Single feature
Production release at the end of each sprint	×	✓
Consolidated release after multiple sprints	✓	×
Final validation concentrated before release	✓	×
Validation distributed across sprints	×	✓
Integration of multiple features per release	✓	×
Incremental delivery per feature	×	✓
Full delivery cycle completed within a sprint	×	✓

IV. Results and Discussions

The comparative analysis considered three main indicators: (i) deadline compliance per sprint, (ii) number of bugs per

sprint and per delivered functionality, and (iii) frequency of production releases. Data were extracted directly from Jira, covering nine sprints in the package-based model and six sprints in the sprint-based model. The comparison was performed using averages and proportions, considering the structural differences between the delivery cycles.

A. Package-Based Delivery Model

In the package-based model, three cycles composed of three sprints each were analyzed, totaling nine sprints. During this period, six functionalities were consolidated into three production releases. A total of 323 bugs were fixed, resulting in an average of approximately 35.9 bugs per sprint and 53.8 bugs per delivered functionality.

Table II presents the empirical data collected for each sprint in the package-based model, including the number of functionalities, user stories, tasks, bugs, and production releases.

TABLE II
EMPIRICAL DATA PER SPRINT – PACKAGE-BASED MODEL

Package	Sprint	Feat.	US	Tasks	Bugs	Release
P1	S1	1	1	21	20	Yes
P1	S2	1	2	20	27	Yes
P1	S3	0	1	15	26	Yes
P2	S1	1	2	21	35	Yes
P2	S2	1	0	15	27	Yes
P2	S3	1	1	22	54	Yes
P3	S1	1	1	16	42	Yes
P3	S2	0	0	4	34	Yes
P3	S3	0	0	5	58	Yes

As shown in Table II, a higher concentration of bugs was observed in the final sprints of each package, indicating an accumulation of validation activities close to the release stage. In addition, not all functionalities initially planned were completed within the original cycle, requiring internal redistribution of demands across sprints.

B. Sprint-Based Delivery Model

In the sprint-based model, six consecutive sprints were analyzed, each resulting in the release of one functionality to production. A total of 153 bugs were fixed, resulting in an average of approximately 25.5 bugs per sprint and 25.5 bugs per delivered functionality.

Table III presents the empirical data collected for each sprint in the sprint-based model, including the number of delivered functionalities, user stories, tasks, bugs, and production releases.

TABLE III
EMPIRICAL DATA PER SPRINT - SPRINT-BASED MODEL

Sprint	Features	US	Tasks	Bugs	Release
S1	1	1	19	23	Yes
S2	1	5	18	20	Yes
S3	1	2	38	11	Yes
S4	1	1	40	24	Yes
S5	1	9	62	45	Yes
S6	1	4	32	30	Yes

As shown in Table III, the distribution of bugs presented lower variation across sprints when compared to the package-based model. All planned functionalities were released within the deadline of their respective sprint, and each sprint resulted in a production deployment, increasing the release frequency throughout the analyzed period.

C. Indicator Comparison

Based on the empirical data presented in Tables II and III, the sprint-based model presented an approximate 52% reduction in the average number of bugs per delivered functionality when compared to the package-based model.

Release frequency also increased, as six functionalities were delivered in six independent releases, whereas in the package-based model the same number of functionalities was consolidated into three production releases. This difference effectively doubled the cadence of production deployments.

Deadline compliance was complete in the sprint-based model, while the package-based model required internal replanning in some cycles. In addition, the average lead time per functionality was lower in the sprint-based model due to the completion and release of each functionality within a single sprint.

D. Analytical Summary

The results indicate structural differences in defect distribution, release cadence, and delivery predictability between the analyzed models. In the package-based model, validation and correction activities tended to concentrate near the consolidated release points, leading to a higher accumulation of defects in the final sprints of each cycle.

In contrast, the sprint-based model distributed validation and deployment activities across consecutive iterations. This continuous validation process reduced the average number of bugs per delivered functionality and increased the frequency of production releases, contributing to greater delivery predictability.

V. FINAL CONSIDERATIONS

This study compared two agile delivery strategies applied to the same web-based system in an RD&I environment: consolidated feature packages across three sprints and single-feature delivery per sprint. The analysis was based on objective metrics extracted from Jira, including bugs per sprint, bugs per functionality, release frequency, lead time, and deadline compliance. The results indicate structural differences in defect distribution, release cadence, and predictability between the two models.

The sprint-based model presented lower average defects per delivered functionality and higher release frequency, while maintaining full deadline compliance across the analyzed sprints. In contrast, the package-based model concentrated corrections near consolidated releases and required internal replanning in some cycles. These findings suggest that the way increments are structured can directly influence validation dynamics and defect accumulation.

From a practical perspective, this experience highlights that incremental delivery at the end of each sprint may improve visibility of progress and distribute validation effort over time. However, the study was conducted in a single academic RD&I

laboratory with a stable team and one system domain, which limits the external validity of the results.

Future work includes replicating the study in industrial environments, expanding the sample size, and applying statistical analyses to strengthen the comparison. Further investigation into how team maturity and backlog volatility influence delivery structure is also recommended.

Additionally, future studies should investigate the infrastructure and operational costs required to sustain sprint-based delivery models, particularly regarding CI/CD pipelines, automated testing, and continuous deployment practices. Although the sprint-based approach demonstrated improvements in defect distribution and delivery cadence, its adoption may require a higher level of technical maturity and investment in automation, which can represent a barrier for teams operating in less structured environments.

Referências

[1] Maria Paasivaara, Benjamin Behm, Casper Lassenius e Minna Hallikainen. "Large-scale agile transformation at Ericsson: a case study". Em: *Empirical Software Engineering* 23 (5 out. de 2018), pp. 2550–2596. ISSN: 15737616. DOI: 10.1007/s10664-017-9555-8.

[2] Marco Kuhrmann et al. "What Makes Agile Software Development Agile?" Em: *IEEE Transactions on Software Engineering* 48 (9 set. de 2022), pp. 3523–3539. ISSN: 19393520. DOI: 10.1109/TSE.2021.3099532.

[3] Elvan Kula, Eric Greuter, Arie Van Deursen e Georgios Gousios. "Factors Affecting On-Time Delivery in Large-Scale Agile Software Development". Em: *IEEE Transactions on Software Engineering* 48 (9 set. de 2022), pp. 3573–3592. ISSN: 19393520. DOI: 10.1109/TSE.2021.3101192.

[4] Paul Temple, Gilles Perrouin, Mathieu Acher, Battista Biggio, Jean Marc Jézéquel e Fabio Roli. "Empirical assessment of generating adversarial configurations for software product lines". Em: *Empirical Software Engineering* 26 (1 jan. de 2021). ISSN: 15737616. DOI: 10.1007/s10664-020-09915-7.

[5] Tingting Bi, Xin Xia, David Lo, John Grundy e Thomas Zimmermann. "An Empirical Study of Release Note Production and Usage in Practice". Em: *IEEE Transactions on Software Engineering* 48 (6 jun. de 2022), pp. 1834–1852. ISSN: 19393520. DOI: 10.1109/TSE.2020.3038881.

[6] Rafaela Sampaio, Cristina Teles Cerdeiral e Gleison Santos. "A Method to Support Continuous Planning at the Team Level". Em: *ACM International Conference Proceeding Series*. Association for Computing Machinery, nov. de 2022. ISBN: 9781450395533. DOI: 10.1145/3493244.3493257.

[7] Olawole Oni e Emmanuel Letier. "Analyzing Uncertainty in Release Planning: A Method and Experiment for Fixed-Date Release Cycles". Em: *ACM Transactions on Software Engineering and Methodology* 31 (2 abr. de 2022). ISSN: 15577392. DOI: 10.1145/3490487.

[8] Jirat Pasuksmit, Patanamon Thongtanunam e Shanika Karunasekera. "Towards Reliable Agile Iterative Planning via Predicting Documentation Changes of Work Items". Em: *Proceedings of the 2022 Mining Software Repositories Conference (MSR 2022)*. Institute of Electrical e Electronics Engineers Inc., out. de 2022, pp. 35–47. ISBN: 9781450393034. DOI: 10.1145/3524842.3528445.

[9] Rezvan Mahdavi-Hezaveh, Nirav Ajmeri e Laurie Williams. "Feature toggles as code: Heuristics and metrics for structuring feature toggles". Em: *Information and Software Technology* 145 (mai. de 2022). ISSN: 09505849. DOI: 10.1016/j.infsof.2021.106813.

[10] Ömer Uludağ, Pascal Philipp, Abheeshta Putta, Maria Paasivaara, Casper Lassenius e Florian Matthes. "Revealing the state of the art of large-scale agile development research: A systematic mapping study". Em: *Journal of Systems and Software* 194 (dez. de 2022). ISSN: 01641212. DOI: 10.1016/j.jss.2022.111473.

[11] Shayan Noei, Heng Li e Ying Zou. "An Empirical Study on Release-Wise Refactoring Patterns". Em: *Proceedings of the ACM on Software Engineering* 2 (FSE jun. de 2025), pp. 403–424. DOI: 10.1145/3715734.

[12] Víctor Pérez-Piqueras, Pablo Bermejo e José A. Gámez. "Intelligent support for agile release planning: an empirical evaluation of ProjectION". Em: *Automated Software Engineering* 33 (2 fev. de 2026). ISSN: 15737535. DOI: 10.1007/s10515-026-00591-3.

Authors



Luis Felipe de Freitas Gomes is an undergraduate student in Systems Analysis and Development at the Federal Institute of Ceará (IFCE). He works in the area of Software Testing and Quality, with experience in quality management and technical leadership in Research, Development, and Innovation (R&D&I) projects at DARM Labs, conducted under the Brazilian Informatics Law, in partnership with national and multinational companies. He is interested in Software Engineering, Software Testing, and Quality Assurance, with a focus on system reliability and continuous improvement.



Leina Soares Lima is an undergraduate student in Systems Analysis and Development at the Federal Institute of Ceará (IFCE) and works in the field of Software Engineering, focusing on requirements elicitation and analysis, as well as experience in User Experience (UX) and User Interface (UI) Design, contributing to the development of user-centered solutions.



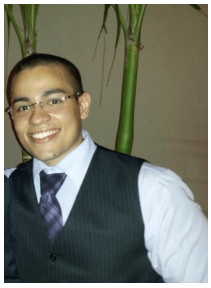
Rubens Abraão da Silva Sousa is a Master's student in Computer Science (UECE), holds a degree in Systems Analysis and Development (IFCE), and is a Electronics Technician (IFCE). He is a member of the Brazilian Computer Society (SBC) and the Brazilian Society for Computational and Applied Mathematics (SBMAC). He holds international certifications in: Scrum Fundamentals, DevOps Essentials, Six Sigma Yellow Belt, Remote Work and Virtual Collaboration, Kanban Foundation, and a national certification in the General Data Protection Law (LGPD). He has experience in software testing, pattern recognition, project management, and multiplatform development, working on state and national Research, Development, and Innovation (RDI) public calls at the Laboratory of Machine Learning and Innovative Projects (LAMPI), and serving as Software Quality and Requirements Leader at DARM Labs in projects under the Brazilian Informatics Law, in partnership with national and multinational companies. His interests include Software Engineering, Software Testing, AI-based Security Testing, and Motion Pattern Recognition.



Manoel Lopes Filho holds a degree in Teleinformatics Engineering from the Federal University of Ceará (2010) and a Professional Master's degree in Applied Computing from the State University of Ceará (2014). He is currently an EBTT professor at the Federal Institute of Ceará - Campus Canindé and is pursuing a Ph.D. in Computing at the Federal University of Ceará. He has experience in the field of Computer Science, with an emphasis on Computer Systems.



Alex Lacerda Ramos serves as a Professor at the Federal Institute of Education, Science and Technology of Ceará (IFCE). He holds a Ph.D. and a Master's degree in Applied Informatics from the University of Fortaleza (UNIFOR), with a PROSUP-CAPES scholarship. He has a specialization in Teaching from the Federal Institute of Education, Science and Technology of Ceará (IFCE) and from the Catholic University Dom Bosco (UCDB). He holds a degree in Systems Analysis and Development from the Federal Institute of Education, Science and Technology of Piauí (IFPI). He has experience in Computer Science, with an emphasis on Information Systems Development, Computer Networks, and Artificial Intelligence. His research interests include Internet of Things (IoT), cloud computing, artificial intelligence, machine learning, and deep learning.



Ronaldo Tadeu Pontes Milfont serves as a Professor at the Federal Institute of Education, Science and Technology of Ceará (IFCE). He is a Ph.D. candidate in Computer Science at the Federal University of Ceará (UFC). He holds a Master's degree in Teleinformatics Engineering from the Federal University of Ceará (UFC) and a degree in Teleinformatics Engineering from the same institution. He has experience in Computing, with an emphasis on Embedded Systems and Computer Networks. His current research interests include Internet of Things (IoT), Cloud Computing, and AI.

O Papel do Laboratório de PD&I no Desenvolvimento de Competências entre a Prática e a Teoria durante a graduação: Um Relato de Experiência

Luis Felipe de Freitas Gomes, Larissa Lacerda, Samuel Sousa Teles
Igor Ramsés Temóteo dos Santos, Rubens Abraão da Silva Sousa, Manoel Lopes Filho
Instituto Federal de Educação, Ciência e Tecnologia do Ceará
luis.gomes@darmlabs.ifce.edu.br, larissa.lacerda@darmlabs.ifce.edu.br, samuel.teles@darmlabs.ifce.edu.br,
igor.ramses@darmlabs.ifce.edu.br, rubens.silva@darmlabs.ifce.edu.br, manoel.filho@ifce.edu.br

***Resumo*—Diante das exigências do mercado de trabalho, a formação em Tecnologia da Informação (TI) requer abordagens que integrem teoria e prática. Este artigo apresenta uma avaliação da experiência de membros de um laboratório de PD&I, destacando sua contribuição ao longo da graduação. Adotou-se uma abordagem empírica, utilizando-se a escala *Likert* para avaliar quatro dimensões: aplicação do conhecimento, desenvolvimento técnico, motivação e orientação profissional. Os resultados demonstraram ganhos na aprendizagem, no engajamento e na percepção de carreira, ressaltando o papel estratégico do laboratório de PD&I na formação acadêmica.**

***Palavras chave*—Desenvolvimento Profissional, Escala *Likert*, Formação em Tecnologia da Informação, Laboratório de Pesquisa e Desenvolvimento**

***Abstract*—Faced with the demands of the job market, training in Information Technology (IT) requires approaches that integrate theory with practice. This article presents an evaluation of the experiences of members of an R&D&I laboratory, highlighting their contributions during their undergraduate studies. We adopted an empirical approach, using a Likert scale to evaluate four dimensions: knowledge application, technical development, motivation, and professional orientation. The results showed gains in learning, engagement, and career perception, highlighting the strategic role of the R&D&I laboratory in academic training.**

***Index Terms*—Information Technology Education, Likert Scale, Professional Development, Research and Development Laboratory**

I. Introdução

No contexto universitário atual, os cursos de graduação em Tecnologia da Informação (TI) registram crescimento no número de formandos para atender à demanda da indústria por profissionais qualificados [1]. Entretanto, a pressão por formação rápida pode negligenciar a aplicação prática do conhecimento, resultando em estudantes com habilidades teóricas amplas, mas com deficiências em competências interpessoais exigidas no ambiente empresarial [2].

Diante desse cenário, destaca-se a importância de estratégias metodológicas ativas que potencializem o processo formativo, como a Aprendizagem Baseada em Problemas (PBL) [3], que desperta o interesse dos alunos por laboratórios de Pesquisa, Desenvolvimento e Inovação (PD&I). Para John Dewey [4], essa iniciativa deve surgir da própria instituição ao apresentar situações do contexto real e permitir a participação dos estudantes em atividades práticas. Nesse modelo, denominado *Learning by Doing*, o aluno assume um papel mais ativo, aplica o conhecimento na prática e desenvolve habilidades para solucionar questões além da sala de aula [4].

Birru [5] ressalta que ambientes interativos alinhados às competências do século XXI exigem reforma educacional na área tecnológica, com mudanças curriculares e adoção de novas metodologias para promover inovação, *soft skills* e habilidades colaborativas. Draaijer [6] aponta que índices convencionais de rendimento teórico não preveem o desempenho em situações profissionais, pois não refletem competências aplicadas, como trabalho em equipe, gestão de tempo e autonomia.

Dessa forma, este estudo tem como objetivo analisar como experiências em ambientes reais, como laboratórios de PD&I, potencializam o desenvolvimento de competências transversais e habilidades sociotécnicas na formação acadêmica e no preparo profissional de estudantes de TI. Romão et al. [7] ressaltam que contextos reais são eficazes no desenvolvimento de *soft skills* e competências técnicas, estimulando a adoção de práticas semelhantes.

Este artigo relata a experiência em um laboratório de PD&I e como essa metodologia ativa facilita o desenvolvimento de habilidades colaborativas. A pesquisa envolveu estudantes dos cursos de Análise e Desenvolvimento de Sistemas, Redes de Computadores e Técnico em Informática, com idades entre 18 e 27 anos. Os resultados foram organizados em quatro eixos temáticos e obtidos por meio de um questionário com 16 perguntas. A análise indicou percepções favoráveis ao ambiente de PD&I, com impacto positivo na formação acadêmica e no desenvolvimento de competências interpessoais. Assim, o artigo propõe uma abordagem que conecta teoria e prática e contribui para a formação de profissionais de tecnologia alinhados às demandas do mercado contemporâneo.

O presente estudo está estruturado em cinco seções: (i) Introdução, que contextualiza o tema e define os objetivos; (ii) Trabalhos relacionados, com a fundamentação teórica sobre metodologias em PD&I; (iii) Método, que descreve o processo de pesquisa; (iv) Resultados e Discussões, com a análise da

experiência prática em ambiente de PD&I; e (v) Conclusão, que apresenta os aprendizados e sugere aplicações futuras da abordagem.

II. Trabalhos Relacionados

A Aprendizagem Baseada em Problemas (PBL) tem se destacado como uma metodologia eficaz no desenvolvimento de competências diversas. Dema et al. [8] demonstram que o PBL promove habilidades como aprendizagem independente, trabalho em equipe, mentoria, aprendizagem experiencial e gestão de projetos. Contudo, os autores não exploram a aquisição de competências técnicas específicas, como programação, controle de versão ou testes automatizados, nem aprofundam os efeitos do PBL sobre a criatividade e a adaptação a contextos tecnológicos. Essa lacuna evidencia a necessidade de estudos que analisem de forma mais direta a articulação entre PBL e o desenvolvimento técnico dos discentes.

De forma complementar, Malik et al. [9] apontam que metodologias ativas, como PBL, práticas laboratoriais e ensino invertido, aumentam significativamente a motivação dos alunos e suas percepções positivas sobre o processo de aprendizagem. Entretanto, a pesquisa foi realizada em apenas um curso de Ciência da Computação, o que limita a generalização dos resultados para outros contextos educacionais.

Por sua vez, Yuldashovna et al. [10] destacam que atividades práticas são fundamentais para reduzir a lacuna entre o conhecimento teórico e sua aplicação no mundo real. Além disso, ressaltam que a aprendizagem interdisciplinar favorece o desenvolvimento da criatividade, das habilidades de resolução de problemas e da inovação, aspectos essenciais em ambientes como os laboratórios de Pesquisa, Desenvolvimento e Inovação (PD&I). Contudo, o estudo não incorpora a perspectiva dos estudantes, o que limita a compreensão do impacto dessas práticas na formação discente.

Adicionalmente, Oliveira et al. [11] demonstram que a Aprendizagem Baseada em Projetos (ABP) é uma abordagem viável e tende a apresentar resultados superiores ao ensino expositivo tradicional quando aplicada de forma estruturada. No entanto, o estudo está restrito ao ensino de matemática no nível médio, o que limita sua aplicação a outros contextos educacionais.

Por fim, May et al. [12] destacam que os laboratórios desempenham papel fundamental na educação moderna, ao potencializar o aprendizado por meio de experimentos práticos e ambientes de realidade cruzada, além de incentivar o aprendizado baseado em pesquisa. Ainda assim, o estudo não apresenta evidências empíricas em contextos institucionais, dificultando a compreensão do impacto prático dessas abordagens.

Em síntese, os trabalhos analisados reforçam a relevância das metodologias ativas e das práticas interdisciplinares no contexto educacional. No entanto, observa-se uma lacuna na literatura quanto à análise empírica da experiência em laboratórios de PD&I como ambiente integrador entre teoria, desenvolvimento técnico e preparação profissional. Nesse sentido, o presente estudo contribui ao investigar, sob a perspectiva dos estudantes, como a participação em um laboratório de PD&I influencia o desenvolvimento de competências acadêmicas e profissionais.

III. Método

Esta pesquisa se caracteriza como quantitativa e exploratória, conforme definido por Gil [13], dado que busca compreender como as experiências vividas por estudantes de um laboratório de Pesquisa, Desenvolvimento e Inovação (PD&I) são potencializadas na formação acadêmica.

Os estudantes participantes são membros do laboratório de PD&I, que mantém parcerias com empresas nacionais e multinacionais, cuja área de atuação dos projetos envolve o desenvolvimento de software, inteligência artificial e redes de computadores. Atualmente, participam do laboratório 1 (um) estudante do curso técnico em Informática, 19 (dezenove) estudantes da graduação em Análise e Desenvolvimento de Sistemas e 2 (dois) estudantes da graduação em Redes de Computadores. Portanto, ao todo, a pesquisa conta com a participação de 22 alunos.

Utilizou-se como instrumento de coleta de dados um formulário com 16 perguntas, estruturadas em escala do tipo *Likert* de 5 pontos, divididas em quatro áreas: i) Aplicação Prática do Conhecimento, que busca compreender em que medida os conhecimentos teóricos adquiridos no curso foram aplicados na prática durante as atividades no laboratório; ii) Desenvolvimento Técnico, que visa compreender o aprimoramento de habilidades técnicas e operacionais por meio das atividades realizadas no laboratório; iii) Motivação e Engajamento Profissional, que analisa em que medida a participação no laboratório despertou o interesse pela área de atuação e fortaleceu o comprometimento com a futura carreira profissional; iv) Orientação Profissional e de Carreira, que busca compreender o apoio proporcionado pelo laboratório na definição de objetivos profissionais e na tomada de decisões sobre a trajetória de carreira. O formulário foi disponibilizado por meio de um grupo de comunicação no aplicativo de mensagens WhatsApp.

A análise foi realizada por meio de estatística descritiva, considerando a frequência absoluta das respostas, e os resultados foram apresentados em tabelas. Para participar do estudo, os respondentes precisaram aceitar o Termo de Consentimento Livre e Esclarecido (TCLE), e todos os dados foram anonimizados, a fim de garantir que os participantes não fossem identificados.

IV. Resultados e Discussões

Esta seção apresenta os resultados obtidos por meio do questionário aplicado aos integrantes do laboratório de Pesquisa, Desenvolvimento e Inovação (PD&I). O objetivo é extrair informações sobre práticas vivenciadas em ambientes reais e analisar como estas podem contribuir significativamente para a formação acadêmica de estudantes de TI. Além disso, os resultados são discutidos à luz da literatura, destacando suas implicações para o desenvolvimento educacional e profissional. A Tabela I apresenta os resultados referentes à pergunta 1 do questionário, que investigou se a aplicação prática do conhecimento reforçou a teoria apresentada no curso.

A análise dos dados revela que 95,5% dos participantes concordaram total ou parcialmente com a afirmação, enquanto apenas 4,5% permaneceram neutros. Esses dados indicam uma percepção amplamente positiva sobre o impacto da prática

TABELA I
A APLICAÇÃO PRÁTICA DO CONHECIMENTO REFORÇOU A TEORIA APRESENTADA NO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	17	77,3%
Concordo	4	18,2%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

na consolidação da teoria, corroborando estudos como o de Mohammed e Ozdamli [2], que evidenciam o papel crucial de experiências práticas bem estruturadas na formação profissional. Dessa forma, reforça-se a importância de atividades que promovam o pensamento crítico e contextualizado, estreitando a relação entre teoria e prática. A Tabela II refere-se à pergunta 2 do questionário, a qual investigou se a aplicação prática complementou o conteúdo teórico apresentado.

TABELA II
A APLICAÇÃO PRÁTICA COMPLEMENTOU O QUE FOI APRESENTADO NO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	14	63,6%
Concordo	8	36,4%
Não concordo, nem discordo	0	0%
Discordo	0	0%
Discordo Totalmente	0	0%

Os dados revelam que todos os participantes perceberam a prática como um complemento direto à teoria, com 100% de concordância. Isso destaca o papel das atividades práticas como mecanismos de ampliação da aprendizagem, conforme defendem autores como Dema e Choden [8], ao enfatizar que metodologias ativas como a Aprendizagem Baseada em Problema (PBL) promovem a integração entre teoria e aplicação real. A Tabela III apresenta a distribuição das respostas à pergunta 3 do questionário, que trata da conexão entre a prática e os conteúdos abordados no curso.

TABELA III
A APLICAÇÃO PRÁTICA SE CONECTOU COM OS CONTEÚDOS ABORDADOS NO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	10	45,5%
Concordo	9	40,9%
Não concordo, nem discordo	2	9,1%
Discordo	1	4,5%
Discordo Totalmente	0	0%

A maioria dos estudantes reconhece uma conexão entre a aplicação prática e os conteúdos do curso, o que reforça a efetividade do laboratório de PD&I na promoção da aprendizagem significativa. Isso está de acordo com estudos de Dema e Choden [8], que destacam o valor das metodologias baseadas em problemas como facilitadoras do engajamento e aplicação do conhecimento em contextos reais. A Tabela IV apresenta as respostas à pergunta 4, que investigou se a prática no laboratório de PD&I facilitou a resolução de exercícios e problemas abordados no curso.

A prática aplicada no laboratório de PD&I foi considerada por 95,4% dos participantes como facilitadora na resolução de problemas, com apenas 4,5% neutros. Estudos como o de

TABELA IV
A APLICAÇÃO PRÁTICA NO LABORATÓRIO DE PD&I FACILITOU A RESOLUÇÃO DE EXERCÍCIOS E PROBLEMAS VISTOS NO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	16	72,7%
Concordo	5	22,7%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

Newby e Fisher [14] reforçam que ambientes bem estruturados estimulam o raciocínio lógico, produtividade e engajamento dos estudantes. A Tabela V apresenta os dados da pergunta 5, que buscou avaliar se o laboratório contribuiu para o desenvolvimento de habilidades técnicas não contempladas no currículo tradicional.

TABELA V
O LABORATÓRIO DE PD&I PERMITIU DESENVOLVER HABILIDADES TÉCNICAS QUE NÃO SÃO VISTAS NO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	16	72,7%
Concordo	5	22,7%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

Com 95,5% de concordância, os resultados apontam que o ambiente laboratorial promove o desenvolvimento de *soft skills*, como trabalho em equipe. Oliveira e Silva [15] defendem que as metodologias ativas aplicadas em ambientes de PD&I ampliam o escopo técnico dos alunos, desenvolvendo habilidades práticas essenciais para o mercado. A Tabela VI analisa as vivências práticas e experiências reais de mercado proporcionadas pelo laboratório.

TABELA VI
O LABORATÓRIO DE PD&I PERMITIU VIVÊNCIAS PRÁTICAS E EXPERIÊNCIAS REAIS DO MERCADO?

Resposta	Quantidade	% Total
Concordo Totalmente	17	77,3%
Concordo	5	22,7%
Não concordo, nem discordo	0	0%
Discordo	0	0%
Discordo Totalmente	0	0%

A unanimidade dos respondentes demonstra que o laboratório de PD&I é reconhecido como um ambiente que aproxima os alunos das exigências do mercado. Succi e Wieandt [16] reforçam que vivências reais promovem o desenvolvimento de *soft skills* e aumentam a empregabilidade. A Tabela VII apresenta os resultados referentes ao desenvolvimento do pensamento crítico e de soluções técnicas.

TABELA VII
O LABORATÓRIO DE PD&I PERMITIU O DESENVOLVIMENTO DO PENSAMENTO CRÍTICO E DE SOLUÇÕES TÉCNICAS?

Resposta	Quantidade	% Total
Concordo Totalmente	15	68,2%
Concordo	6	27,3%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

Os dados indicam que 95,5% dos participantes perceberam que o ambiente laboratorial favorece o desenvolvimento do pensamento crítico. Isso está alinhado com Newby e Fisher [14], que afirmam que ambientes desafiadores promovem engajamento intelectual e autonomia. A Tabela VIII aborda o desenvolvimento de habilidades interpessoais.

TABELA VIII
O LABORATÓRIO DE PD&I PERMITIU O DESENVOLVIMENTO DE HABILIDADES INTERPESSOAIS?

Resposta	Quantidade	% Total
Concordo Totalmente	14	63,6%
Concordo	6	27,3%
Não concordo, nem discordo	2	9,1%
Discordo	0	0%
Discordo Totalmente	0	0%

Cerca de 91% dos participantes apontam que desenvolveram habilidades interpessoais. Essa conclusão reforça a ideia defendida por Fifani et al. [17], segundo a qual ambientes colaborativos são essenciais para o desenvolvimento de competências sociais. A Tabela IX explora a influência do laboratório no interesse pelas disciplinas do curso.

TABELA IX
O LABORATÓRIO DE PD&I MOTIVOU O INTERESSE NAS DISCIPLINAS DO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	10	45,5%
Concordo	8	36,4%
Não concordo, nem discordo	3	13,6%
Discordo	1	4,5%
Discordo Totalmente	0	0%

Com 81,9% de concordância, os dados indicam que o laboratório é uma fonte de motivação para o interesse em disciplinas acadêmicas, como mostrado por Romão et al. [7] em estudos que integram o PBL. A Tabela X trata da motivação para entrada no mercado de trabalho.

TABELA X
O LABORATÓRIO DE PD&I MOTIVOU A ENTRADA NO MERCADO DE TRABALHO?

Resposta	Quantidade	% Total
Concordo Totalmente	18	81,8%
Concordo	3	13,6%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

Quase a totalidade dos participantes apontou que a experiência em PD&I motivou a entrada no mercado de trabalho. Isso reforça a ideia de que experiências práticas e resolução de problemas reais favorecem a empregabilidade [7]. A Tabela XI apresenta os resultados da pergunta 11, que investigou se o laboratório de PD&I motivou os participantes a seguir uma carreira relacionada ao curso.

Dos participantes, 86,3% concordam total ou parcialmente que o laboratório os motivou a seguir carreira na área do curso. Esses dados reforçam os achados de Subasman e Aliyyah [18], que destacam o papel de ambientes com metodologias ativas no incentivo à permanência e interesse pelas áreas STEM. A Tabela XII apresenta as respostas à pergunta 12, relacionada à visualização da atuação profissional.

TABELA XI
O LABORATÓRIO DE PD&I MOTIVOU A SEGUIR A CARREIRA RELACIONADA AO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	14	63,6%
Concordo	5	22,7%
Não concordo, nem discordo	3	13,6%
Discordo	0	0%
Discordo Totalmente	0	0%

TABELA XII
O LABORATÓRIO DE PD&I AJUDOU A VISUALIZAR MINHA ATUAÇÃO PROFISSIONAL?

Resposta	Quantidade	% Total
Concordo Totalmente	15	68,2%
Concordo	7	31,8%
Não concordo, nem discordo	0	0%
Discordo	0	0%
Discordo Totalmente	0	0%

Com 100% de concordância, os dados revelam que os estudantes reconheceram no laboratório de PD&I um espaço de projeção da própria carreira. Florambel et al. [19] defendem que projetos de pesquisa e inovação auxiliam na construção de perfis profissionais e na tomada de decisões sobre o futuro. A Tabela XIII examina se a experiência no laboratório de PD&I trouxe clareza sobre possíveis atuações profissionais proporcionadas pelo curso.

TABELA XIII
A EXPERIÊNCIA NO LABORATÓRIO DE PD&I ME DEU CLAREZA SOBRE AS POSSÍVEIS ATUAÇÕES PROFISSIONAIS PROPORCIONADAS PELO CURSO?

Resposta	Quantidade	% Total
Concordo Totalmente	16	72,7%
Concordo	6	27,3%
Não concordo, nem discordo	0	0%
Discordo	0	0%
Discordo Totalmente	0	0%

A unanimidade das respostas aponta que o laboratório contribui para o entendimento das oportunidades profissionais decorrentes da formação. Borrego et al. [20] apontam que esse tipo de vivência favorece decisões conscientes quanto à trajetória profissional. A Tabela XIV traz os resultados da pergunta 14, que questiona se a experiência foi considerada importante para o currículo profissional dos participantes.

TABELA XIV
O LABORATÓRIO DE PD&I ME PROPORCIONOU UMA EXPERIÊNCIA QUE CONSIDERO IMPORTANTE PARA O MEU CURRÍCULO PROFISSIONAL?

Resposta	Quantidade	% Total
Concordo Totalmente	19	86,4%
Concordo	3	13,6%
Não concordo, nem discordo	0	0%
Discordo	0	0%
Discordo Totalmente	0	0%

Os dados reforçam a percepção de que o ambiente de PD&I fortalece a formação profissional. Thakur et al. [21] destacam que vivências práticas como estágios e projetos aumentam significativamente a empregabilidade e a relevância curricular dos estudantes. A Tabela XV refere-se ao impacto do laboratório no desenvolvimento da colaboração em equipe.

Com 95,5% de concordância, os resultados apontam que o ambiente laboratorial promove o desenvolvimento de *soft*

TABELA XV
O LABORATÓRIO DE PD&I CONTRIBUIU PARA O APRIMORAMENTO DA MINHA CAPACIDADE DE COLABORAÇÃO EM EQUIPE?

Resposta	Quantidade	% Total
Concordo Totalmente	15	68,2%
Concordo	6	27,3%
Não concordo, nem discordo	1	4,5%
Discordo	0	0%
Discordo Totalmente	0	0%

skills como trabalho em equipe. Alsmadi et al. [22] defendem que projetos colaborativos baseados em PBL são ideais para o fortalecimento dessas competências. A Tabela XVI encerra a seção, avaliando se o laboratório facilitou a criação de uma rede de contatos profissionais.

TABELA XVI
O LABORATÓRIO DE PD&I FACILITOU A CRIAÇÃO DE UMA REDE DE CONTATOS COM PROFISSIONAIS DA ÁREA DE TECNOLOGIA?

Resposta	Quantidade	% Total
Concordo Totalmente	14	63,6%
Concordo	6	27,3%
Não concordo, nem discordo	1	4,5%
Discordo	1	4,5%
Discordo Totalmente	0	0%

Os dados indicam que 90,9% dos participantes reconhecem que o laboratório favoreceu o networking com profissionais da área. Esse dado está alinhado com Costa et al. [23], que apontam a colaboração universidade-indústria como estratégia para ampliar oportunidades de inserção no mercado.

V. CONCLUSÃO

Este estudo buscou compreender de que forma a experiência adquirida em ambientes reais, como um laboratório de PD&I, pode contribuir para a formação acadêmica de estudantes de TI. A partir dos resultados, nota-se que limitações foram consideradas, como o número reduzido de participantes na amostra, o recorte amostral restrito aos estudantes de um único laboratório de Pesquisa, Desenvolvimento e Inovação (PD&I) e a delimitação do estudo a um único contexto educacional na área de TI.

Os resultados revelam que a vivência prática reforça e complementa a teoria: 95,5% dos estudantes afirmaram que a aplicação prática reforçou a teoria do curso e 100% concordam que a prática complementou o conteúdo teórico. Ademais, 86,4% reconheceram que as atividades realizadas estavam conectadas com os conteúdos do curso, e 95,4% relataram que tais experiências facilitaram a resolução de problemas abordados em sala de aula.

Além disso, 95,4% afirmam que o laboratório desenvolveu habilidades técnicas não abordadas no curso, o que demonstra que o desenvolvimento técnico obtido vai além da grade curricular oferecida. O ambiente de PD&I aproxima os estudantes das dinâmicas profissionais reais e aumenta seu engajamento com o futuro profissional, com 100% dos participantes reconhecendo que o laboratório ofereceu vivências reais do mercado e 95,4% relatando maior motivação para ingressar no mercado após a experiência.

Com base nesta investigação empírica, conclui-se que a experiência em laboratórios de PD&I contribui de maneira ampla e significativa tanto para a formação acadêmica quanto para o preparo profissional dos estudantes de TI. Futuras pesquisas podem ser conduzidas visando à ampliação da amostra e à integração de diferentes áreas, permitindo análises em múltiplos contextos educacionais.

REFERÊNCIAS

[1] Elvan Kula, Eric Greuter, Arie van Deursen e Georgios Gousios. "Factors Affecting On-Time Delivery in Large-Scale Agile Software Development". Em: *IEEE Transactions on Software Engineering* (2022). DOI: 10.1109/TSE.2021.3101192.

[2] Farhad Sadik Mohammed e Fezile Ozdamli. "A Systematic Literature Review of Soft Skills in Information Technology Education". Em: *Behavioral Sciences* 14.10 (2024). DOI: 10.3390/bs14100894.

[3] Elaine H. J. Yew e Karen Goh. "Problem-Based Learning: An Overview of its Process and Impact on Learning". Em: *Health Professions Education* 2.2 (2016), pp. 75–79. DOI: 10.1016/j.hpe.2016.01.004.

[4] John Dewey. *How We Think*. Disponível em: Global Grey eBooks. Boston: D. C. Heath e Company, 1933.

[5] Yonas Birru. "The Integration of 21st-Century Skills into the Higher Education Curriculum". Em: *Teacher Education and Curriculum Studies* 9.3 (2024), pp. 60–68. DOI: 10.11648/j.tecs.20240903.12.

[6] S. Draaijer. *A Report on the Current State of Online Proctoring Practices in Higher Education within the EU*. Rel. técn. OP4RE Project, 2017.

[7] Lucas Romao, Allysson Allex Araújo, Marcos Kalinowski, Simone D. J. Barbosa, Clarissa Barbosa e Helio Lopes. "Agile Minds, Innovative Solutions, and Industry-Academia Collaboration: Lean R&D Meets Problem-Based Learning in Software Engineering Education". Em: 2024.

[8] Chimi Dema e Ugyen Choden. "Impact of Project-Based Learning on Computer Science Education". Em: *Educational Innovation and Practice* 7.1 (2024). DOI: 10.17102/eip.7.2024.08.

[9] Khalid Malik e Meina Zhu. "Do project-based learning and flipped teaching enhance student learning". Em: *Education and Information Technologies* 28.3 (2023), pp. 3581–3604. DOI: 10.1007/s10639-022-11350-8.

[10] Qalandarova Orinposhsha Yuldashovna. "Formation of Professional Competence of Students of Technological Education". Em: *International Journal of Pedagogics* 4.10 (2024), pp. 172–179. DOI: 10.37547/ijp/Volume04Issue10-30.

[11] Sebastião Luís de Oliveira, Adriano Francisco Siqueira e Estaner Claro Romão. "Aprendizagem Baseada em Projetos no Ensino Médio: estudo comparativo entre métodos de ensino". Em: *Bolema: Boletim de Educação Matemática* 34.67 (2020), pp. 764–785. DOI: 10.1590/1980-4415v34n67a20.

[12] Dominik May e Claudius Terkowsky. "Contemporary educational approaches in instructional laboratories". Em: *European Journal of Engineering Education* 48.5

(2023), pp. 783–801. DOI: 10.1080/03043797.2023.2248819.

[13] Antônio Carlos Gil. *Como Elaborar Projetos de Pesquisa*. São Paulo: Atlas, 2002.

[14] Michael Newby e Darrell Fisher. "A Model of the Relationship Between University Computer Laboratory Environment and Student Outcomes". Em: *Learning Environments Research* 3 (2000), pp. 51–66.

[15] Oliveira Silva e Silva. "RCMOS - Revista Científica Multidisciplinar O Saber". Em: *RCMOS* (2024). ISSN: 2675-9128.

[16] Chiara Succi e Michaela Wieandt. "Walk the talk: soft skills' assessment of graduates". Em: *European Journal of Management and Business Economics* 28.2 (2019), pp. 114–125. DOI: 10.1108/EJMBE-01-2019-0011.

[17] Samah Fifani, Dounia Rabhi e Abdelhakim Qachar. "Exploring the Role of Interpersonal Skills and Educational Technologies in Higher Education: Perspectives from Teachers and Students". Em: *Multidisciplinary Reviews* 8.10 (2025). DOI: 10.31893/multirev.2025332.

[18] Iman Subasman e Rusi Rusmiati Aliyyah. "The Impact of Technological Transformation on Career Choices in the STEM Sector". Em: *Jurnal Kajian Pendidikan dan Psikologi* 1.2 (2023), pp. 129–142. DOI: 10.61397/jkpp.v1i2.94.

[19] Heveline Cristina Alves De Vasconcelos Florambel, Solonildo Almeida Da Silva e Sandro César Silveira Jucá. "A inserção de discentes em projetos de Pesquisa, Desenvolvimento e Inovação". Em: *Research, Society and Development* 8.11 (2019). DOI: 10.33448/rsd-v8i11.1471.

[20] Maura Borrego, Marialice Mastronardi, Hyungsok Choe e Risa Hartman. "The Impact of Undergraduate Research Experiences on Participants' Career Decisions". Em: *Journal of STEM Education* 22.2 (2021).

[21] Rashmi Thakur e Manoj Chavan. "Bridging Academia and Industry: A Study on the Skill Development Effects". Em: *Educational Administration Theory and Practices* (2024). DOI: 10.53555/kuey.v30i4.2276.

[22] Hiba Alsmadi e Gok Kandasamy. "Empowering computing students through multidisciplinary project based learning". Em: *Social Sciences & Humanities Open* 10 (2024), p. 101180. DOI: 10.1016/j.ssaho.2024.101180.

[23] Joana Costa, Ana Rita Neves e João Reis. "University-Industry Collaboration and Open Innovation as Enhancers of Firm Performance". Em: *Sustainability* 13.7 (2021), p. 3866. DOI: 10.3390/su13073866.

AUTORES



Luis Felipe de Freitas Gomes Graduando em Análise e Desenvolvimento de Sistemas pelo Instituto Federal do Ceará. Atua na área de Testes e Qualidade de Software, com experiência em gestão da qualidade e liderança técnica em projetos de Pesquisa, Desenvolvimento e Inovação no DARM Labs, em parceria com empresas nacionais e multinacionais. Tem interesse em Engenharia de Software, Testes de Software e Garantia da Qualidade, com foco na confiabilidade e melhoria contínua de sistemas.



Larissa Lacerda Graduada em Análise e Desenvolvimento de Sistemas pelo IFCE. Atua com análise de requisitos no Laboratório DARM e participa de projetos de pesquisa no Laboratório LAMPI.



Samuel Sousa Teles Mestrando em Ciência da Computação pela Universidade Estadual do Ceará, é graduado em Redes de Computadores pelo Instituto Federal do Ceará (2025). Possui experiência em Engenharia de Software, com ênfase em Qualidade de Software e Detecção de Bugs, atuando principalmente nos temas de desenvolvimento seguro e detecção de bugs.



Igor Ramsés Temóteo dos Santos Mestrando em Ciência da Computação pela Universidade Estadual do Ceará, é graduado em Análise e Desenvolvimento de Sistemas pelo Instituto Federal do Ceará (2025). Possui experiência em Engenharia de Software, com ênfase em Qualidade de Software e Detecção de Bugs, atuando principalmente nos temas de desenvolvimento seguro e detecção de bugs.



Rubens Abraão da Silva Sousa Mestrando em Ciência da Computação pela UECE, graduado em Análise e Desenvolvimento de Sistemas e técnico em Eletrônica pelo IFCE. Atua em projetos de Pesquisa, Desenvolvimento e Inovação (PD&I) nos laboratórios LAMPI e DARM Labs. Seus interesses incluem Engenharia de Software, Testes de Software e Reconhecimento de Padrões.



Manoel Lopes Filho Graduado em Engenharia de Teleinformática pela Universidade Federal do Ceará (2010) e mestre em Computação Aplicada pela Universidade Estadual do Ceará (2014). Atualmente é professor EBTT do Instituto Federal do Ceará (Campus Canindé) e doutorando em Computação pela Universidade Federal do Ceará. Atua na área de Ciência da Computação, com ênfase em Sistemas de Computação.

Métricas de Efetividade para Análise de Vulnerabilidades em Aplicações com Bancos de Dados: Um Comparativo NIST e DevSecOps

Bruno Rezende Higino de Cuba, Laura Pivoto Ambrósio, Álvaro Lúcio Almeida Ribeiro,
Guilherme Pedro Aquino, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
bruno.cuba@get.inatel.br, pivoto.laura@ges.inatel.br, alvaro.lucio@ges.inatel.br, guilhermeaquino@inatel.br, evandro.cesar@inatel.br

***Abstract*—The security of modern applications depends on protecting data stored in databases, as vulnerabilities in queries and access controls represent critical vectors for incidents. The analysis of these flaws often relies on heterogeneous methods, compromising reproducibility and increasing triage costs. This article presents a conceptual comparison between NIST framework–based approaches and DevSecOps practices applied to systems using relational and non-relational databases. As a contribution, it proposes an effectiveness measurement model grounded in the PDCA cycle, defining the object of measurement and indicators such as confirmation rate, false positive rate, triage time, scope coverage, recurrence after remediation, and reproducibility. The results indicate that NIST strengthens standardization and formal evidence criteria, while DevSecOps prioritizes automation and continuous retesting, making them complementary in different contexts.**

***Index Terms*—Database Security, Vulnerability Analysis, NIST, DevSecOps, Metrics, SQL/NoSQL Injection.**

***Resumo*—A segurança de aplicações modernas depende da proteção de dados em bancos de dados, sendo vulnerabilidades em consultas e controles de acesso vetores críticos de incidentes. A análise dessas falhas costuma empregar métodos heterogêneos, comprometendo a reprodutibilidade e elevando o custo de triagem. Este artigo apresenta uma comparação conceitual entre abordagens baseadas em frameworks do NIST e práticas DevSecOps aplicadas a sistemas com bancos relacionais e não relacionais. Como contribuição, propõe um modelo de medição de efetividade fundamentado no ciclo PDCA, definindo objeto de medição e indicadores como taxa de confirmação, falsos positivos, tempo de triagem, cobertura, reincidência e reprodutibilidade. Conclui-se que NIST reforça padronização e evidência formal, enquanto DevSecOps prioriza automação e retestes contínuos, sendo abordagens complementares em diferentes contextos.**

***Palavras chave*—Segurança de Banco de Dados, análise de vulnerabilidades, NIST, DevSecOps, métricas, injeção de SQL/NoSQL.**

I. Introdução

Aplicações modernas dependem de dados para operar, fazendo com que bancos de dados se tornem o núcleo de praticamente qualquer serviço digital. Sistemas de autenticação, comércio eletrônico, redes sociais, serviços de saúde, logística e plataformas educacionais armazenam e processam informações sensíveis como dados de usuários e clientes, endereços, históricos de compra, registros de atendimento, documentos e, em muitos casos, transações financeiras. Essa centralização transforma o banco de dados em um ativo crítico. Logo, quando há exposição, alteração indevida ou indisponibilidade desses dados, o impacto deixa de ser apenas técnico e passa a afetar diretamente indivíduos e organizações.

O risco é ainda maior quando o comprometimento envolve privilégios elevados. Se um agente malicioso obtém acesso administrativo, abre-se a possibilidade de exfiltração em larga escala, manipulação de registros, fraudes e sabotagem operacional. Em cenários corporativos, esse tipo de incidente pode interromper serviços, causar prejuízo financeiro e reputacional e, em casos extremos, comprometer a continuidade do negócio. Além disso, a diversidade tecnológica amplia a superfície de ataque. Bancos de dados relacionais convivem com bancos não relacionais em arquiteturas híbridas, frequentemente distribuídas (microsserviços) e integradas a provedores em nuvem. Em consequência, vulnerabilidades ligadas à forma como a aplicação constrói consultas, filtra dados e aplica controles de acesso podem surgir em diferentes camadas e tecnologias, exigindo um processo de análise que seja consistente e replicável para ambientes distintos [1].

Apesar de existirem diversas técnicas e ferramentas para análise de vulnerabilidades em sistemas que interagem com bancos de dados, a prática costuma ser heterogênea. Equipes combinam análise estática, análise dinâmica, revisões manuais e ferramentas específicas de varredura e configuração, muitas vezes com resultados que variam significativamente conforme o conjunto de ferramentas, suas configurações e o contexto do sistema. Esse cenário dificulta a padronização e a reprodutibilidade, pois dois analistas, ou duas ferramentas, podem produzir relatórios diferentes para o mesmo sistema [2]. Um ponto crítico nesse processo é a qualidade dos achados reportados. Falsos positivos, alertas que indicam incorretamente a presença de uma vulnerabilidade, aumentam o custo de triagem, reduzem a confiança no processo de segurança e geram ruído operacional [3]. Na prática, isso pode levar a desperdício de tempo, priorização inadequada e atrasos na correção de falhas reais, especialmente em ambientes com ciclos de entrega frequentes [4].

Nesse contexto, duas abordagens têm se destacado como referência para estruturar e operacionalizar segurança em

organizações. A primeira é baseada em frameworks do NIST, que oferecem diretrizes para gestão de risco e para execução de avaliações técnicas, enfatizando planejamento, escopo, evidências e relatórios [5, 2]. A segunda é o modelo DevSecOps, que busca integrar segurança ao ciclo de desenvolvimento e operação por meio de automação, testes contínuos e validações recorrentes ao longo do pipeline de entrega [6, 7]. Embora ambas sejam amplamente adotadas, elas diferem no foco, no nível de abstração e na forma como lidam com reprodutibilidade, evidência e controle de falsos positivos.

Diante desse cenário, esse trabalho compara como uma abordagem orientada por NIST e uma abordagem DevSecOps se comportam quando aplicadas à análise de vulnerabilidades relacionadas a bancos de dados, incluindo cenários relacionais e não relacionais. Em particular, busca-se compreender como cada abordagem favorece a consistência do processo, a qualidade dos achados e a validação de correções, reduzindo ruído e aumentando a confiabilidade dos resultados [8, 9]. O estudo considera bancos relacionais (com ênfase em vulnerabilidades como SQL Injection) e também ataques relevantes em bancos não relacionais (por exemplo, NoSQL injection), com foco especial na redução de falsos positivos por meio de critérios objetivos de evidência e validação. Como contribuição, pretende-se sistematizar pontos fortes e limitações de cada abordagem e orientar a construção de um processo de análise mais consistente e aplicável a cenários atuais de desenvolvimento e operação.

Este artigo está estruturado em cinco seções. A Seção II apresenta os fundamentos teóricos sobre superfície de ataque, vulnerabilidades relacionadas a consultas e critérios de validação de achados. A Seção III define o modelo de mensuração e o conjunto de métricas adotado para comparar as abordagens. A Seção IV discute os resultados da análise comparativa entre NIST e DevSecOps à luz dos indicadores propostos. Por fim, são apresentadas as conclusões e limitações do estudo na Seção V.

II. Fundamentos

A. Superfície de ataque e fluxo aplicação–banco de dados

Em sistemas orientados a dados, a interação entre aplicação e banco de dados define uma cadeia crítica de processamento: entradas fornecidas por usuários (ou outros serviços) são recebidas pela aplicação, passam por regras de validação e lógica de negócio, são transformadas em consultas por meio de *drivers*, bibliotecas ou camadas de abstração, e então executadas no mecanismo do banco. O resultado dessa execução retorna à aplicação, que o formata e entrega ao cliente. Esse fluxo concentra os principais pontos de risco, pois qualquer falha na forma como a aplicação constrói consultas, filtra dados ou aplica controles de acesso pode permitir comportamentos indevidos, exposição de informações e alteração não autorizada de registros [1].

Em aplicações web e APIs, parâmetros de URL, corpo de requisições (por exemplo, JSON), cabeçalhos, filtros de busca, ordenações, paginação, identificadores de recursos e dados enviados em etapas distintas de um fluxo (cadastro, atualização e recuperação) podem influenciar consultas ao banco [1]. Além disso, rotinas internas como tarefas agendadas, integrações entre serviços, importação de arquivos e módulos administrativos frequentemente operam com permissões elevadas, ampliando o impacto caso sejam exploradas. Em arquiteturas híbridas, nas quais coexistem bancos relacionais e não relacionais, o risco se estende a mecanismos de consulta baseados em documentos, filtros estruturados e operadores específicos de cada tecnologia. Logo, qualquer dado controlável por um agente externo que seja utilizado para compor uma consulta ou um filtro, sem tratamento adequado, pode introduzir condições inesperadas na execução do lado do banco [1].

B. Vulnerabilidades relacionadas a consultas

A SQL Injection é uma classe de vulnerabilidade em que dados de entrada influenciam indevidamente a estrutura lógica de uma consulta SQL, permitindo que o banco interprete parte do conteúdo fornecido como elemento da própria instrução, e não apenas como dado [1]. A causa raiz costuma estar associada à construção dinâmica de consultas sem separação adequada entre código de consulta e valores (por exemplo, concatenação de *strings*), combinada a validações insuficientes. Embora a manifestação varie conforme o contexto, é útil organizar SQL Injection em categorias conceituais para orientar a análise e a priorização [1].

Uma classificação recorrente diferencia SQL Injection conforme o canal de obtenção de resultado. Em cenários de retorno direto, a aplicação expõe efeitos da consulta em sua resposta, o que facilita a observação de anomalias. Em outras situações, os efeitos aparecem de modo indireto, por mudanças consistentes de comportamento, fluxo, mensagens ou tempo de resposta. Também é relevante distinguir casos em que a aplicação propaga erros do banco (gerando indícios) daqueles em que erros são tratados ou suprimidos, exigindo validação por outros sinais [2]. Há ainda casos de segunda ordem, em que a entrada é armazenada e executada posteriormente em outro ponto do sistema, fazendo com que o ponto de inserção e o ponto de execução não coincidam [1].

Ao ampliar o escopo para bancos não relacionais, surgem ataques conceitualmente análogos, frequentemente referidos como NoSQL injection. Nesses casos, a falha não envolve sintaxe SQL, mas sim a construção insegura de filtros e consultas em modelos baseados em documentos ou estruturas. O padrão de risco, entretanto, é semelhante: quando a aplicação incorpora entradas do usuário diretamente em estruturas de consulta, permitindo alteração indevida de condições lógicas, podem ocorrer *bypass* de autenticação, acesso indevido a documentos e exposição de dados [1]. Assim, ainda que SQL Injection seja um estudo de caso central, a lógica de análise precisa contemplar também a possibilidade de injeções e falhas equivalentes em bancos não relacionais.

C. Critérios de evidência, validação e falsos positivos

Para tornar a análise reprodutível e reduzir falsos positivos, é fundamental definir critérios objetivos do que constitui um achado confirmado [8, 9]. Um achado deve ser tratado como suspeito quando há indícios iniciais (por exemplo, alerta de ferramenta, padrão de código potencialmente inseguro ou comportamento anômalo), mas ainda sem evidência suficiente de

que uma vulnerabilidade é explorável no contexto avaliado. Ele passa a ser confirmado quando é possível demonstrar que uma entrada controlável influencia a lógica de consulta ou filtragem no banco, resultando em impacto de segurança, como acesso indevido, alteração não autorizada ou violação de controles de acesso. Ou quando a análise técnica comprova a construção insegura de consultas aliada à ausência de barreiras eficazes [2].

Como evidências válidas, prioriza-se consistência e triangulação. Sinais observáveis devem ser reproduzíveis, preferencialmente combinando mais de uma fonte [2]. Exemplos incluem diferenças repetíveis no comportamento da aplicação sob condições controladas, respostas que indiquem falhas de interpretação ou execução de consulta no banco, registros do lado do servidor que confirmem a consulta gerada, quando disponíveis em ambiente de teste, e verificação por inspeção técnica do ponto de construção da consulta ou filtro, demonstrando incorporação insegura de valores externos. A validação deve evidenciar impacto sem depender de dados reais, utilizando ambiente controlado e dados sintético, preservando segurança e ética do experimento.

Por fim, um achado só deve ser considerado resolvido após validação pós-correção. Isso envolve confirmar que a correção elimina a causa raiz. Por exemplo, separando parâmetros da lógica da consulta, ajustando filtros e controles de acesso e reforçando validações e permissões. Além disso, o reteste não deve reproduzir os sinais observados anteriormente. Esse ciclo de confirmação e reteste é essencial para reduzir falsos positivos, aumentar a confiabilidade do processo e permitir que diferentes avaliadores cheguem a conclusões equivalentes ao aplicar a mesma metodologia [4]. A partir desses fundamentos, a próxima seção estabelece um modelo de medição de efetividade e um conjunto de métricas para comparar, de forma objetiva, abordagens baseadas em NIST e em DevSecOps no contexto de aplicações que utilizam bancos de dados.

III. Medição de Efetividade e Métricas para Comparar Abordagens

A comparação entre abordagens de segurança tende a se tornar subjetiva quando não há critérios mensuráveis e indicadores de efetividade. A literatura sobre medição de controles em Sistemas de Gestão de Segurança da Informação reforça que a efetividade de controles precisa ser mensurada para apoiar decisão, desempenho e *accountability* [8, 9]. Nesse contexto, a mensuração deve ser vista como parte de um ciclo de melhoria contínua, frequentemente descrito pelo modelo PDCA (*Plan–Do–Check–Act*), onde a organização planeja metas e controles, executa, mede/monitora e então melhora com base no que foi observado [8].

Aplicado ao problema deste trabalho, o foco da mensuração é responder de forma objetiva em qual abordagem (orientada por NIST ou baseada em DevSecOps) fornece melhores resultados práticos para reduzir risco em aplicações com bancos de dados, especialmente no que diz respeito à reprodutibilidade do processo, qualidade dos achados e redução de falsos positivos.

A. Objeto, atributos, medidas e indicadores

Para tornar a comparação reprodutível, é necessário explicitar o caminho que transforma dados brutos em indicadores úteis. A abordagem de mensuração discutida na literatura baseada no ISO/IEC 27004 estrutura esse processo definindo: objeto de medição e seus atributos, medidas base, medidas derivadas, indicadores e critérios de decisão [8]. Medida base é a medida mais simples obtida diretamente a partir de um atributo, medida derivada é construída a partir de uma ou mais medidas base, e indicador é o resultado interpretável que estima ou avalia o estado de um atributo a partir de um modelo analítico [8]. A análise do indicador deve ser feita com critérios de decisão, que orientam se ações adicionais são necessárias e viabilizam análises de tendência ao longo do tempo [9].

Além disso, para que as métricas sejam úteis operacionalmente, recomenda-se que elas atendam a qualidades como SMART (*Specific, Measurable, Attainable, Repeatable, Time-Dependent*) e que sejam significativas, reprodutíveis e coletáveis de forma consistente [9]. No contexto deste trabalho, isso implica escolher métricas que possam ser coletadas tanto em processos mais *framework-driven* (NIST) quanto em *pipelines* automatizados (DevSecOps), mantendo comparabilidade.

B. Métricas e indicadores para comparação NIST vs DevSecOps

Para comparar NIST e DevSecOps no domínio de vulnerabilidades relacionadas a bancos de dados, define-se como objeto de medição o processo de análise e validação de vulnerabilidades aplicado a uma aplicação (ou conjunto de aplicações) que consulta um banco relacional e/ou não relacional. Os atributos observados nesse objeto de medição se concentram em três eixos: (i) qualidade dos achados (sinal vs ruído), (ii) eficiência operacional (tempo e esforço) e (iii) consistência do processo (replicabilidade e reteste). Esses atributos refletem diretamente o problema motivador descrito na introdução: volume de alertas, falsos positivos, validação de evidências e sustentação do processo ao longo do tempo [8].

Com base no modelo de medidas base $\rightarrow$ derivadas $\rightarrow$ indicadores, este trabalho adota métricas que permitem avaliar efetividade de forma prática, com foco em bancos de dados (incluindo SQL Injection e equivalentes em bancos não relacionais), conforme apresentado na Tabela I [8, 9]. Esses indicadores seguem o princípio de mensuração voltada à decisão: ao definir critérios de decisão e comparar indicadores ao longo do tempo, torna-se possível identificar se o processo está melhorando ou se a abordagem adotada está gerando ruído operacional, além de permitir comparações entre execuções em momentos distintos (análise de tendência) [9].

Para evitar interpretações subjetivas, recomenda-se estabelecer critérios de decisão para os indicadores [8]. Neste trabalho, os limiares serão definidos conforme o contexto do estudo de caso (por exemplo, aceitabilidade de falsos positivos e metas de tempo de triagem), e a coleta dos dados ocorrerá por ciclo de análise, permitindo comparar resultados de uma abordagem orientada por NIST (mais estruturada por etapas e evidências) com uma abordagem DevSecOps (mais contínua e automatizada) [2, 7].

TABELA I
MÉTRICAS E INDICADORES PARA COMPARAÇÃO ENTRE ABORDAGENS

Métrica	Medida Base	Medida Derivada	Indicador
Taxa de confirmação de achados	Número de achados reportados e número de achados confirmados	Percentual de achados confirmados sobre o total	Mede diretamente a redução de falsos positivos e a consistência do processo de triagem [4]
Taxa de falsos positivos	Contagem de achados não confirmados após critérios de evidência	Percentual de falsos positivos sobre o total reportado	Quanto menor, maior a precisão do processo [4]
Tempo médio de triagem/validação por achado	Tempo necessário para concluir a validação de um achado (suspeito → confirmado ou descartado)	Média por período ou ciclo de análise	Avalia eficiência operacional [8]
Cobertura do escopo avaliado	Número de *endpoints* ou fluxos mapeados e número efetivamente avaliados	Percentual de cobertura	Quanto maior, melhor a abrangência do processo
Taxa de reincidência após correção	Número de vulnerabilidades corrigidas e número que reaparecem em ciclos posteriores (ou falham em reteste)	Percentual de reincidência	Quanto menor, melhor a efetividade do ciclo de correção [2]
Reprodutibilidade do resultado	Concordância entre avaliadores (ou execuções independentes) quanto ao status do achado	Percentual de concordância	Quanto maior, maior a consistência metodológica [8]

C. Enquadramento das abordagens sob as métricas: aplicações novas versus legadas

Com a base de mensuração definida, a comparação entre NIST e DevSecOps pode ser feita considerando dois cenários. Em aplicações novas (*greenfield*), práticas DevSecOps tendem a favorecer detecção precoce, reteste contínuo e coleta automática de evidências [6, 7]. Já em aplicações legadas (*brownfield*), abordagens orientadas por NIST podem favorecer definição de escopo, priorização por risco, documentação e consolidação de evidências para correção planejada [5, 2].

Uma abordagem orientada por NIST tende a começar pela organização do problema em termos de risco, escopo e evidências, reduzindo ambiguidade e aumentando a reprodutibilidade do processo [5]. O NIST Cybersecurity Framework (CSF) estrutura resultados de cibersegurança por funções, enquadrando a análise de vulnerabilidades como um ciclo completo: governança, identificação, proteção, detecção, resposta e recuperação [5]. No nível operacional, guias como o NIST SP 800-115 enfatizam planejamento, execução estruturada, registro de evidências e reporte consistente [2]. Esse modelo favorece indicadores como taxa de confirmação, redução de falsos positivos, rastreabilidade e análise de tendência ao longo do tempo. Para aplicações novas, o Secure Software Development Framework (SSDF) reforça integração de requisitos de segurança desde o início [7]. Para aplicações em produção, o ponto forte é a padronização de avaliação, documentação e comparação entre ciclos.

DevSecOps integra práticas de segurança ao ciclo de desenvolvimento com ênfase em automação, feedback rápido e responsabilidade compartilhada [6]. Seu principal impacto nas métricas definidas está na frequência de reteste e na detecção precoce de regressões, influenciando diretamente a taxa de reincidência e a cobertura ao longo do tempo [7]. A abordagem fortalece o *shift-left*, permitindo identificar falhas relacionadas a acesso a banco de dados ainda nas fases iniciais [6]. Contudo, sem critérios claros de evidência e calibração adequada das ferramentas, a automação pode aumentar o volume de alertas e amplificar falsos positivos [4]. Assim, a efetividade do modelo depende da definição explícita de critérios de confirmação, priorização e aceitação de risco.

IV. DISCUSSÃO E COMPARAÇÃO ENTRE ABORDAGENS

A comparação entre abordagens orientadas por NIST e por DevSecOps deve basear-se nos indicadores definidos anteriormente, considerando o contexto organizacional e a maturidade do processo de desenvolvimento e operação [8]. Embora ambas busquem reduzir riscos e aumentar a confiabilidade dos sistemas, diferem em seus mecanismos operacionais e ênfases práticas.

Sob a perspectiva da qualidade dos achados, a abordagem NIST tende a oferecer maior rigor na definição de escopo, critérios de evidência e documentação [2]. Esse rigor favorece a distinção entre achados suspeitos e confirmados, contribuindo para maior taxa de confirmação e redução de falsos positivos [4]. Além disso, o planejamento estruturado e o registro consistente fortalecem a rastreabilidade e a justificativa técnica das decisões [2]. Em contraste, DevSecOps enfatiza a frequência de verificação e a detecção precoce de falhas por meio de automação e integração contínua [6]. Esse modelo permite identificar e retestar vulnerabilidades ao longo do ciclo de vida do software, reduzindo reincidência e ampliando cobertura [7].

Quanto à eficiência operacional, a abordagem NIST pode demandar maior esforço inicial devido à formalização de escopo e relatórios [2], mas tende a reduzir retrabalho em ambientes regulados [5]. DevSecOps, por sua vez, favorece ciclos curtos e feedback rápido, desde que os mecanismos automatizados estejam adequadamente calibrados [6]. No que se refere à reprodutibilidade, o NIST promove consistência por meio de etapas e critérios bem definidos [2], enquanto DevSecOps alcança repetibilidade pela execução automatizada de *pipelines* [6]. Contudo, a qualidade dos resultados automatizados depende da calibração das ferramentas, especialmente para controle de falsos positivos [4]. Em aplicações *greenfield*, DevSecOps favorece prevenção precoce e redução de custos de correção [7, 6]. Já em contextos *brownfield*, abordagens estruturadas inspiradas em NIST tendem a oferecer melhor definição de escopo e priorização por risco [5, 2].

As abordagens não são excludentes, mas complementares: o NIST fornece estrutura conceitual e alinhamento à gestão de risco [5], enquanto DevSecOps operacionaliza verificação contínua [6]. Sua integração pode maximizar indicadores como confirmação de achados, cobertura e reprodutibilidade [8, 9]. A escolha ou combinação entre elas deve considerar maturidade organizacional, requisitos regulatórios e criticidade dos dados [5], sendo a efetividade condicionada à adoção de métricas monitoráveis e decisões baseadas em evidência [8].

Este trabalho apresenta uma análise conceitual fundamentada em princípios consolidados de gestão de segurança [8, 5], sem validação empírica em estudo de caso. Assim, os efeitos discutidos devem ser interpretados como tendências esperadas, influenciadas por fatores como maturidade da equipe, configuração de ferramentas e características do ambiente. Como limitações, destaca-se a possível variação dos resultados em diferentes contextos organizacionais (validade externa) e a impossibilidade de estabelecer causalidade direta entre abordagem e melhoria dos indicadores (validade interna). Ainda assim, o modelo proposto permanece útil como estrutura replicável para avaliações futuras [8, 9].

V. Conclusão

Este artigo discutiu o problema de reprodutibilidade e qualidade de achados na análise de vulnerabilidades em aplicações que utilizam bancos de dados, destacando o impacto operacional dos falsos positivos e a necessidade de critérios objetivos de evidência e validação. A partir disso, foram apresentadas métricas e indicadores para avaliar efetividade do processo e foi realizada uma comparação entre abordagens orientadas por NIST e por DevSecOps. Observa-se que NIST tende a contribuir com estrutura, rastreabilidade e padronização de evidências, enquanto DevSecOps tende a contribuir com automação, reteste contínuo e redução de regressões ao longo do ciclo de entrega.

Como conclusão prática, a abordagem mais adequada depende do cenário: aplicações novas tendem a se beneficiar fortemente de DevSecOps quando combinado com critérios claros de evidência; aplicações legadas tendem a se beneficiar de estruturação por NIST para escopo e priorização, com posterior sustentação por automação onde possível. Como trabalhos futuros, propõe-se validar empiricamente o modelo por meio de um estudo de caso em ambiente controlado (com bases relacionais e não relacionais), coletando métricas como taxa de confirmação, falsos positivos, tempo de triagem, cobertura e reincidência, além de avaliar reprodutibilidade entre avaliadores.

VI. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] OWASP Foundation. *Web Security Testing Guide (WSTG) v4.1 — Introduction*. https://owasp.org/www-project-web-security-testing-guide/. Version 4.1. 2020.

[2] Karen Scarfone et al. *Technical Guide to Information Security Testing and Assessment*. Rel. técn. Special Publication 800-115. Gaithersburg, MD, USA: National Institute of Standards e Technology, set. de 2008. DOI: 10.6028/NIST.SP.800-115.

[3] National Institute of Standards and Technology (NIST). *False Positive*. https://csrc.nist.gov/glossary. NIST Computer Security Resource Center (CSRC) Glossary. n.d.

[4] Farhad Cheirdari e George Karabatis. "Analyzing False Positive Source Code Vulnerabilities Using Static Analysis Tools". Em: *Proceedings of the 2018 IEEE International Conference on Big Data (BigData)*. IEEE, 2018, pp. 4782–4788. DOI: 10.1109/BigData.2018.8621911.

[5] National Institute of Standards and Technology (NIST). *The NIST Cybersecurity Framework (CSF) 2.0*. Rel. técn. CSWP 29. Gaithersburg, MD, USA: National Institute of Standards e Technology, fev. de 2024. DOI: 10.6028/NIST.CSWP.29.

[6] Microsoft Security. *What Is DevSecOps? Definition and Best Practices*. https://www.microsoft.com/security/blog/. n.d.

[7] Murugiah Souppaya, Karen Scarfone e David Dodson. *Secure Software Development Framework (SSDF) Version 1.1: Recommendations for Mitigating the Risk of Software Vulnerabilities*. Rel. técn. Special Publication 800-218. Gaithersburg, MD, USA: National Institute of Standards e Technology, fev. de 2022. DOI: 10.6028/NIST.SP.800-218.

[8] ISO/IEC. *ISO/IEC 27004:2016 Information technology — Security techniques — Information security management — Monitoring, measurement, analysis and evaluation*. International Organization for Standardization. 2016.

[9] A. P. Aldya, S. Sutikno e Y. Rosmansyah. "Measuring effectiveness of control of information security management system based on SNI ISO/IEC 27004: 2013 standard". Em: *IOP Conference Series: Materials Science and Engineering* 550 (2019), p. 012020. DOI: 10.1088/1757-899X/550/1/012020.

Caderno Instrutivo sobre Redes PON

Aryel Lopes de Almeida Lima, Bruno Rezende Higino de Cuba,
Felipe Penha Pereira Duarte, Evandro César Vilas Boas, Guilherme Pedro Aquino
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
aryellopes@get.inatel.br, bruno.cuba@get.inatel.br, felipe.duarte@ges.inatel.br, evandro.cesar@inatel.br,
guilhermeaquino@inatel.br

***Abstract*—This article presents a practical guide for configuring GPON (Gigabit Passive Optical Network) networks, enabling students at the National Institute of Telecommunications (Inatel) to apply their knowledge in a laboratory environment. Using the Huawei MA5600T OLT (Optical Line Terminal), the document provides detailed instructions for network configuration, including the use of splitters and ONUs (Optical Network Units), simulating a scenario close to the reality of Brazilian ISPs.**

***Index Terms*—GPON, OLT MA5600T, passive optical networks, network configuration.**

***Resumo*—Este artigo apresenta um guia prático para configuração de redes GPON (*Gigabit Passive Optical Network*), permitindo que alunos do Instituto Nacional de Telecomunicações (Inatel) pratiquem seus conhecimentos em um ambiente laboratorial. Utilizando a OLT (*Optical Line Terminal*) MA5600T da Huawei, o documento fornece instruções detalhadas para configuração de rede, incluindo o uso de *splitters* e ONUs (*Optical Network Units*), simulando um cenário próximo à realidade de ISPs brasileiros.**

***Palavras chave*—GPON, OLT MA5600T, redes ópticas passivas, configuração de rede.**

I. Introdução

O avanço acelerado da sociedade digital exige redes de comunicação capazes de suportar altas taxas de transmissão de dados, com baixa latência e elevada confiabilidade. Diante desse cenário, as redes ópticas surgem como uma alternativa superior às tecnologias baseadas em meios cabeados baseados em fios de cobre, oferecendo maior desempenho, maior alcance e redução no consumo energético. Entre as diversas soluções ópticas, a tecnologia GPON (*Gigabit-capable Passive Optical Network*) destaca-se como uma das mais difundidas no mundo, especialmente em projetos de acesso FTTH (*Fiber to the Home*).

O GPON é um padrão de rede óptica passiva que adota uma topologia ponto-multiponto, permitindo que uma única fibra seja compartilhada entre diversos usuários por meio de dispositivos passivos, como *splitters*. Isso proporciona economia de infraestrutura e escalabilidade eficiente, ideal para atender desde pequenos bairros até grandes centros urbanos. O padrão GPON, definido pela série de recomendações ITU-T G.984.x, oferece taxas de transmissão de até 2,5 Gbps no sentido descendente (*downstream*) e 1,25 Gbps no sentido ascendente (*upstream*), suportando serviços como Internet de alta velocidade, voz sobre IP (VoIP) e IPTV, com recursos avançados de qualidade de serviço (*QoS*) e gerenciamento remoto.

No contexto brasileiro, a adoção massiva do GPON por provedores regionais e grandes operadoras foi impulsionada pela demanda por banda larga em áreas residenciais, comerciais e rurais. Essa tecnologia tornou-se essencial para a infraestrutura de telecomunicações do país, permitindo a inclusão digital e dando suporte à transformação digital de empresas e instituições públicas. Logo, torna-se fundamental que os alunos de Engenharia de Telecomunicações tenham contato direto com a tecnologia GPON durante sua formação. Para isso, foi estruturado, no Inatel, um ambiente laboratorial que simula uma rede óptica real, utilizando a OLT MA5600T da Huawei e os demais elementos que compõem uma rede GPON completa.

Este artigo tem como objetivo apresentar um guia prático e detalhado para a configuração da rede GPON no laboratório do Inatel. Inicialmente, são abordados os conceitos teóricos essenciais para a compreensão da tecnologia. Em seguida, descreve-se a topologia implementada no laboratório e, por fim, apresenta-se um passo a passo da configuração, visando capacitar alunos e professores a operar e compreender, na prática, os principais recursos de uma rede GPON.

II. Fundamentos em redes GPON

As redes GPON oferecem conectividade de alta velocidade com eficiência energética e menor custo, especialmente em cenários que exigem o atendimento de múltiplos usuários por meio de infraestrutura compartilhada e passiva. A Figura 1 apresenta a topologia básica de uma rede GPON, utilizada como referência neste artigo para ilustrar seus principais elementos no laboratório do Inatel. Nessa arquitetura, a OLT (*Optical Line Terminal*) atua como ponto central, conectando-se via fibra óptica às ONUs/ONTs (*Optical Network Units/Terminals*) instaladas nos usuários finais. Entre a OLT e as ONUs, o *splitter* óptico divide o sinal para diversos usuários sem o uso de equipamentos ativos intermediários, caracterizando a rede como passiva.

A transmissão na rede GPON utiliza múltiplos protocolos e mecanismos para garantir eficiência e qualidade de serviço. No *downstream* (OLT para ONU), os dados são enviados de forma contínua por meio de TDM (*Time Division Multiplexing*). Já no *upstream* (ONU para OLT), o acesso ao meio é coordenado por meio de janelas de tempo alocadas dinamicamente com base no protocolo DBA (*Dynamic Bandwidth Allocation*) [7]. Esse protocolo permite que a OLT gerencie o tráfego de forma otimizada, alocando largura de banda conforme a demanda de cada ONU. A multiplexação é feita sobre diferentes comprimentos de onda na mesma fibra: tipicamente 1490 nm para downstream e 1310 nm para upstream, segundo a recomendação ITU-T

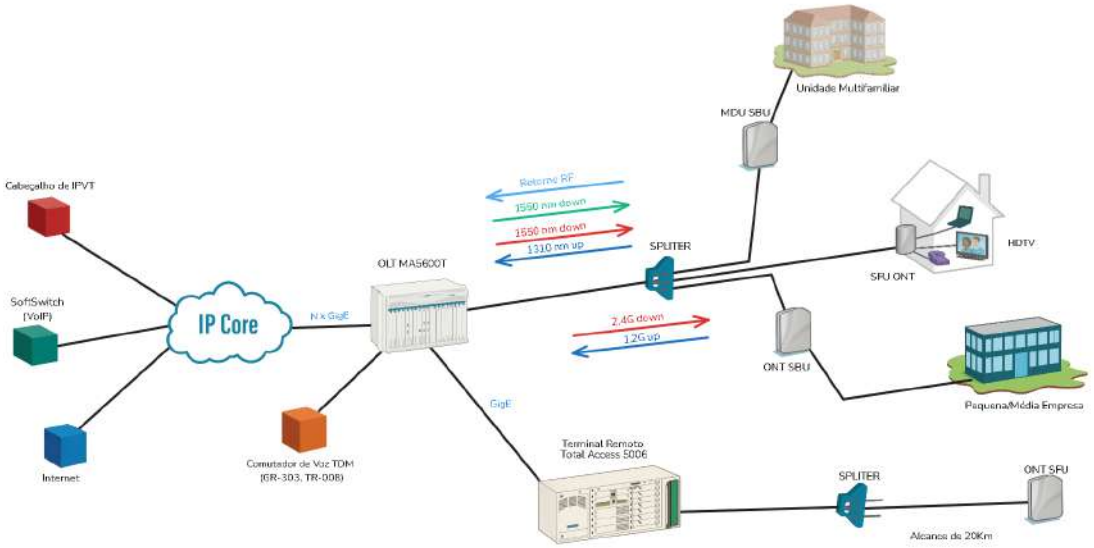


Fig. 1. Topologias utilizadas para configurações de redes GPON.

G.984.5 [8]. Esse mecanismo é chamado de WDM (*Wavelength Division Multiplexing*) e permite comunicação bidirecional simultânea.

Cada ONU pode ser configurada e gerenciada remotamente via protocolo OMCI (*ONT Management and Control Interface*) [9], conforme definido na recomendação ITU-T G.988. A partir da OLT, é possível associar perfis de linha e de serviço, definir políticas de QoS (*Quality of Service*), VLANs e parâmetros de tráfego para cada ONU, sem intervenção física no local. A comunicação lógica entre a OLT e cada ONU é baseada em três conceitos principais:

- **ONT-ID**: identificador único da ONU dentro da interface da OLT.
- **T-CONT (Transmission Container)**: unidade lógica de alocação de banda associada a perfis de QoS.
- **GEM Port (GPON Encapsulation Method Port)**: canal lógico de encapsulamento para serviços específicos.

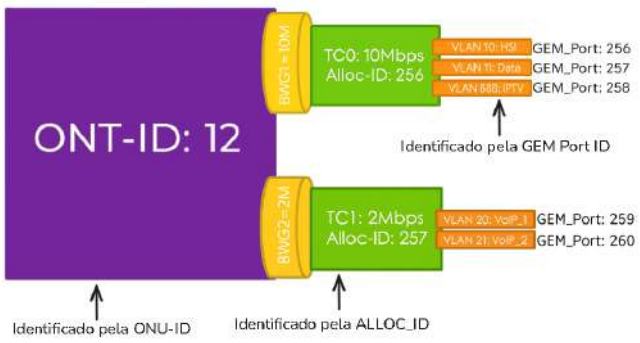


Fig. 2. Relação entre ONT-ID, T-CONT e GEM Port na estrutura lógica GPON.

Esses elementos estão representados na Figura 2. Pode-se observar diferentes serviços mapeados para GEM Ports distintas (como Internet e VoIP), cada um com seu respectivo controle de banda via T-CONT. Essa granularidade permite

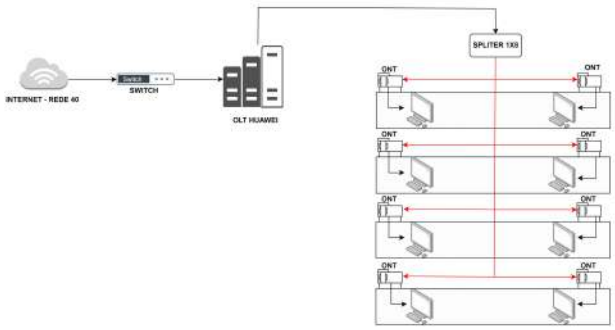


Fig. 3. Topologia prática implementada no laboratório do Inatel com OLT Huawei, splitter 1x8 e ONTs.

que a rede GPON ofereça serviços múltiplos com garantias de qualidade, priorização de tráfego e controle refinado sobre o uso da banda. Essa base teórica é essencial para a compreensão das configurações práticas que serão realizadas no laboratório do Inatel, conforme será detalhado na Seção IV.

III. Topologia montada no laboratório do Inatel

A Figura 3 apresenta a topologia real montada no laboratório do Inatel, em que é possível observar a conexão entre a OLT Huawei, o splitter 1x8 e as ONTs distribuídas em diferentes estações, todas conectadas a dispositivos como computadores, simulando uma rede GPON de acesso banda larga. A seguir, os comandos e procedimentos necessários para configurar a rede GPON no equipamento Huawei MA5600T.

A. OLT MA5600T

A OLT, denominada MA5600T, da Huawei, é um equipamento central em uma rede GPON e responsável por gerenciar e controlar a transmissão de dados às ONUs por meio da rede passiva de fibra óptica. A OLT MA5600T é projetada para suportar uma grande quantidade de ONUs, conectando-as a partir de um único ponto central. Cada ONU se comunica com



Fig. 4. Placas instaladas na OLT MA5600T utilizada no laboratório.

a OLT através de uma rede óptica passiva, onde a largura de banda é dividida dinamicamente entre os diferentes usuários, garantindo que todos recebam a quantidade adequada de recursos, conforme suas necessidades. Dependendo da configuração, a MA5600T pode suportar até 16 portas GPON, com cada porta com capacidade para atender até 128 ONUs, totalizando um limite teórico de até 2048 ONUs conectadas à mesma OLT.

No contexto da rede GPON, a OLT MA5600T implementa uma série de funções essenciais para a operação da rede:

- **Controle de Largura de Banda:** A OLT gerencia a alocação dinâmica de largura de banda usando o protocolo *DBA (Dynamic Bandwidth Allocation)*, que ajusta a quantidade de banda fornecida a cada ONU com base em suas demandas de tráfego, assegurando uma utilização eficiente dos recursos disponíveis.
- **Gerenciamento de Tráfego:** A OLT realiza o gerenciamento de tráfego, incluindo controle de congestionamento, priorização de tráfego e implementação de políticas de *QoS (Quality of Service)*, para garantir que serviços críticos, como voz e vídeo, recebam prioridade sobre tráfegos menos sensíveis.
- **Segurança e Proteção de Dados:** A OLT MA5600T inclui recursos de segurança avançados para proteger os dados transmitidos, evitando acessos não autorizados e ataques à rede.
- **Configuração e Provisionamento de ONUs:** A OLT utiliza o protocolo *OMCI (ONT Management and Control Interface)* para o provisionamento e a configuração remota das ONUs conectadas. Isso permite que a OLT controle e monitore o desempenho das ONUs, ajustando as configurações conforme necessário para otimizar a operação da rede.
- **Monitoramento e Diagnóstico:** A OLT MA5600T possui capacidades de monitoramento em tempo real, permitindo a detecção de falhas e a realização de diagnósticos avançados. Isso inclui o monitoramento do desempenho dos enlaces ópticos, identificação de atenuações anormais e verificação da integridade do sinal.

No laboratório do Inatel, foi implementada uma OLT como mostrada na Figura 4. É possível notar que as principais placas da OLT MA5600T em uso no laboratório do Inatel são (da esquerda para a direita): i) PRTE (*Power Rectifier Terminal Equipment*), responsável pela alimentação em -48 Vdc da OLT; ii) GPBD (*Gigabit-capable Passive Optical Network Board*), que contém as portas GPON; iii) SCUN (*Service Control Unit*), que atua como unidade de controle e gerenciamento, fornecendo as interfaces *CON (console)* e *ETH (Ethernet)* para gerência via *IP*; e, por fim, iv) GICF (*Gigabit Interface Card – Fiber*), que é responsável pelas interfaces *Ethernet* do tipo *GE* e será utilizada como porta de *uplink*, por onde os dados trafegarão entre a OLT e a rede de dados.



Fig. 5. ONU.

B. Optical Network Units (ONUs)

As ONUs são dispositivos instalados nos locais dos usuários finais, como residências ou empresas, e são responsáveis por converter o sinal óptico recebido da OLT em sinais elétricos que podem ser utilizados pelos dispositivos do cliente, como computadores, telefones *VoIP*, e *set-top boxes* de TV. As ONUs também desempenham um papel fundamental na comunicação entre os dispositivos do usuário com a OLT, enviando dados e solicitações de serviço conforme necessário. A ONU utilizada no laboratório do Inatel é representada na Figura 5. O modelo da ONU escolhida é EchoLife HG8010H, da fabricante Huawei. Entre as principais características desse modelo de ONU, destacam-se a existência de: i) uma porta óptica GPON, conector *SC/APC (Subscriber Connector / Angled Physical Contact)*, na cor azul; ii) uma porta *Gigabit Ethernet (GE)* para conexão em rede local; iii) LEDs indicadores para status de funcionamento; e iv) alimentação externa por fonte 12 Vdc.

C. Rede óptica passiva

Basicamente, na rede óptica passiva existem dois elementos: a fibra óptica e o *splitter* óptico. A fibra óptica usada na rede PON é do modelo padrão ITU-T G.652, denominada *Standard Single Mode Fiber* [11], operando em 1310 nm para a janela de *upstream* e 1490 nm para *downstream*.

Os *splitters* ópticos são dispositivos passivos que dividem o sinal óptico de uma única fibra proveniente da OLT em múltiplos sinais que são enviados para diferentes ONUs. Um exemplo de *splitter* pode ser visto na Figura 6. Esses *splitters* permitem que uma única OLT possa servir várias ONUs, aumentando a eficiência do uso da fibra óptica. Os *splitters* são tipicamente configurados em razões de divisão como 1:2, 1:4, 1:8, 1:16, ou 1:32, dependendo da densidade de clientes que precisam ser atendidos e da distância entre a OLT e as ONUs. A escolha da razão de divisão influencia diretamente a potência do sinal e a qualidade da conexão em cada ONU.

IV. Configuração da rede GPON no laboratório do Inatel

Nesta seção, serão mostrados todos os passos para a configuração da rede GPON que foi montada no laboratório do Inatel. Todas as interfaces de configurações, gráficas e *CLIs*

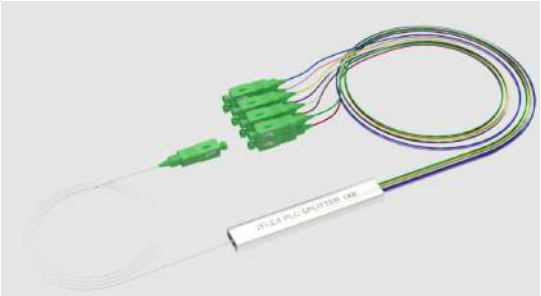


Fig. 6. Splitter.

(Command Line Interface) serão detalhadas, permitindo que professores e alunos possam configurar, operar e gerenciar a rede.

A. Acessando a OLT

Para acessar a OLT MA5600T, é necessário instalar o programa *MobaXterm* para fazer o acesso via terminal serial. A OLT MA5600T possui interface de configuração gráfica. Contudo, o Inatel ainda não possui licença do software da Huawei que permite a configuração via interface gráfica. Por esse motivo, todas as configurações da OLT foram feitas usando terminal serial. Com o *MobaXterm* instalado, conecte o cabo *USB/Ethernet* na porta *CON* da placa *SCUN* da OLT, como ilustrado na Figura 7. A extremidade *USB* do cabo deve ser conectada no computador que possui o *MobaXterm* instalado.

Após a conexão feita no passo anterior, no computador, abra o *MobaXterm*, clique no ícone `Session` e, depois, selecione a opção `Serial`, como ilustrado na Figura 8. A porta *COM* será ativada automaticamente e o computador já terá acesso à porta console da OLT e, portanto, terá acesso também à CLI (*Command Line Interface*), conforme pode ser visto na Figura 9. Nesse ponto, o usuário deverá entrar com as credenciais de acesso, que são: Login: `root` e Senha: `admin`. Nesse ponto o usuário deve fazer a configuração e a ativação das placas da OLT. Para isso, no prompt `MA5600T>` digite o comando `enable`, para que seja possível habilitar o modo privilegiado. Após esse comando, perceba que o prompt muda para `MA5600T#`.

Em seguida, para exibir as placas conectadas na OLT e seus estados, digite:

```
MA5600T# display board 0
```

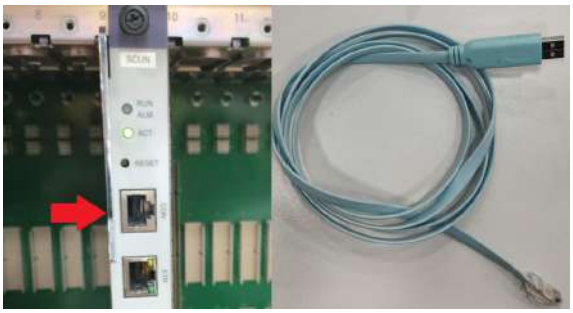


Fig. 7. Porta console da OLT e o cabo USB/Ethernet usado para configuração da OLT

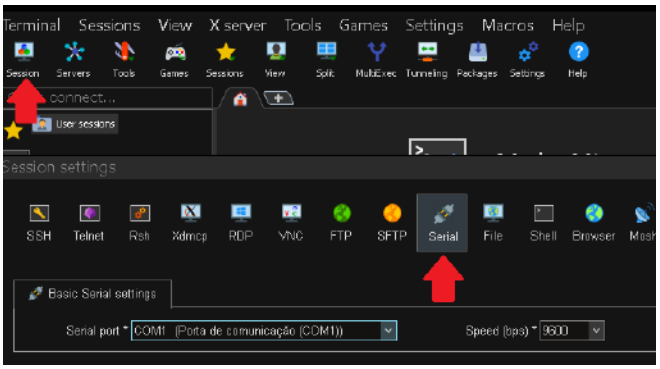


Fig. 8. Acesso via MobaXterm



Fig. 9. CLI da OLT MA5600T

Observe que a resposta desse comando está ilustrada na Figura 10. Caso algumas placas estejam com *status* `Auto_Find`, isso indica que ainda não foram confirmadas e que, portanto, os próximos três comandos serão necessários:

```
MA5600T# config
MA5600T(config)# board confirm 0/4
MA5600T(config)# board confirm 0/20
```

Após os comandos de confirmação, as placas devem aparecer com *status* idênticos aos relacionados na Figura 10. Depois das verificações feitas nos passos anteriores, pode-se iniciar o processo de configuração da conexão entre OLT e ONU e, também, a configuração dos serviços que irão ser fornecidos nessa conexão. Para tal, deve-se seguir o fluxo de configuração ilustrado na Figura 11. De acordo com a Figura 11, são necessários 7 passos para a ativação do serviço *PON* para uma ONU específica. A seguir, estão listados todos os comandos relativos a cada passo.

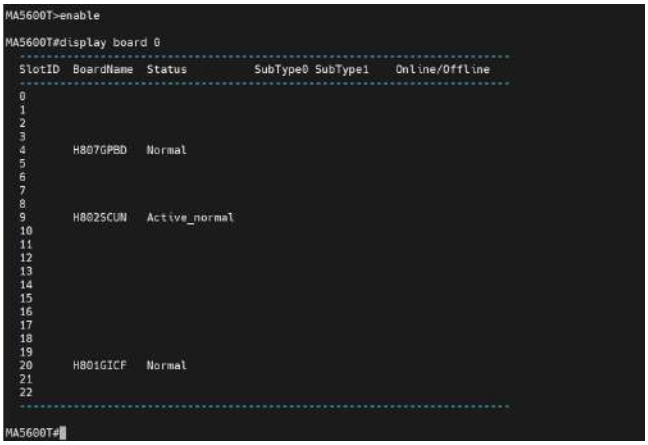


Fig. 10. Board display

B. Criação da VLAN de Uplink

No modo de configuração, `MA5600T(config)#`, digite os comandos listados a seguir para criação e associação de uma *VLAN (Virtual Local Area Network)* com uma porta de *uplink*. No exemplo, será criada uma VLAN 10 e essa VLAN será associada à porta *ETH* 0 da placa de *uplink* 0/20.

```
MA5600T(config)#vlan 10 smart
MA5600T(config)#port vlan 10 0/20 0
```

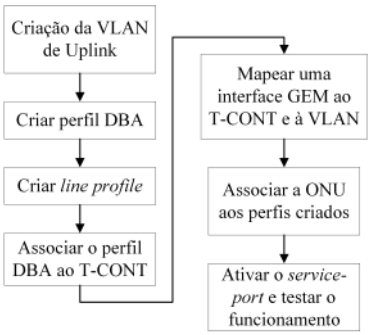


Fig. 11. Fluxograma das configurações para o serviço PON

C. Criação do perfil DBA

Para criação do perfil DBA, será necessário digitar o comando a seguir na *CLI*. O comando mostrado no exemplo possui a seguinte estrutura `dba-profile add profile-id X type4 max PIR`. Ou seja, designa-se uma identificação para o perfil e define-se a *PIR (Peak Information Rate)* que o controle DBA irá oferecer para a ONU que se associar a esse perfil. No exemplo em questão, definiu-se o perfil 10 com *PIR* igual a 1 *Gbps (gigabit por segundo)*.

```
MA5600T(config)#dba-profile add profile-id 10
    type4 max 1000000
```

D. Criação do line profile

Para a criação do *line profile* será necessário o comando a seguir. Deve-se criar um nome qualquer para o *line profile* e associá-lo a um perfil DBA. No exemplo a seguir, é criado um *line profile* com nome `1-svc-vlan10` e esse é associado ao perfil DBA 10 (`profile-id 10`), criado no comando anterior.

```
MA5600T(config)#ont-lineprofile gpon profile-id
    10 profile-name 1-svc-vlan10
```

E. Associação do perfil DBA ao T-CONT

Para a criação da estrutura lógica mostrada na Figura 2, é necessário que se crie um T-CONT (*Transmission Container*) e associe-o ao perfil DBA criado anteriormente. Para isso, usa-se o comando a seguir. Ou seja, dentro do *line profile* é necessário fazer a associação de um T-CONT qualquer com o perfil DBA criado. No exemplo, criou-se o T-CONT 1 e o associou ao perfil DBA 10.

```
MA5600T(config-gpon-lineprofile-10)#tcont 1 dba-
    profile-id 10
```

F. Mapeamento da interface GEM ao T-CONT e à VLAN de uplink

A última parte da criação da estrutura lógica mostrada na Figura 2 é feita fazendo a associação do método de encapsulamento GEM (*GPON Encapsulation Method*) ao T-CONT criado e, depois, associando toda essa estrutura à VLAN de *uplink*. Para fazer isso, é necessário o comando mostrado a seguir. É importante notar que, no exemplo, cria-se um GEM de número 10 e o associa ao T-CONT 1, criado no passo anterior e, logo em seguida, faz-se toda essa estrutura associar à VLAN 10. Por fim, é necessário encerrar o processo com os comandos `commit` e `quit`.

```
MA5600T(config-gpon-lineprofile-10)#gem add 10
    eth tcont 1
MA5600T(config-gpon-lineprofile-10)#gem mapping
    10 1 vlan 10
MA5600T(config-gpon-lineprofile-10)#commit
MA5600T(config-gpon-lineprofile-10)#quit
```

G. Associação da ONU aos perfis criados

No modo de configuração, deve-se acessar a interface GPON que foi usada para conectar a OLT à ONU e habilitar o *autofind* na porta *PON (Passive Optical Network)* para encontrar todas as ONUs conectadas a essa porta. Logo em seguida, deve-se verificar as informações das ONUs que foram encontradas. As três linhas de código são listadas abaixo e o resultado da consulta é ilustrado na Figura 12.

```
MA5600T(config)#interface gpon 0/4
MA5600T(config-if-gpon-0/4)#port 1 ont-auto-find
    enable
MA5600T(config-if-gpon-0/4)#display ont autofind
    1
```

As ONUs listadas no passo anterior já estão prontas para serem provisionadas. Sendo assim, é necessário confirmar a ONU, autenticando-a pelo seu número de série e associando-a a perfis de linha e serviço e definir o protocolo OMCI (*ONT Management and Control Interface*) para gerenciar a ONU. Para isso, precisa-se entrar com o código a seguir na *CLI*.

```
MA5600T(config-if-gpon-0/4)#ont confirm 1 sn-
    auth 48575443A922769D omci ont-lineprofile-
    id 10 ont-srvprofile-id 0 desc "vlan10"
```

Para verificar se a ONU foi autenticada, é necessário colocar o código a seguir na *CLI*. Se a ONU foi autenticada, o resultado mostrado na Figura 13 deve aparecer.

```
MA5600T(config-if-gpon-0/4)#display ont info 0
    all
```

```
MA5600T(config-if-gpon-0/4)#display ont autofind 1
-----------------------------------------------------------------------------
Number              : 1
F/S/P               : 0/4/1
Ont SN              : 48575443A922769D
Password            : 0x69323131333833376E65(i2113837ne)
Loid                :
Checkcode           :
VendorID            : HWTC
Ont Version         : 635.A
Ont SoftwareVersion : V3R015C10S150
Ont EquipmentID     : EG8010H
Ont autofind time   : 2024-08-16 15:21:24+08:00
-----------------------------------------------------------------------------
```

Fig. 12. Informações das ONUs encontradas pelo *autofind*

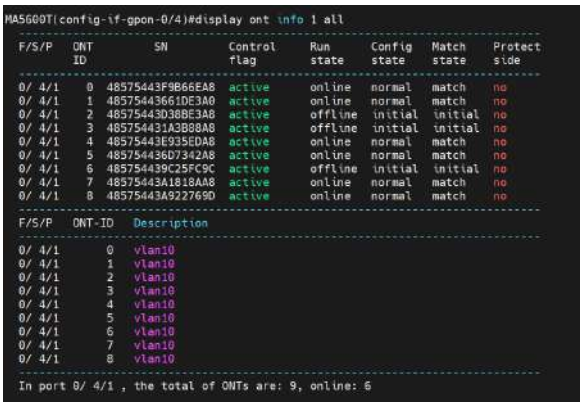


Fig. 13. ONT provisionadas

H. Ativar o service-port na ONT e testar o seu funcionamento

O último passo para que a ONT passe a ter acesso à rede é mostrado nesta subsessão. Primeiramente, é necessário associar a VLAN 10 às portas *Ethernet* da ONU para permitir o encaminhamento do tráfego pelo *uplink* da OLT. O comando `MA5600T(config-if-gpon-0/4)#ont port native-vlan 1 X eth 1 vlan 10` deve ser usado para tal propósito. Deve-se observar que o valor colocado na variável `X` do comando é referente ao **ID** da ONT, que pode ser consultado no resultado, Figura 13. No caso do exemplo abaixo, foi usada a ONT de **ID** igual a 0.

```
MA5600T(config-if-gpon-0/4)#ont port native-vlan
    1 0 eth 1 vlan 10
```

Por fim, deve-se criar um *service-port* (*porta de serviço*) e vinculá-lo à ONU que foi previamente provisionada. Isso permitirá que a ONU trafegue dados pelo *uplink* da OLT.

```
MA5600T(config)#service-port 9 vlan 10 gpon
    0/4/1 ont 0 gemport 10 multi-service user-
    vlan 10
```

Para testar se a rede está funcionando perfeitamente, conecte um computador à porta *Ethernet* da ONU, e este deverá conseguir acesso à Internet por meio da OLT.

Boas práticas e observações didáticas

Durante a configuração da rede GPON, é importante considerar aspectos como:

- **Segurança**: Em ambientes reais, deve-se alterar credenciais padrão da OLT e ONU e limitar o acesso via CLI e SNMP.
- **Escalabilidade**: Recomenda-se a separação de serviços por VLANs específicas e o uso de perfis bem definidos, facilitando a replicação do ambiente para múltiplos assinantes.
- **Diagnóstico e troubleshooting**: Comandos como `display ont info`, `display alarm` e `display ont optical-info` são úteis para identificar problemas como perda de sinal óptico, falhas de autenticação ou baixa potência de recepção.
- **Gerenciamento de banda**: A criação de perfis DBA adequados ao tipo de serviço (residencial, VoIP, IPTV) é essencial para garantir qualidade de serviço (*QoS*) e evitar congestionamentos.

Durante as configurações realizadas, os alunos puderam compreender, na prática, o provisionamento de uma rede GPON, observando a autenticação da ONU, a criação de perfis de serviço e o estabelecimento do tráfego de dados. A estrutura mostrou-se eficiente para fins didáticos, consolidando conceitos como identificação da ONU, mapeamento de VLANs, configuração de GEM Ports e gerenciamento de banda via DBA. Além disso, os comandos em CLI proporcionam uma visão direta do funcionamento interno da rede.

V. Conclusão

Este artigo apresenta um guia prático para a configuração de redes GPON utilizando a OLT MA5600T da Huawei, com o objetivo de apoiar professores e alunos do Inatel no aprimoramento de conhecimentos em redes de telecomunicações por meio de atividades em laboratório. O material oferece um passo a passo da configuração, contribuindo para a formação prática em comissionamento e gerenciamento de redes de acesso em banda larga, alinhada às demandas do mercado. A escolha da MA5600T deve-se à sua ampla utilização no cenário nacional e à interface CLI, que favorece o aprendizado técnico, embora o equipamento não suporte tecnologias mais recentes, como XG-PON. Como melhorias, sugerem-se a simulação de cenários de falhas, testes com diferentes perfis DBA e a integração de serviços de voz com telefones VoIP disponíveis no laboratório, ampliando a aderência do ambiente didático à realidade de ISPs e operadoras. O material pode ser aplicado em disciplinas como Redes de Telecomunicações I e áreas correlatas, contribuindo para a formação de futuros profissionais da área.

VI. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Referências

[1] Hood, Dave. *Gigabit-capable Passive Optical Networks*. John Wiley & Sons, 2009.
[2] Derahman, Mohd Noor. *Fiber to the Home (FTTH) Architecture: Design and Optimization*. Lap Lambert Academic Publishing GmbH KG, 2010.
[3] International Telecommunication Union (ITU-T). *Recommendation ITU-T G.984.1: Gigabit-capable passive optical networks (G-PON): General characteristics*. ITU-T, February 2022.
[4] Gast, Matthew S. *802.11 Wireless Networks: The Definitive Guide*. O'Reilly Media, 2012.
[5] Starr, Thomas, et al. *DSL: A Modern Perspective*. Prentice Hall PTR, 1999.
[6] Gallo, Ivaldo. *Redes de Acesso Banda Larga*. Editora Érica, 2010.
[7] Li, Ling, et al. "A comparison of dynamic bandwidth allocation for EPON, GPON, and next-generation TDM PON."*IEEE Communications Magazine*, vol. 47, no. 11, 2009, pp. 92-100.
[8] International Telecommunication Union (ITU-T). *Recommendation ITU-T G.984.5: Gigabit-capable passive optical networks (G-PON): Enhancement band*. ITU-T, February 2022.
[9] International Telecommunication Union (ITU-T). *Recommendation ITU-T G.988: Gigabit-capable passive optical networks (G-PON): Optical network unit management and control interface (OMCI) specification*. ITU-T, October 2022.
[10] Huawei Technologies Co., Ltd., "Huawei SmartAX MA5600T Series,"[Online]. Available: https://support.huawei.com/enterprise/en/doc/EDOC1100036801. [Accessed: Aug. 19, 2024].
[11] International Telecommunication Union (ITU-T). *Recommendation ITU-T G.652: Characteristics of a single-mode optical fibre and cable*. ITU-T, October 2023.

Implementação de uma Plataforma de Monitoramento Remoto para Sistemas Power-over-Fiber

Breno C. do Nascimento, Eduardo M. Bertozzi, Letícia C. de Souza e Francisco M. Portelinha Junior
Instituto Nacional de Telecomunicações - Inatel
breno.carvalho@ges.inatel.br, eduardo.bertozzi@ges.inatel.br, leticia.c@inatel.br, franciscoportelinha@inatel.br

***Abstract*—This work presents the development and experimental validation of an embedded remote monitoring platform applied to Power-over-Fiber (PoF) systems, a technology that enables power transmission through optical fibers. In this context, an autonomous architecture was designed and experimentally validated to measure, process, and supervise in real time the electrical parameters delivered by the photovoltaic receiver stage, including voltage, current, and power. The proposed solution integrates an *ESP32* microcontroller and a *MAX471* sensor for local data acquisition and processing, along with an embedded web interface providing dynamic visualization and configurable alert features. The PoF system was experimentally characterized, including the identification of the maximum power point of the photovoltaic converter and the evaluation of both optical-to-electrical conversion efficiency and overall system energy efficiency. The results demonstrate the technical feasibility of the proposed platform for energy performance optimization and continuous monitoring of remote loads, contributing to the broader applicability of PoF systems in industrial environments, embedded systems, and Industrial Internet of Things (IIoT) scenarios.**

***Index Terms*—Embedded systems, optical power transmission, power-over-fiber, remote monitoring.**

***Resumo*—Este trabalho apresenta o desenvolvimento e a validação experimental de uma plataforma embarcada de monitoramento remoto aplicada a sistemas *Power-over-Fiber* (PoF), tecnologia que realiza a transmissão de energia por meio de fibras ópticas. Nesse contexto, foi desenvolvida e validada experimentalmente uma arquitetura autônoma capaz de medir, processar e supervisionar em tempo real os parâmetros elétricos fornecidos pelo estágio receptor fotovoltaico, como tensão, corrente e potência. A solução integra um microcontrolador *ESP32* e um sensor *MAX471* para aquisição e processamento local dos dados, além de uma interface web embarcada com recursos de visualização dinâmica e alertas configuráveis. O sistema PoF foi caracterizado experimentalmente, incluindo a identificação do ponto de máxima potência do conversor fotovoltaico e a avaliação da eficiência de conversão óptico-elétrica e da eficiência energética global do sistema. Os resultados demonstram a viabilidade técnica da plataforma para otimização do desempenho energético e monitoramento contínuo de cargas remotas, contribuindo para ampliar a aplicabilidade de sistemas PoF em ambientes industriais, sistemas embarcados e cenários de Internet das Coisas Industrial (IIoT).**

***Palavras chave*—Monitoramento remoto, power-over-fiber, sistemas embarcados, transmissão óptica de energia.**

I. Introdução

A crescente demanda por soluções de sensoriamento e monitoramento em ambientes de alta tensão, isolados ou sujeitos à interferência eletromagnética (EMI) tem impulsionado o desenvolvimento de tecnologias capazes de aliar segurança, confiabilidade e eficiência energética [1]–[3]. Nesse contexto, a tecnologia *power-over-fiber* (PoF) surge como uma alternativa promissora para a transmissão segura e totalmente isolada de energia através de fibras ópticas, eliminando o uso de condutores metálicos e garantindo isolamento galvânico total entre a fonte de alimentação e o dispositivo energizado [4], [5]. Embora o PoF não envolva a transmissão direta de dados, sua integração é particularmente relevante em sistemas que combinam sensores eletrônicos, atuadores ou módulos de aquisição que necessitam de alimentação elétrica para operar. Diversos sensores eletrônicos utilizados nesses ambientes exigem potência para circuitos de condicionamento de sinal, processamento e comunicação. Para esses casos, o PoF oferece uma alternativa viável e segura. Além disso, essa tecnologia apresenta vantagens expressivas em relação aos métodos convencionais de alimentação, como imunidade à EMI, baixa perda de transmissão, peso reduzido e elevada flexibilidade de instalação [1], [2]. Tais características tornam o PoF particularmente adequado para aplicações em ambientes industriais sensíveis, sistemas embarcados, sensores remotos que dependem de eletrônica ativa e arquiteturas avançadas de monitoramento. Além disso, sua capacidade de fornecer energia de forma totalmente isolada permite sua integração com tecnologias emergentes como a *Industrial Internet of Things* (IIoT) e redes de comunicação móvel, incluindo quinta geração (5G) e sexta geração (6G), nas quais módulos eletrônicos de baixa potência podem ser energizados de maneira segura e imunes a ruídos [6].

Sistemas PoF têm sido amplamente explorados na literatura em aplicações de sensoriamento remoto e monitoramento em ambientes de alta tensão. Essa tecnologia, de forma simples, baseia-se na transmissão de luz gerada por um *high-power laser diode* (HPLD), que é então convertida em energia elétrica por um *photovoltaic power converter* (PPC), permitindo o fornecimento seguro e totalmente isolado de potência por fibra óptica. Matsuura apresentou uma revisão abrangente sobre os avanços tecnológicos em PoF, discutindo a evolução dos PPCs e dos HPLDs utilizados na transmissão óptica de energia [2]. Já Kim e Park investigaram as tendências de patentes e o crescimento industrial relacionado à tecnologia PoF, destacando o aumento do interesse em aplicações seguras e eletricamente isoladas [4]. Em [5], a implementação de nós sensores sem fio alimentados opticamente, voltados para ambientes industriais, demonstrou a viabilidade prática do PoF em sistemas embarcados e redes de monitoramento. Além dessas análises de

caráter conceitual e industrial, diversos estudos apresentaram resultados experimentais que forneceram uma visão clara dos limites de potência e eficiência alcançáveis em sistemas PoF. Os estudos de Matsuura *et al*. demonstraram a transmissão de mais de 40 W de potência elétrica por fibra óptica em conjunto com comunicação de dados, representando um dos maiores valores já reportados na literatura [7]. Já Fafard e Masson apresentaram um enlace PoF capaz de entregar cerca de 1 W após 5 km de fibra, com eficiência de conversão próxima de 49%, além de demonstrações superiores a 2 W em enlaces curtos, inclusive em condições criogênicas [8]. Esses resultados ofereceram uma noção prática do desempenho energético possível em diferentes cenários de aplicação.

Apesar das vantagens apresentadas, a eficiência global de um sistema PoF depende fortemente das condições de propagação óptica e do desempenho do PPC, responsável por converter a potência óptica recebida em energia elétrica [5], [9]. Dessa forma, o monitoramento contínuo dos parâmetros elétricos do sistema é essencial para garantir um desempenho estável e uma operação confiável. A medição em tempo real de tensão, corrente e potência possibilita a identificação de variações anômalas, a prevenção de falhas e o ajuste dinâmico do ponto de operação. A utilização de sensores inteligentes e de um sistema embarcado de aquisição e controle permite a implementação de uma abordagem autônoma de monitoramento, capaz de reagir às variações de forma adaptativa e eficiente. Estudos anteriores, como os de López *et al*. [10] e Altuna *et al*. [11], exploraram o conceito de monitoramento inteligente em sistemas ópticos e elétricos, evidenciando seu potencial para aprimorar a eficiência e a confiabilidade de processos automatizados. Entretanto, essas abordagens geralmente apresentam alta complexidade de implementação, exigindo arquiteturas específicas, sensores dedicados e algoritmos de análise avançados, o que restringe sua aplicação a contextos muito nichados e de difícil replicação prática.

Este artigo propõe o desenvolvimento de um sistema de monitoramento inteligente aplicado a um enlace PoF, com capacidade de medição e análise de tensão, corrente e potência, para a otimização do desempenho energético. A proposta visa oferecer uma arquitetura embarcada projetada para operar de forma autônoma, com coleta e processamento local dos dados, permitindo a integração com plataformas de supervisão remota e análise em nuvem. Essa abordagem busca democratizar o uso de tecnologias PoF, tornando-as mais acessíveis e viáveis para cenários práticos que demandam monitoramento contínuo, eficiência energética e isolamento elétrico total. O manuscrito está estruturado em quatro seções. A Seção II descreve a arquitetura e a implementação do sistema proposto, contemplando a caracterização do enlace PoF, o desenvolvimento da plataforma embarcada de aquisição e processamento de dados, bem como a implementação da interface web e dos mecanismos de supervisão e proteção. A Seção III apresenta os resultados experimentais e a análise de desempenho do sistema. Por fim, as conclusões e os trabalhos futuros são apresentados na Seção IV.

II. Arquitetura e Implementação do Sistema de Monitoramento

Esta seção apresenta a arquitetura e a implementação do sistema de monitoramento proposto, abrangendo o enlace PoF, o estágio de conversão óptico-elétrica e a plataforma embarcada de aquisição, processamento e supervisão de dados. Inicialmente, são descritos os componentes do sistema PoF e os parâmetros adotados na implementação experimental. Em seguida, detalha-se o módulo de aquisição baseado no microcontrolador *ESP32* e no sensor de corrente e tensão *MAX471*, incluindo aspectos de calibração e condicionamento dos sinais. Por fim, são apresentados o firmware responsável pelo processamento das medições e a interface web embarcada destinada à visualização e supervisão remota das grandezas elétricas do sistema.

A. Sistema Power-over-Fiber (PoF)

De forma geral, um sistema PoF é constituído por três blocos fundamentais: a fonte óptica HPLD, a fibra óptica utilizada como meio de transmissão e o PPC empregado na etapa de conversão eletro-óptica final, como ilustrado na Figura 1 [12]. O HPLD é alimentado de modo a manter uma corrente constante através do diodo, garantindo operação estável e emissão contínua de potência luminosa. A radiação óptica é acoplada ao núcleo da fibra, propagando-se ao longo do enlace sem suscetibilidade a EMI. No módulo receptor, um PPC opera em regime fotovoltaico, isto é, sem polarização externa, convertendo diretamente a energia óptica recebida em energia elétrica por meio do campo interno do dispositivo. Dessa forma, a potência gerada pode ser utilizada para alimentar uma carga ou subsistema eletrônico da estação remota [2].

A Tabela I apresenta as principais especificações dos componentes empregados na implementação do sistema PoF. O comprimento de onda de 808 nm foi selecionado por apresentar elevada maturidade tecnológica em HPLD, permitindo operação estável em níveis de potência óptica da ordem de dezenas de watts. Além disso, o PPC empregado apresenta resposta espectral otimizada nessa faixa, possibilitando maior eficiência de conversão fotoelétrica. A utilização de fibra multimodo 62,5/125 μm justifica-se pela maior área de núcleo, que facilita o acoplamento de potência óptica proveniente do HPLD, reduzindo a densidade de potência e mitigando riscos de danos ópticos e efeitos térmicos no conector e na face da fibra. No módulo receptor, o PPC apresenta eficiência máxima de 60%. A tensão elétrica disponibilizada na condição nominal de operação é da ordem de 6 V, sendo posteriormente elevada para 12 V por meio de um conversor DC/DC do tipo *step-up*.

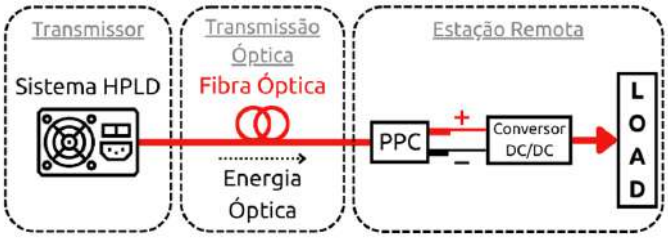


Fig. 1. Diagrama simplificado de um sistema PoF.

TABELA I
PRINCIPAIS PARÂMETROS DOS COMPONENTES DO SISTEMA POF.

Componente	Parâmetro	Valor
HPLD	Comprimento de onda	808 nm
HPLD	Potência óptica máxima	12 W
HPLD	Eficiência eletro-óptica	45 %
Fibra óptica	Tipo	Multimodo
Fibra óptica	Dimensões	62,5/125 μm
Fibra óptica	Atenuação	1 dB/km
PPC	Eficiência máxima	60 %
PPC	Tensão em circuito aberto	6 V
PPC	Máxima potência óptica de entrada	6 W

A avaliação da eficiência energética em sistemas PoF requer métricas que considerem não apenas o desempenho isolado dos componentes, mas o comportamento global do sistema. Nesse contexto, define-se a eficiência de energia global (EEG) como um indicador que quantifica a eficiência elétrica total do enlace, incorporando tanto a eficiência do transmissor óptico quanto a do conversor fotovoltaico na recepção, a qual pode ser descrita como [13]:

$$\mathrm{EEG} = \left(\frac{P_{\mathrm{PPC}}}{P_{\mathrm{HPLD}}} \right) 100\%, \tag{1}$$

onde P_{PPC} é a potência elétrica entregue pelo PPC e P_{HPLD} é a potência elétrica consumida pelo HPLD.

B. Sistema de Aquisição e Processamento de Dados

O sistema desenvolvido é constituído por uma plataforma embarcada baseada no microcontrolador *ESP32*, integrada ao sensor de corrente/tensão *MAX471*, além dos elementos principais do enlace PoF: o HPLD, a fibra óptica e o PPC. A Figura 2 apresenta a arquitetura geral do sistema, evidenciando o enlace óptico, o estágio de conversão óptico-elétrica e o circuito simplificado de aquisição de dados.

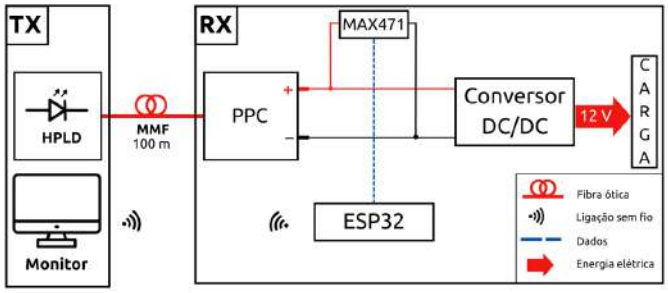


Fig. 2. Arquitetura do sistema embarcado de aquisição de dados do sistema PoF.

A partir da potência elétrica gerada pelo PPC, inicia-se o processo de aquisição e monitoramento das grandezas associadas ao desempenho do sistema. A energia elétrica disponível na saída é aplicada a uma carga resistiva, permitindo a avaliação controlada das condições operacionais. As grandezas de interesse (tensão e corrente) são monitoradas de forma contínua, possibilitando a caracterização dinâmica da potência entregue à carga. A aquisição dessas grandezas é realizada por meio do sensor *MAX471*, um circuito integrado dedicado à medição de corrente contínua baseado em amplificação diferencial de alta precisão. O dispositivo opera com detecção de corrente por meio de um resistor de *shunt* interno de baixo valor ôhmico, permitindo medições típicas na faixa de até aproximadamente 3 A, com tensão de operação entre 3 V e 36 V.

O sensor *MAX471* fornece como saída um sinal analógico proporcional à corrente medida, além da leitura direta da tensão do barramento monitorado. Os ensaios experimentais indicaram que, para a carga empregada, o fator de conversão mais representativo situou-se na faixa de aproximadamente 0,37–0,38 V/A. Esse valor foi, portanto, adotado como parâmetro de sensibilidade no *firmware*, visando aprimorar a calibração do sistema e aumentar a precisão das medições realizadas. O *MAX471* foi conectado em série com a carga, possibilitando a medição simultânea da corrente fornecida e da tensão no ponto de operação do sistema. Essa configuração permite que o próprio sensor seja alimentado pela linha monitorada, reduzindo a complexidade do circuito e eliminando a necessidade de fonte auxiliar dedicada.

O *firmware* do sistema, disponível em [14], foi desenvolvido em linguagem C/C++, utilizando a plataforma *Arduino IDE* configurada para compilação e gravação do código na placa baseada no microcontrolador *ESP32*. A implementação explorou os periféricos internos do dispositivo, incluindo as interfaces de *general-purpose input/output* (GPIO), o *analog-to-digital converter* (ADC) e o módulo de comunicação *universal asynchronous receiver/transmitter* (UART). Os sinais analógicos provenientes do *MAX471* são encaminhados às entradas do ADC interno do microcontrolador *ESP32*, cuja resolução é de 12 bits (valor máximo de 4095). Os valores amostrados são convertidos em valores de corrente, considerando a resolução do ADC e a tensão de referência de 3 V. Por fim, as grandezas calculadas são transmitidas por meio da interface serial UART para acompanhamento em tempo real, adotando-se um intervalo de atualização de 500 ms entre amostragens consecutivas. Para assegurar a confiabilidade das medições e minimizar interferências elétricas, todos os referenciais de terra do PPC, do sensor e do microcontrolador foram interligados, estabelecendo um plano de referência comum e eletricamente estável.

C. Interface Web e Funcionalidades de Supervisão

Após a etapa de aquisição e processamento das medições em tempo real, o *ESP32* passa a atuar como um servidor web embarcado, disponibilizando os dados por meio de sua interface Wi-Fi integrada. O servidor *hypertext transfer protocol* (HTTP) é executado localmente no microcontrolador, permitindo que qualquer dispositivo conectado à mesma rede acesse a plataforma de monitoramento por meio de um navegador convencional, sem necessidade de software adicional. A interface gráfica do usuário (*front-end*) foi desenvolvida utilizando as tecnologias *hyperText markup language* (HTML) 5, *cascading style sheets* (CSS) 3 e *JavaScript*, estruturada no modelo de *Single Page Application* (SPA). Essa arquitetura permite que a atualização dos dados ocorra de forma assíncrona e fluida, eliminando a necessidade de recarregamentos constantes da página pelo navegador.

A área central da plataforma é dedicada à visualização gráfica dinâmica das grandezas elétricas monitoradas: corrente, tensão e potência. O sistema oferece alta interatividade, permitindo ao operador selecionar quais variáveis deseja visualizar

simultaneamente ou isoladamente. O eixo temporal dispõe de taxas de atualização ajustáveis (0,5s, 1s, 2s ou 5s), além de funções para pausar a plotagem para análise estática e visualização em tela cheia. Ao interagir com o gráfico (*hovering*), o usuário obtém via *tooltips* o valor exato da medição e o respectivo carimbo de tempo (*timestamp*). A Figura 3 ilustra a disposição dos elementos gráficos na interface.

Paralelamente à visualização gráfica, o *firmware* processa estatísticas da sessão em tempo real. Painéis laterais informam a duração da monitoria e realizam a detecção automática de valores de pico (*Peak Hold*) para todas as grandezas. Indicadores de tendência visual sinalizam instantaneamente se a corrente está em ascensão, queda ou estabilizada. Além disso, a integridade da conexão com o microcontrolador é monitorada constantemente via sinal de *heartbeat*, alertando o operador em caso de perda de comunicação. Um recurso fundamental para a análise pós-operatória é a capacidade de exportação integral do histórico da sessão (incluindo Tempo, Corrente, Tensão e Potência) em formato *.CSV*.

No que tange à segurança do sistema PoF, foram implementados mecanismos de proteção configuráveis por software. O usuário pode definir limites máximos admissíveis de corrente diretamente na interface. Caso o valor medido ultrapasse o limiar estabelecido ou seja detectada corrente nula (indicando possível interrupção no circuito), o sistema aciona notificações multimodais, combinando *feedback* visual (intermitência da tela em cores amarela para aviso e vermelha para crítico) com alarmes sonoros distintos, conforme demonstrado nas Figuras 4 e 5. Todos os eventos de alerta são registrados automaticamente em um *log* de notificações auditável, acessível pelo menu de relatórios.

Adicionalmente, a plataforma incorpora funcionalidades vol-



Fig. 3. Interface web de monitoramento exibindo gráficos de corrente e tensão em tempo real.



Fig. 4. Alerta de advertência acionado quando é detectada corrente nula.



Fig. 5. Alerta crítico acionado quando a corrente ultrapassa o limite configurado.

tadas à segurança e ergonomia. O acesso às configurações administrativas é protegido por autenticação via *login*, restringindo alterações críticas a usuários autorizados. A interface também permite alternância dinâmica entre modos visuais *Light* e *Dark*, assegurando melhor adaptação a diferentes condições de iluminação. Parâmetros globais, como taxas de atualização e limites padrão, podem ser personalizados e restaurados aos valores originais de fábrica quando necessário.

III. Resultados Experimentais

Para avaliar o desempenho prático da solução proposta, foi estruturado um arranjo experimental integrando todos os elementos de *hardware* do sistema óptico e do módulo de aquisição de dados, operando de forma coordenada com a plataforma de monitoramento *web* desenvolvida. A Figura 6 apresenta a disposição completa do arranjo experimental utilizado nos ensaios.

Inicialmente, foi realizada a caracterização elétrica do PPC com o objetivo de determinar a potência máxima elétrica passível de extração sob diferentes níveis de potência óptica incidente. Para operação em 808 nm, o dispositivo apresenta comportamento fotovoltaico típico, com aumento aproximadamente proporcional da corrente de curto-circuito (I_{SC}) à medida que a potência óptica incidente é elevada, enquanto a tensão de circuito aberto (V_{OC}) apresenta variação moderada. Para potência óptica incidente de 1,5 W, o PPC apresenta $I_{SC} \approx 0{,}137$ A e $V_{OC} \approx 7{,}14$ V com *maximum power point* (MPP) em torno de 5,50 V e 0,137 A, resultando em potência elétrica máxima de aproximadamente 0,75 W e eficiência óptico-elétrica de cerca de 50%. Para 3 W, observa-se $I_{SC} \approx 0{,}28$ A e $V_{OC} \approx 7{,}16$ V, com MPP em 5,47 V e 0,28 A, entregando cerca de 1,53 W e eficiência de aproximadamente 51%. Já para 6 W de potência óptica incidente, $I_{SC} \approx 0{,}565$ A e $V_{OC} \approx 7{,}22$ V, com MPP em 5,46 V e 0,565 A, resultando em potência máxima de aproximadamente 3,09 W e eficiência óptico-elétrica de 51,5%. A partir do MPP, tornou-se

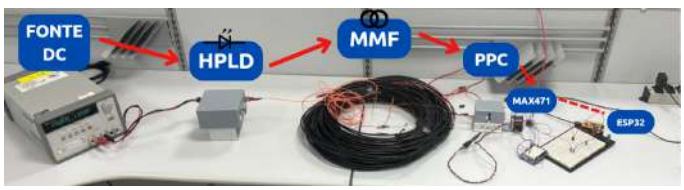


Fig. 6. Arranjo experimental do sistema PoF, destacando a integração entre o enlace óptico e a plataforma embarcada de aquisição de dados.

possível calcular a potência elétrica máxima entregue à carga e, consequentemente, estimar a EEG do sistema PoF por meio da relação entre a potência elétrica útil na carga e a potência elétrica consumida na fonte laser. Considerando o HPLD, que apresenta eficiência eletro-óptica de cerca de 45% e fornece aproximadamente 6,5 W de potência óptica mediante consumo de 14,4 W, e um enlace de 100 m de fibra multimodo com cerca de 6 W atingindo o PPC, a EEG do sistema PoF é estimada em aproximadamente 21,5%.

Para avaliar o comportamento da carga e verificar a precisão da plataforma de monitoramento, foi realizado um ensaio experimental utilizando uma carga resistiva nominal de 390 Ω, projetada para operação em 12 V. Nessa condição, a corrente elétrica teórica esperada é de aproximadamente 30,77 mA, enquanto a potência dissipada calculada é da ordem de 369,23 mW. A partir dos dados exportados pelo aplicativo, procedeu-se à análise das grandezas elétricas adquiridas em tempo real. Em regime estacionário, a plataforma registrou valores médios de tensão de 11,49 V e corrente de 30,23 mA, indicando estabilidade operacional. A corrente aferida apresentou elevada concordância com o valor teórico estimado, resultando em erro relativo de aproximadamente 1,75%, o que evidencia a consistência da calibração adotada e a confiabilidade do circuito de sensoriamento implementado. Esses resultados confirmam a adequação do sistema de monitoramento para a medição de grandezas elétricas em aplicações de baixa potência, demonstrando precisão compatível com as limitações inerentes aos componentes utilizados.Além disso, todos os sistemas de alarme funcionaram corretamente durante os ensaios, acionando as notificações visuais e sonoras de acordo com os limites configurados.

IV. Conclusões

Este trabalho apresentou o desenvolvimento e a implementação de uma plataforma embarcada para monitoramento inteligente de parâmetros elétricos em sistemas PoF, integrando o microcontrolador *ESP32* e o sensor *MAX471* em uma arquitetura autônoma capaz de medir, processar e disponibilizar, em tempo real, dados de tensão, corrente e potência elétrica por meio de uma interface web embarcada. Os resultados experimentais validaram a confiabilidade da solução, confirmando a precisão do sistema de aquisição e o desempenho energético do enlace óptico, bem como a identificação adequada do MPP do PPC e a EEG compatível com aplicações de alimentação remota de baixa potência. A plataforma demonstrou operação estável, segura e adequada a ambientes que exigem isolamento galvânico e imunidade a EMI, apresentando potencial para aplicações industriais, IIoT e futuras redes de comunicação. Como perspectivas futuras, destacam-se a implementação de técnicas de *Maximum Power Point Tracking* (MPPT) no firmware para otimização dinâmica da eficiência, a integração com serviços em nuvem para supervisão remota e análise em larga escala e a alimentação óptica do próprio sistema de monitoramento, tornando a arquitetura integralmente alimentada pelo sistema PoF.

Referências

[1] S. Silva and C. Marques, "Optical fibre sensors and sensing networks: overview of the main techniques and their field of application," *Sensors*, vol. 22, no. 21, p. 8253, 2022.

[2] M. Matsuura, "Recent advancement in power-over-fiber technologies," in *Photonics*, vol. 8, no. 8. Multidisciplinary Digital Publishing Institute, 2021, p. 335.

[3] M. A. da Rosa, C. M. Henrique, G. S. Bolacell, H. I. Del Monego, and P. C. R. de Lima Junior, "High-voltage measurement infrastructure based on optical technology for transmission lines," *Energies*, vol. 18, no. 4, p. 830, 2025.

[4] H. Kim and Y. Park, "Technology update on patent and development trend of power-over-fiber," *Journal of Photonics for Energy*, vol. 13, no. 1, p. 011001, 2023.

[5] M. Alonso, P. Garcia, I. Lopez, and C. Mateo, "Optically-powered wireless sensor nodes towards industrial applications," *Sensors*, vol. 21, no. 2, p. 424, 2021.

[6] L. C. Souza, E. R. Neto, E. S. Lima, and A. C. Sodré Jr., "Optically-powered wireless sensor nodes towards industrial internet of things," *Sensors*, vol. 22, no. 1, p. 57, 2021.

[7] M. Matsuura, K. Aoyama, M. Yamada, and T. Nagashima, "Over 40-w electric power and optical data transmission using an optical fiber," *IEEE Transactions on Power Electronics*, vol. 36, no. 5, pp. 4532–4540, 2021.

[8] S. Fafard and D. Masson, "Demonstration of power-over-fiber with watts of output power capabilities over kilometers or at cryogenic temperatures," *Photonics*, vol. 11, no. 7, p. 596, 2024.

[9] H. Yusoff, F. Rahman, and N. Hashim, "Power-over-fiber using a multimode optical power with a core diameter of 50 µm," *Optical and Quantum Electronics*, vol. 56, no. 2, p. 154, 2024.

[10] J. D. López-Cardona, D. Sánchez Montero, and C. Vázquez, "Smart remote nodes fed by power over fiber in internet of things applications," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7328–7334, 2019.

[11] R. Altuna, J. D. López-Cardona, and C. Vázquez, "Monitoring of power over fiber signals using intercore crosstalk in arof 5g nr transmission," *Journal of Lightwave Technology*, vol. 41, no. 23, pp. 7155–7161, 2023.

[12] L. C. De Souza, V. D. P. Souto, and A. C. Sodré, "Radio-and power-over-fiber integration for 6g networks: Challenges and future prospects," *IEEE access*, 2024.

[13] C. Vázquez, J. D. López-Cardona, P. C. Lallana, D. S. Montero, F. M. A. Al-Zubaidi, S. Pérez-Prieto, and I. P. Garcilópez, "Multicore fiber scenarios supporting power over fiber in radio over fiber systems," *IEEE Access*, vol. 7, pp. 158 409–158 418, 2019.

[14] E. M. Bertozzi and B. C. do Nascimento, "Pof - monitoring: Plataforma de monitoramento remoto para sistemas Power-over-Fiber," https://github.com/EduBertozzi/pof-monitoring, 2026.

Emprego de Rádio Definido por Software como MAGE-COM em Veículos de Superfície Não Tripulados

Daniel Scorzello Lopes, Paulo Victor Gonçalves Sampaio e Arthur Pinheiro de Araújo Costa
Marinha do Brasil
daniel.scorzello@marinha.mil.br, paulo.sampaio@marinha.mil.br e arthur.costa@marinha.mil.br

***Abstract*—Naval technological evolution has driven the use of Unmanned Surface Vehicles (USV) in patrols. This research evaluates Software-Defined Radios (SDR) as Communications Electronic Support Measures (ESM-COM) in USVs. Using the RTL-SDR V3 receiver and MATLAB software, real-world measurements were conducted in Guanabara Bay and compared with theoretical VHF propagation models. The results estimated reception ranges and projected patrol areas in regions of interest to the Navy, in accordance with the National Defense Strategy. It is concluded that SDR enhances maritime surveillance capability and the protection of critical infrastructures.**

***Index Terms*—Unmanned Surface Vehicles, Software-Defined Ra- dio, Communications Electronic Warfare Support Measures.**

***Resumo*—A evolução tecnológica naval impulsionou o uso de Veículos de Superfície Não Tripulados (USV) em patrulhas. Esta pesquisa avalia Rádios Definidos por Software (RDS) como Medidas de Apoio à Guerra Eletrônica de Comunicações (MAGE-COM) em USV. Utilizando o receptor RTL-SDR V3 e o software MATLAB, realizaram-se medições na Baía de Guanabara, comparadas a modelos teóricos de propagação em VHF. Os resultados estimaram alcances de recepção e projetaram áreas de patrulha em áreas de interesse da Marinha, conforme a Estratégia Nacional de Defesa. Conclui-se que o RDS potencializa a capacidade de vigilância marítima e proteção de infraestruturas críticas.**

***Palavras chave*—— Veículos de Superfície Não Tripulados, Rádio Definido por Software, Medidas de Apoio à Guerra Eletrônica de Comunicações.**

I. Introdução

As Medidas de Apoio à Guerra Eletrônica de Comunicações (MAGE-COM) são ações com a intenção de coletar informações de forma não cooperativa sobre um adversário, por meio da interceptação de emissões de radiofrequência (RF) oriundas de sistemas de comunicações. As informações obtidas podem ser utilizadas para fins de inteligência, por meio de uma análise detalhada dos dados internos do sinal, ou para fins táticos, que se baseiam em uma análise imediata dos dados aparentes para um uso rápido em operações de combate [**1**]. A partir do conhecimento das intenções do oponente, pode-se adotar ações adequadas visando proteção e ataque eletrônico no âmbito das comunicações [**2**].

Em contextos de conflitos contemporâneos, observa-se uma significativa evolução em diversos setores, dentre eles no ambiente marítimo e no setor de equipamentos voltados para emprego militar. A evolução dos meios navais de patrulha e de guerra tem tornado os navios cada vez mais adaptados à realidade do cenário global, levando em consideração a tecnologia para sua produção e os recursos tecnológicos embarcados, que necessitam de constante atualização. Nesse contexto, o desenvolvimento dos Veículos de Superfície Não Tripulados, ou Unmanned Surface Vessels (USV), tem sido impulsionado pelo aumento global do interesse em diversas atividades no mar, como o emprego dual (militar e civil) visando a proteção de infraestruturas críticas, que são instalações cuja destruição, total ou parcial, provoque sério impacto social, ambiental, econômico, político, internacional ou à segurança nacional e da sociedade, como, por exemplo, bases militares [**3**].

Visando aprimorar o patrulhamento em mar aberto e águas interiores e obter uma vantagem estratégica em termos de consciência situacional na área de atuação, é relevante a utilização de equipamentos capazes de realizar a função de MAGE-COM. A coleta passiva das comunicações inimigas pode auxiliar a reconhecer as intenções hostis e permitir a geração de um banco de dados de sinais de comunicações com alto valor estratégico, possibilitando antecipar ações, identificar e localizar os possíveis alvos [**1**].

Nesse sentido, o RDS é um equipamento que pode ser utilizado para coletar informações de comunicações. Ele é capaz de realizar o monitoramento e gravação de ampla faixa de frequências, incluindo aquelas utilizadas no ambiente marítimo, com um baixo custo de investimento e dimensões reduzidas. Além disso, permite a alteração de parâmetros via software, eliminando a necessidade de troca de componentes para reconfiguração [**4**], acrescentando alto grau de flexibilidade ao sistema. Essas características somadas possibilitam a utilização do RDS como ferramenta de MAGE-COM eficaz para obtenção de sinais de comunicações inimigas, contribuindo para a proteção de infraestruturas críticas e vigilância marítima.

Face ao exposto, torna-se de extrema importância o emprego de novas tecnologias visando o patrulhamento do litoral brasileiro, com intuito de combater ilícitos nas Águas Jurisdicionais Brasileiras (AJB), e em especial nas áreas de interesse da Marinha do Brasil, conforme a Estratégia Nacional de Defesa (END), sendo elas a área marítima compreendida entre Santos-SP e Vitória-ES, Margem Equatorial, na qual a foz do rio Amazonas se encontra, e Elevação do Rio Grande [**5**]. Para atender à essa necessidade, será verificada a viabilidade da utilização de RDS aplicado para fins de MAGE-COM nas frequências de Very High Frequency (VHF), faixa do espectro eletromagnético amplamente utilizada no mar.

O uso de RDS aplicados em USV possui relevância devido

ao seu baixo custo e facilidade de aquisição, e visa a redução de custos provenientes da utilização de meios navais tripulados, como gastos com pagamento de pessoal e combustível, além do aumento de autonomia na área marítima de operação e redução da exposição da vida humana aos riscos inerentes da missão.

Como objetivo do trabalho, busca-se avaliar o desempenho do RDS escolhido para o projeto para sua possível aplicação como MAGE-COM em USV, por meio da utilização efetiva de um RDS como receptor de sinais em VHF transmitidos por uma unidade transmissora fixa, localizada na baía de Guanabara. Após a recepção dos sinais transmitidos, estes serão comparados com modelos teóricos de perda por propagação, com auxilio do software MATLAB, a fim de permitir a análise do alcance máximo de interceptação de comunicações em VHF. Por fim, este trabalho pretende sugerir áreas de patrulha para USV nos locais propostos na END.

II. MODELOS DE PROPAGAÇÃO UTILIZADOS

Nesta seção são apresentados os modelos de perdas por propagação de ondas eletromagnéticas a serem utilizados no trabalho, que são os Modelos de Perda de Percurso no Espaço Livre (FSPL) e Modelo de Perda de Dois Raios (PL_{2r}), condições para propagação em linha de visada (LoS) e definição de Horizonte Rádio.

A. Elementos associados à Perda de Percurso

Os sinais transmitidos pelo canal de Rádio Frequência (RF) de comunicações móveis estão sujeitos aos fenômenos de reflexão, refração e dutos de evaporação, que culminam na variação do sinal recebido. Esses problemas podem ser mitigados se as características de propagação do canal puderem ser previstas, por meio de modelos de propagação.

Esses modelos são conjuntos de expressões matemáticas usados para representar as características de um canal de RF em um determinado cenário e ambiente, considerando fatores como a frequência de operação, a distância entre o transmissor e o receptor e as condições atmosféricas. A escolha incorreta de um modelo para um cenário pode levar à superestimação ou subestimação da Perda de Percurso, ou Path Loss (PL) [**6**].

A PL entre um par de antenas é dada pela razão entre a potência transmitida e a potência recebida, comumente expressa em decibéis (dB), podendo ser determinada convencionalmente pela equação a seguir:

$$PL(dB) = 10\log\left(\frac{P_t}{P_r}\right) \tag{1}$$

Porém, diversos elementos dos sistemas de transmissão e recepção estão associados à PL. Visando o correto cálculo, deve-se considerar os ganhos das antenas de transmissão e recepção, sendo usadas como referência antenas isotrópicas, bem como as perdas dos cabos utilizados nas linhas de transmissão e recepção e as potências transmitidas e recebidas [**7**]. A equação a seguir expressa a PL, em dB:

$$PL(dB) = P_t + G_t + G_r - P_r - L_t - L_r \tag{2}$$

B. Modelos de Perda de Percurso no Espaço Livre

O modelo de PL no Espaço Livre, ou Free-Space Path Loss (FSPL) é o modelo mais simples utilizado, prevendo a intensidade do sinal de RF recebido a uma certa distância da fonte emissora. Esse modelo é utilizado para terrenos com Linha de Visada direta, ou Line of Sight (LoS), entre as fontes emissoras e receptoras [**6**].

A equação abaixo mostra o cálculo da FSPL:

$$FSPL(\text{dB}) = 20\log(d) + 20\log(f) + 20\log\left(\frac{4\pi}{c}\right) \tag{3}$$

onde ”d”é a distância entre as antenas de transmissão e recepção, em metros, e ”f”a frequência de transmissão utilizada, em Hertz.

C. Modelos de Dois Raios

Outro modelo utilizado para predição da perda de intensidade do sinal que chega à antena receptora é o Modelo de Dois Raios. Segundo [**7**], diferente do modelo de Espaço Livre, esse método considera a existência de dois caminhos principais para o sinal, sendo eles o raio direto e o raio refletido no solo. Esse modelo pode ser utilizado para retratar a propagação em ambientes marítimos.

A reflexão da superfície do mar é muito importante para projetar um canal, por se tratar de uma superfície que possui comportamento varido, ora agitada e rugosa, ora calma e lisa. Em algumas ocasiões, essa reflexão pode ter influência significativa no sinal resultante, podendo ocasionar interferências construtivas ou destrutivas no sinal recebido. Quando ocorre a LoS, a PL pode ser medida usando o Modelo de Dois Raios [**8**], por meio da equação a seguir, que considera as alturas das antenas de transmissão e recepção, além dos parâmetros já considerados no FSPL:

$$PL_{2r}(\text{dB}) = -10\log\left[\left(\frac{\lambda}{4\pi d}\right)^2\left(2\sin\left(\frac{2\pi h_r h_t}{\lambda d}\right)\right)^2\right] \tag{4}$$

D. Condições para estabelecimento de Linha de Visada (LoS)

A Linha de Visada (LoS) é a condição em que existe um caminho desobstruído entre a antena transmissora e a receptora, permitindo que as ondas de rádio viajem diretamente de um ponto a outro. Para ocorrer a LoS, é necessário que sejam estabelecidas duas condições fundamentais.

A primeira está relacionada à zona de Fresnel. As Zonas de Fresnel são volumes elipsoidais concêntricos que se formam ao entorno da linha direta entre uma estação transmissora e receptora. Para haver LoS, é necessária a desobstrução total de, pelo menos, 60% do elipsóide da primeira zona de Fresnel (Δ/rd $\approx$ 0.6) [**9**].

A segunda condição está relacionada a correção da curvatura da Terra, por meio da utilização do fator K, gerando um Raio Terrestre Equivalente. Adota-se o valor de $K_{\text{médio}}$ = 4/3 para caracterizar a atmosfera padrão e $K_{\text{mínimo}}$ cujo valor é determinado de acordo com a recomendação ITU-R. P.530 e deve ser excedido em 99,9% do tempo no pior mês do ano [**10**]. Contudo, é fundamental considerar que o fator k é dinâmico,

alterando-se conforme as condições locais de temperatura, pressão atmosférica e umidade relativa do ar.

E. Horizonte Rádio

Graças ao fenômeno da refração, a energia eletromagnética acaba se encurvando e acompanhando, até uma certa distância, a curvatura da Terra, fazendo com que o alcance máximo em ondas de RF seja maior que o horizonte visual ou geométrico, conhecido como horizonte rádio [**11**].

Esse fenômeno é comumente modelado pela utilização do $K_{\text{médio}}$ = 4/3 do raio terrestre real. A distância do horizonte rádio é, teoricamente, a distância máxima que um sinal de rádio desobstruído percorrerá a partir do transmissor antes de tangenciar a superfície [**9**]. A distância do horizonte rádio, obtida em quilômetros, depende da altura das antenas do transmissor e do receptor, em metros, e do valor de k, podendo ser obtida por meio da equação dada por:

$$D_{hor} = 4.2\left(\sqrt{h_r} + \sqrt{h_t}\right) \tag{5}$$

III. DESCRIÇÃO E ANÁLISE DOS RESULTADOS

De modo a tornar exequível a realização das medições, foi necessário o estabelecimento de locais que representassem a propagação no ambiente marítimo com linha de visada direta, a fim de gerar uma representação próxima à realidade encontrada por USV em patrulhas. Logo, foram escolhidas duas Organizações Militares separadas fisicamente pela baía de Guanabara, no estado do Rio de Janeiro: Corveta Barroso, Navio da Marinha do Brasil, localizado no cais norte do Arsenal de Marinha do Rio de Janeiro (AMRJ) e Base Almirante Castro e Silva (BACS), localizado em Niterói-RJ. A figura a seguir apresenta ambos os pontos utilizados, distantes 4.117 metros, aproximadamente.



Fig. 1. Distância entre as estações envolvidas

A Corveta Barroso atuou como a estação transmissora, conduzindo a emissão de mensagens padronizadas “CGAEM, CGAEM, aqui Corveta Barroso, experiência fonia, câmbio”. A transmissão ocorreu no canal 08 do VHF Marítimo, correspondente à frequência de 156,4 MHz. O equipamento de transmissão utilizado foi o transceptor VHF marítimo IC-M304, da marca ICOM, acoplado a uma antena Shakespeare HS-2774. Durante a campanha de medições, foram realizadas 30 aquisições de dados com o objetivo de caracterizar o desempenho do enlace de comunicação por meio de uma quantidade satisfatória de amostras.

A estação de recepção foi instalada na BACS, composta pelo Rádio Definido por Software RTL-SDR V3, conectado a um notebook utilizando o software RDS++, e uma antena omnidirecional AP0191, da marca STEELBRAS. Ademais, foram empregados cabos coaxiais RG58 de 6 metros de comprimento em ambas as estações. A configuração física do sistema de recepção utilizado durante a campanha de medições é ilustrada na figura a seguir.



Fig. 2. Sistema de recepção usando RDS

Dada a necessidade de se conhecer os elementos associados às perdas de percurso, a tabela apresenta um resumo dos parâmetros dos equipamentos utilizados em ambas as estações envolvidas nas medições. A partir dos dados a seguir, torna-se possível a realização dos cálculos de perdas ocorridas durante as medições, bem como de cada modelo utilizado.

TABELA I
PARÂMETROS DOS EQUIPAMENTOS USADOS

Parâmetro	Valor
Potência de Transmissão	30 dBm
Ganho da antena de transmissão	4 dBi
Ganho da antena de recepção	3.15 dBi
Altura da antena de transmissão	13,5 m
Altura da antena de recepção	3 m
Perdas nos cabos (Tx e Rx)	1 dB

A. Discussão dos Resultados Obtidos

Após as transmissões realizadas pela Corveta Barroso, os dados obtidos indicaram uma potência média recebida variando entre -35 dBm e -30 dBm e relação sinal-ruído (SNR) média estimada em 60 dB, possibilitando o cálculo da potência do ruído por meio da equação abaixo.

$$\text{SNR}_{\text{dB}} = 10 \times \log_{10}\left(\frac{P_s}{P_n}\right) \tag{6}$$

A partir desse cálculo, obteve-se um valor de -90 dBm para a potência de ruído. Considerando que a sensibilidade média do RTL-SDR V3 é de, aproximadamente, -130 dBm em sua faixa de operação, o valor de -90 dBm foi adotado como o limiar de inteligibilidade para os cálculos de predição de alcance. Isso se justifica pelo fato de que sinais com potência inferior a -90 dBm não seriam detectáveis de forma confiável a partir da demodulação do sinal recebido.

Com base nos parâmetros técnicos dos componentes do sistema, dos dados obtidos através da campanha de medições e

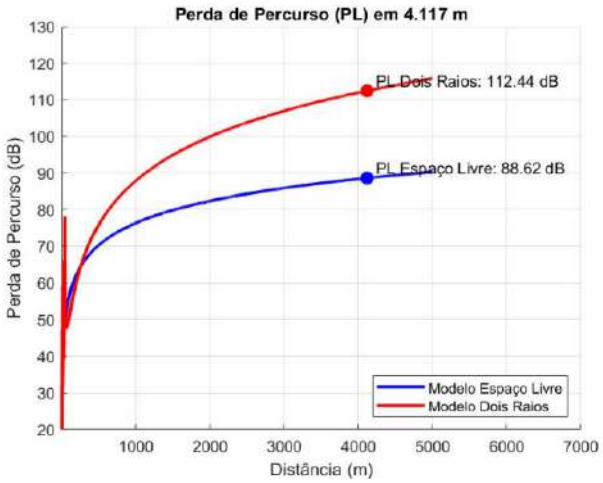


Fig. 3. Gráfico da Perda de Percurso em 4.117 metros

da frequência de operação, pode-se calcular o alcance máximo do enlace de comunicação. Para isso, foi desenvolvido um código no software MATLAB, que implementa os modelos de propagação em Espaço Livre e Dois Raios.

Os resultados das predições teóricas serão comparados aos dados obtidos nas medições de campo. Essa comparação permitirá validar a precisão de cada modelo e determinar qual deles oferece a melhor aproximação para as condições reais do ambiente de propagação. A figura 3 mostra a análise gráfica das perdas em cada modelo, especificamente para a distância utilizada entre as estações na campanha de medições.

A predição da PL foi realizada com base em uma distância de 4.117 metros. A aplicação do modelo de Dois Raios resultou em uma PL2r de 112,44 dB, aplicando a equação 4. Com base neste valor, a potência recebida, de acordo com a equação 2, foi calculada como -56,59 dBm. De forma análoga, a aplicação da equação 3 do modelo de Espaço Livre resultou em uma FSPL de 88,62 dB. Este valor de perda corresponde a uma potência recebida de -32,77 dBm, de acordo com a equação 2. A figura 4 mostra os valores das potências recebidas de cada modelo para a distância de 4.117 metros.

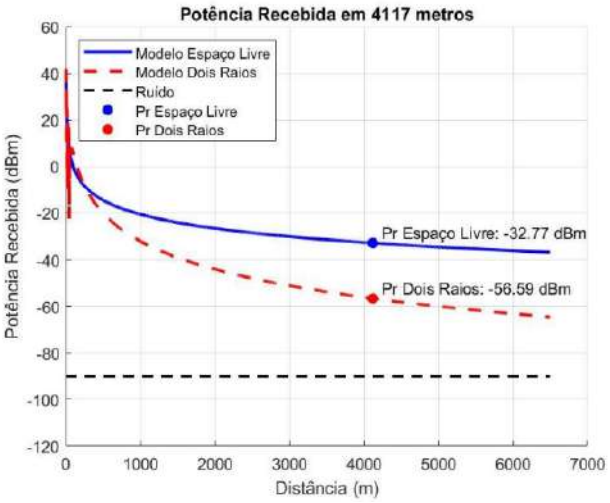


Fig. 4. Gráfico da Potência Recebida em 4.117 metros

Uma comparação entre os valores preditos e os dados empíricos, que indicaram uma potência recebida entre -35 dBm e -30 dBm, revela que o modelo de Espaço Livre demonstrou maior precisão para as condições específicas desta campanha de medições. A pequena discrepância entre o valor teórico de -32,76 dBm e os valores reais obtidos sugere o modelo de Espaço Livre como uma aproximação mais adequada para a predição do alcance neste cenário.

Após a validação do modelo do Espaço Livre para a predição de perdas de propagação para a campanha de medições realizada, foi gerado um script no MATLAB para simular a variação da potência recebida em um intervalo de 1 a 37.000 metros, aproximadamente 20 milhas náuticas (MN).

À medida em que a distância aumenta, a potência do sinal recebido diminui. Nesse cenário, o alcance máximo do enlace de comunicação pode ser limitado por dois fatores, sendo eles a distância do horizonte rádio ou a potência do ruído ambiente. Considerando que o nível de ruído obtido nas medições (-90 dBm) foi significativamente superior à sensibilidade do receptor (-130 dBm), o ruído do ambiente se torna o principal fator limitante para a inteligibilidade do sinal e, inicialmente, para o alcance máximo do sistema.

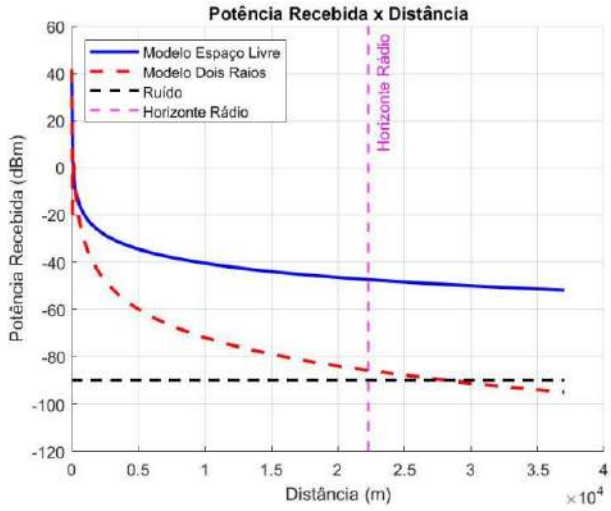


Fig. 5. Cálculo do alcance máximo de comunicações

A partir da figura 5 pode-se verificar que, para o limite de distância aplicado para o gráfico de 20 MN, a curva de potência recebida pela estação receptora, para o modelo de Espaço Livre, não intercepta a linha de nível de ruído ambiente. Isso indica que, para este modelo e condições de medição, a limitação do alcance não é causada pela degradação do sinal pelo ruído, mas sim pela distância do horizonte rádio. A distância do horizonte rádio depende das alturas das antenas, conforme observado na equação 5. Comparando com o modelo de Dois Raios, o sinal no modelo de Espaço Livre permanece com uma potência superior ao ruído ao longo de toda a distância do horizonte, garantindo a inteligibilidade da comunicação.

Após determinado que a limitação do alcance é obtida por meio do horizonte rádio, é possível obter, a partir da equação 5 contida no script do MATLAB, o valor do horizonte rádio como sendo 22.274 metros (aproximadamente 12 MN). Logo, o valor encontrado pode ser definido como o alcance máximo

para interceptar comunicações, utilizando um RTL-SDR V3, para fins de MAGE-COM visando interceptação de sinais de comunicações.

IV. ESTIMAÇÃO DE ÁREAS DE PATRULHA DE USV

Pretende-se, neste tópico, apresentar possibilidades de utilização de USV nas áreas estratégicas de patrulha, conforme estabelecido na END, a fim de otimizar a eficiência das atividades de patrulhamento, e demonstrar a aplicabilidade operacional e estratégica do estudo realizado. A adoção dos USV visa reduzir custos operacionais, ampliar a permanência em áreas de operações de patrulha e, quando dotados de equipamentos de MAGE-COM para frequências em VHF, viabilizar a detecção antecipada de comunicações hostis. Esse alerta prévio permite o acionamento oportuno de plataformas tripuladas equipadas com sistemas de armas, conferindo maior capacidade de resposta às ameaças identificadas para proteção de bases militares ou recursos considerados críticos para soberania nacional.

A estimação das áreas de patrulha baseou-se diretamente nos resultados obtidos durante a campanha de medições conduzida na baía de Guanabara, na qual foi possível verificar a coerência entre os valores medidos e o modelo teórico de propagação no Espaço Livre, resultando em um alcance máximo de 12 MN, limitado pelo horizonte rádio.

Dessa forma, os parâmetros obtidos experimentalmente serviram de suporte técnico-científico para subsidiar as propostas das áreas de patrulha dos USV, considerando cenários reais de operação, como a foz do rio Amazonas, o Porto de Santos e o Porto do Rio de Janeiro, conferindo credibilidade operacional à solução proposta. Entre as variáveis críticas para essa medição, destaca-se a altura das antenas de recepção e de transmissão, que podem variar em função das embarcações empregadas, impactando diretamente o alcance do horizonte rádio, além das condições da atmosfera, que variam com temperatura, umidade e pressão.

A. Foz do rio Amazonas

A foz do rio Amazonas é considerada uma área estratégica para a patrulha da Marinha do Brasil [**5**]. A figura a seguir ilustra as áreas propostas para a implementação dos USV na região, as quais, se patrulhadas em conjunto, possibilitam a cobertura de todas as entradas do rio, conforme estabelecido pela END. As áreas de patrulha são representadas pelos círculos internos, concêntricos aos círculos que definem a área de cobertura, e possuem raio de 5 km. A utilização de três USV permite o monitoramento simultâneo das três áreas sugeridas, resultando em maior eficiência no patrulhamento e no controle das comunicações.

B. Porto de Santos-SP

Assim como a foz do rio Amazonas, as águas próximas ao porto de Santos, no estado de São Paulo, são de relevância estratégica para a Marinha do Brasil, por fazerem parte da área marítima compreendida entre Santos-SP e Vitória-ES [**5**]. A figura a seguir ilustra a localização sugerida para a implementação de USV na entrada desse porto, um dos mais movimentados do país. Similar à região anterior, a área de patrulha é representada pelo círculo interno, com 5 km de raio,



Fig. 6. Possíveis áreas de patrulha na foz do rio Amazonas

que é concêntrico ao círculo que delimita a área de cobertura de comunicações.

A figura ilustra a capacidade de um USV cobrir uma área extensa nas proximidades da baía de Santos, região caracterizada pela grande quantidade de navios mercantes fundeados. No entanto, a utilização do USV posicionado nessa área não permite um monitoramento eficiente do porto de Santos, pois não há visada direta devido ao relevo e às construções urbanas. Portanto, para monitorar o porto e as embarcações que navegam nas proximidades, é necessário também empregar um USV nas águas interiores de Santos.



Fig. 7. Possível área de patrulha no porto de Santos

C. Porto do Rio de Janeiro

A análise do porto do Rio de Janeiro é justificada por sua importância estratégica, visto que abriga a sede da Esquadra Brasileira, que pode ser visto como uma das mais importantes infraestruturas críticas nacionais por abrigar Navios de guerra. A figura abaixo mostra a área de cobertura de um USV posicionado no interior da baía de Guanabara, entre o Aeroporto Santos Dumont, localizado no Rio de Janeiro, e a ilha de Boa Viagem, em Niterói. A localização sugerida para a utilização do USV é representada pelo círculo interno, que possui 1,3 km de raio.

Esta posição do USV permite o monitoramento de praticamente toda a área das águas interiores, incluindo navios atracados no porto e embarcações fundeadas no interior da baía de Guanabara. Adicionalmente, é possível interceptar comunicações de embarcações que adentram a baía, o que garante um

monitoramento abrangente. A cobertura sobre terra, no entanto, não será eficiente, pois a visada direta é comprometida pelo relevo e pelas estruturas urbanas.

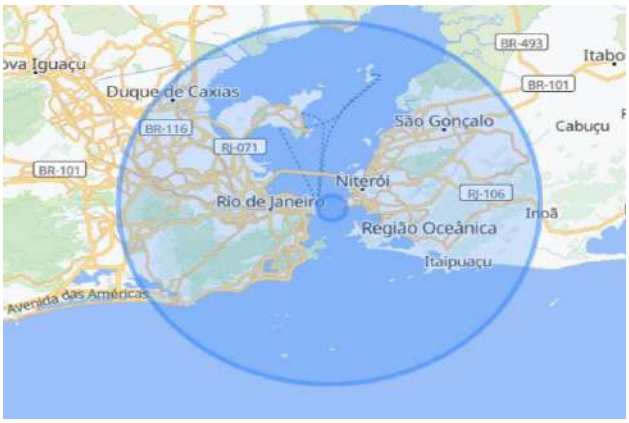


Fig. 8. Possível área de patrulha na baia de Guanabara

V. Conclusão

O presente trabalho teve como objetivo principal avaliar o desempenho do RDS para sua aplicação como MAGE-COM em USV. Para isso, foram realizadas medições na baía de Guanabara, com o intuito compará-las com os modelos de propagação no Espaço Livre e Dois Raios e, após verificação do modelo mais apropriado, sugerir áreas de aplicação de USV para operações de patrulha marítima.

A pesquisa de campo revelou que, nas condições de medição, a potência média do sinal recebido foi próxima à potência teórica calculada para o modelo no Espaço Livre, o qual foi escolhido para realização dos cálculos de predição de alcance. Por meio do uso do software MATLAB, foram gerados códigos para visualização das perdas de propagação para cada modelo e os seus respectivos alcances, sendo obtido o valor de 12 MN de alcance, limitado pelo horizonte rádio. Os resultados das medições confirmam que mesmo um equipamento de baixo custo, como o RTL-SDR V3, é capaz de detectar comunicações hostis dentro do raio de cobertura definido, demonstrando a viabilidade técnica e operacional da proposta.

Apesar de não terem sido feitas a bordo de um USV, buscou- se alternativas de representar de forma exequível a propagação em ambiente marítimo, tendo em vista que as estações envolvidas no enlace de comunicações se encontravam separadas fisicamente pela baía de Guanabara.

A partir da análise dos resultados, a pesquisa reforça a relevância prática da utilização de USV equipados com RDS para fins de Defesa, alinhado à END, destacando que é essencial ao Poder Naval possuir meios com recursos de detec- ção, identificação e neutralização de ameaças nas AJB, além de estabelecer as áreas de importância estratégica abordadas.

O uso de USV para patrulhamento, conforme sugerido pelo estudo, poderia aumentar a eficiência e reduzir os custos operacionais, permitindo maior autonomia e permanência no local de patrulha, comparado aos meios tripulados. Já a aplicação dos RDS, representado pelo RTL-SDR V3, devido ao seu baixo custo e flexibilidade, pode ser uma solução economicamente viável para coletar informações de comuni- cações, auxiliando na segurança de portos e na detecção de atividades ilícitas.

Dessa forma, a pesquisa buscou contribuir para o debate sobre o patrulhamento de áreas sensíveis da costa brasileira sem a necessidade da utilização inicial de meios navais tripulados, por meio da aplicação de tecnologias e recursos ainda pouco exploradas no cenário nacional.

Referências

[1] POISEL, Richard A. *Introduction to Communication Electronic Warfare Systems*. Boston: Artech House, 2002.

[2] NERI, Filippo. *Introduction to Electronic Defense Systems*. Boston: Artech House, 2018.

[3] BRASIL. *Decreto nº 9.573, de 22 de novembro de 2018 — Aprova a Política Nacional de Segurança de Infraestruturas Críticas*. Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Decreto/D9573.htm.

[4] JANSON, J. *Radio Definido por Software: Estudo e Realização de Teste com uma Plataforma Livre*. 2012. 60 f. Monografia (Tecnólogo em Sistemas de Telecomunicações) – Centro Federal de Educação Tecnológica de Santa Catarina, São José, 2012.

[5] BRASIL. *Decreto nº 12.725, de 18 de novembro de 2025 — Aprova a Política Nacional de Defesa, a Estratégia Nacional de Defesa e o Livro Branco de Defesa Nacional*. Brasília, DF: Presidência da República, 18 nov. 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/D12725.htm.

[6] ERUNKULU, O. O. et al. Cellular Communications Coverage Prediction Techniques: A Survey and Comparison. *IEEE Access*, v. 8, p. 113052–113077, 2020. Disponível em: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9119431. Acesso em: 10 ago. 2025. DOI: 10.1109/ACCESS.2020.3003247.

[7] SAUNDERS, Simon R.; ZAVALA, Alejandro Aragón. *Antennas and propagation for wireless communication systems*. 2. ed. Chichester, West Sussex: John Wiley e Sons, 2007.

[8] HABIB, A.; MOH, S. Wireless channel models for over-the-sea communication: A comparative study. *Applied Sciences*, v. 9, n. 3, 2019. Disponível em: https://www.mdpi.com/2076-3417/9/3/443. Acesso em: 12 ago. 2025. DOI: 10.3390/app9030443.

[9] DUGGAL, G.; BUEHRER, R.; TRIPATHI, N.; REED, J. Line-of-Sight Probability for Outdoor-to-Indoor UAV-Assisted Emergency Networks, 2023. DOI: 10.48550/arXiv.2302.14709.

[10] ITU-R, *Recommendation ITU-R P.530-19: Propagation data and prediction methods required for the design of terrestrial line-of-sight systems*, P Series: Radio-wave propagation, 09/2025.

[11] BRASIL. GE-201: Medidas de Guerra Eletrônica - Medidas de Apoio à Guerra Eletrônica. 2ª ed. Niterói, RJ. Marinha do Brasil. Centro de Guerra Acústica e Eletrônica da Marinha, 2023.

Autores

Daniel Scorzello Lopes Possui graduação em Ciências Navais (2018) pela Escola Naval e Pós-Graduação em Sistemas de Defesa Eletrônica (2025) pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

Paulo Victor Gonçalves Sampaio Possui graduação em Ciências Navais (2010) pela Escola Naval e Mestrado em Engenharia Elétrica e de Telecomunicações, com Área de Concentração em Sistemas de Telecomunicações (2024), pela Universidade Federal Fluminense (UFF).

Arthur Pinheiro de Araújo Costa Possui graduação em Ciências Navais (2015) pela Escola Naval e Mestre em Sistemas e Computação pelo Instituto Militar de Engenharia (IME).

Engenharia reversa de um oxímetro de pulso comercial P-01 para a aquisição de dados fisiológicos por meio de ESP32 em aplicações IoT

Luis Diego Ferreyra Contreras & Melissa Andrea Herrera Tambo

Faculdade de Informática e Eletrônica – Univalle, Cochabamba - Bolivia
fcl0031576@est.univalle.edu, htm0031168@est.univalle.edu
https://orcid.org/0009-0003-7018-3375, https://orcid.org/0009-0008-5085-9725

***Abstract*—** Pulse oximeters are biomedical devices used for the non-invasive measurement of blood oxygen saturation (SpO_2) and heart rate, essential parameters for assessing the cardiorespiratory status of patients [10]. This paper presents a reverse engineering process applied to a commercial pulse oximeter to analyze its electronic architecture and integrate its data into a biomedical monitoring system. The study included disassembling the device, identifying components, and analyzing the communication protocol used for data transmission. Using an embedded platform based on ESP32, the signals generated by the sensor were acquired and decoded. The results allowed for the identification of a UART communication interface and the extraction of the transmitted physiological parameters, demonstrating the feasibility of integrating commercial oximeters into remote monitoring systems based on IoT technologies [8], [7].

***Index Terms*—** Pulse oximeter, reverse engineering, UART communication, biomedical monitoring, embedded systems.

***Resumo*—** Os oxímetros de pulso são dispositivos biomédicos utilizados para a medição não invasiva da saturação de oxigênio no sangue (SpO_2) e da frequência cardíaca, parâmetros essenciais para a avaliação do estado cardiorrespiratório dos pacientes [10]. Neste trabalho é apresentado um processo de engenharia reversa aplicado a um oxímetro de pulso comercial com o objetivo de analisar sua arquitetura eletrônica e integrar seus dados em um sistema de monitoramento biomédico para aplicações IoT. O estudo incluiu a desmontagem do dispositivo, a identificação de componentes e a análise do protocolo de comunicação utilizado para a transmissão de dados. Por meio de uma plataforma embarcada baseada em ESP32, foi possível adquirir e decodificar os sinais gerados pelo sensor. Os resultados permitiram identificar uma interface de comunicação UART e extrair os parâmetros fisiológicos transmitidos, demonstrando a viabilidade de integrar oxímetros comerciais em sistemas de monitoramento remoto baseados em tecnologias IoT [8], [7].

***Palavras-chave*—** Oxímetro de pulso, engenharia reversa, comunicação UART, monitoramento biomédico, sistemas embarcados.

I. INTRODUÇÃO

O monitoramento contínuo de sinais vitais por meio de dispositivos eletrônicos tornou-se uma ferramenta fundamental nos sistemas modernos de atenção médica, especialmente em aplicações de telemedicina, monitoramento domiciliar e atendimento pré-hospitalar [5], [1]. Nesse contexto, o uso de tecnologias da Internet das Coisas (IoT) permite a interconexão de dispositivos biomédicos com plataformas digitais para a supervisão remota de pacientes e a transmissão de dados fisiológicos em tempo real [3], [8]. Diversos estudos demonstraram que a integração de sensores biomédicos com plataformas embarcadas e sistemas de comunicação sem fio permite desenvolver soluções eficientes para a aquisição e transmissão de parâmetros fisiológicos, facilitando a supervisão contínua de pacientes em ambientes clínicos e remotos [9], [14]. Essas arquiteturas tecnológicas fazem parte da denominada Internet das Coisas Médicas (IoMT), que integra dispositivos biomédicos, redes de comunicação e plataformas de processamento de dados para melhorar a qualidade e acessibilidade dos serviços de saúde [7].

A integração de dispositivos médicos comerciais em sistemas de monitoramento personalizados representa um dos desafios técnicos mais frequentes no desenvolvimento de soluções biomédicas de baixo custo. Em numerosos contextos, especialmente em ambientes com recursos limitados, o desenvolvimento de sensores fisiológicos do zero é inviável devido aos custos de projeto, fabricação e validação, recorrendo-se assim a dispositivos comerciais disponíveis no mercado como alternativa tecnológica acessível. No entanto, a ausência de documentação técnica por parte dos fabricantes, juntamente com a falta de interfaces de comunicação abertas e especificações de interoperabilidade, impede sua integração direta com plataformas embarcadas. Essa limitação dificulta o aproveitamento dos dados fisiológicos gerados por esses dispositivos em sistemas externos de monitoramento, gerando a necessidade de aplicar técnicas de engenharia reversa para extrair a informação medida e torná-la acessível a aplicações de monitoramento remoto e telemedicina.

A engenharia reversa aplicada a dispositivos eletrônicos consiste na análise sistemática do hardware e do comportamento de um sistema com o objetivo de compreender seu funcionamento interno sem acesso à sua documentação original. No contexto de dispositivos biomédicos, esse processo adquire uma dimensão adicional, pois os dados que se busca recuperar correspondem a parâmetros fisiológicos críticos como a saturação de oxigênio no sangue (SpO_2) e a frequência cardíaca (BPM), amplamente utilizados para avaliar o estado fisiológico dos pacientes em ambientes clínicos e de monitoramento remoto [6], [10].

O presente trabalho tem como objetivo identificar e decodificar o protocolo de comunicação interno de um oxímetro comercial tipo dedo modelo P-01, visando a aquisição dos parâmetros fisiológicos transmitidos pelo dispositivo e sua integração com sistemas embarcados para aplicações de monitoramento biomédico remoto. Além disso, busca-se avaliar a viabilidade do uso de dispositivos comerciais de oximetria como fonte de aquisição de sinais fisiológicos em plataformas eletrônicas externas, contribuindo para o desenvolvimento de soluções de monitoramento em tempo real orientadas a aplicações de telemedicina e tecnologias da Internet das Coisas (IoT).

II. MATERIAIS E MÉTODOS

A. Dispositivo de oximetria analisado

Para a aquisição de parâmetros fisiológicos foi utilizado um oxímetro de pulso comercial tipo dedo Finger Pulse Oximeter modelo P-01, selecionado devido à sua disponibilidade comercial, baixo custo e precisão clínica adequada para aplicações de monitoramento não invasivo.

Este dispositivo permite a medição simultânea da saturação de oxigênio no sangue (SpO_2) e da frequência cardíaca (BPM) usando técnicas de espectrofotometria e fotopletismografia, empregando emissores de luz no espectro vermelho e infravermelho juntamente com um receptor de fotodiodo que detecta variações de absorção produzidas pelo fluxo sanguíneo pulsátil [15]. A Figura 1 mostra o dispositivo usado durante os testes experimentais.



Fig 1. Oxímetro de pulso comercial Finger Pulse Oximeter modelo P-01

No entanto, durante a etapa inicial de integração identificou-se que o fabricante não fornecia documentação técnica relacionada a interfaces de comunicação digital ou protocolos de transmissão de dados. Essa limitação impedia a conexão direta do dispositivo com plataformas embarcadas ou sistemas de monitoramento remoto IoT.

Por essa razão, foi proposto um processo de engenharia reversa experimental, orientado a determinar se o oxímetro transmitia internamente os dados medidos para sua tela por meio de algum canal de comunicação acessível. As principais características técnicas do dispositivo são apresentadas na Tabela I.

TABELA I
CARACTERÍSTICAS TÉCNICAS DO OXÍMETRO FINGER PULSE OXIMETER P-01

Características	**Especificação**
Parâmetros medidos	Saturação de oxigênio (SpO_2) e frequência cardíaca (BPM)
Precisão	±2 % SpO_2, ±2 lpm
Faixa de medição	SpO_2: 70–100 % / BPM: 30–240 lpm
Tela	Tela LED/LCD
Alimentação	2 baterias AAA
Consumo	~20 mA
Normativas	CE, FDA 510(k), ISO 13485

B. Exploração física do dispositivo

Com o objetivo de identificar possíveis pontos de acesso às informações internas do sistema, foi realizado um processo de desmontagem controlada do oxímetro, removendo a carcaça externa para acessar a placa eletrônica interna.

Durante a inspeção do hardware, foram identificados os componentes do oxímetro P01, destacando um Front-End Analógico (AFE) em formato SMD para o condicionamento do sinal e um microcontrolador de processamento local. Na placa de controle principal, foi validada a presença de um regulador de tensão LDO para a estabilidade do sistema, um cristal oscilador para a sincronização de processos, capacitores de desacoplamento para a filtragem de ruído, uma tela para exibir os dados e um botão para ligar o equipamento. A Figura 2 mostra a placa eletrônica do oxímetro após a desmontagem.



Fig 2. Placa eletrônica interna do oxímetro de pulso

Por meio de medições com multímetro e análise de continuidade elétrica foram identificadas duas linhas de comunicação adicionais associadas ao microcontrolador interno do dispositivo.

Essas linhas correspondiam aos pinos TX (Transmissão de dados) e RX (Recepção de dados).

A presença dessas conexões sugeriu a existência de um canal de comunicação serial interno, possivelmente utilizado para transmitir a informação processada ao módulo de visualização do dispositivo. A Figura 3 mostra a identificação das linhas TX e RX na placa do oxímetro.

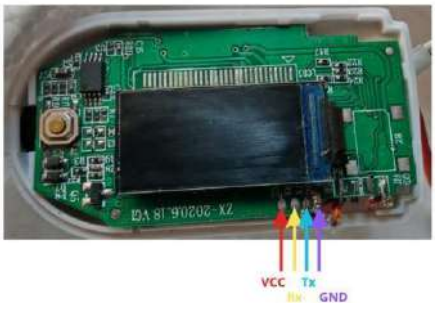


Fig 3. Identificação dos pinos TX e RX no circuito do oxímetro

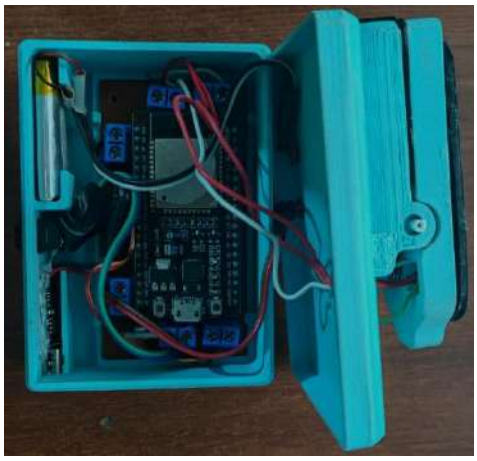
C. Testes iniciais de comunicação serial

Uma vez identificadas as linhas de transmissão, foi estabelecida uma conexão física entre o pino TX do oxímetro e a porta UART do microcontrolador ESP32, configurado como receptor serial.

Para a implementação do protótipo experimental, foi utilizada uma placa de desenvolvimento baseada no microcontrolador Espressif Systems ESP32, conectada diretamente à linha de transmissão do oxímetro e configurada para recepção contínua de dados seriais. O firmware desenvolvido permitiu capturar as tramas transmitidas, identificar os delimitadores de pacote e processar a informação recebida para sua posterior conversão e interpretação dos parâmetros fisiológicos, conforme apresentado na Figura 4. Devido à ausência de documentação

técnica fornecida pelo fabricante relativamente ao protocolo de comunicação utilizado, foi necessário realizar um processo experimental para determinar os parâmetros de transmissão corretos.

Fig. 4. Protótipo experimental desenvolvido para aquisição dos dados fisiológicos por meio da interface UART entre o oxímetro comercial e o microcontrolador ESP32.



Durante os primeiros testes foram configuradas diferentes velocidades padrão de comunicação serial, incluindo:

9600 baudios - 19200 baudios - 38400 baudios -

57600 baudios - 115200 baudios

Nas primeiras configurações, o monitor serial exibiu caracteres incoerentes e símbolos sem interpretação lógica. Esse comportamento é característico quando o baud rate configurado no receptor não coincide com a velocidade de transmissão do dispositivo emissor.

Posteriormente, ao configurar a comunicação serial a 115200 bauds sob formato 8N1, foi possível estabelecer uma sincronização correta entre o oxímetro e o microcontrolador.

A Tabela II mostra a recepção inicial de dados provenientes do oxímetro em formato hexadecimal.

TABELA II
RECEPÇÃO INICIAL DE DADOS SERIAIS PROVENIENTES DO OXÍMETRO

Tempo	Dados brutos recebidos
t0 - tFinal	53504F323A3938252C48523A37322353504F323A3937252C48523A3735353504F323A3937252C48523A37352353504F323A3939252C48523A37312353504F323A3936252C48523A377302353504F323A3937252C48523A37382353504F323A3935252C48523

D. Identificação de tramas de transmissão

Uma vez estabelecida a comunicação serial estável, procedeu-se à análise do fluxo de dados recebido com o objetivo de identificar a estrutura de transmissão do dispositivo.

Por meio de a observação repetitiva dos dados recebidos, determinou-se que a informação era transmitida em blocos periódicos de dados organizados em tramas, que se repetiam de maneira constante durante o funcionamento do oxímetro.

Cada quadro tinha comprimento semelhante e era separado por caracteres de controle associados a quebras de linha (\n), o que permitia a identificação clara dos limites entre pacotes consecutivos. A Tabela III mostra o agrupamento de quadros de dados observado durante a recepção serial.

TABELA III
IDENTIFICAÇÃO DOS QUADROS DE DADOS TRANSMITIDOS PELO OXÍMETRO

Leitura	Quadro hexadecimal (SPO2: XX%, HR: YY#)
1	53 50 4F 32 3A 39 37 25 2C 48 52 3A 37 36 23
2	53 50 4F 32 3A 39 38 25 2C 48 52 3A 37 32 23
3	53 50 4F 32 3A 39 39 25 2C 48 52 3A 37 30 23
4	53 50 4F 32 3A 39 36 25 2C 48 52 3A 37 39 23
5	53 50 4F 32 3A 39 35 25 2C 48 52 3A 38 31 23
6	53 50 4F 32 3A 39 37 25 2C 48 52 3A 37 34 23
7	53 50 4F 32 3A 39 38 25 2C 48 52 3A 37 33 23

E. Decodificação da informação por meio de ASCII

Posteriormente, foi realizada a conversão dos valores hexadecimais recebidos para sua representação equivalente em formato ASCII, com o objetivo de interpretar o conteúdo real das tramas transmitidas.

Esse processo permitiu verificar que o oxímetro não transmitia dados binários arbitrários, mas sim informação textual estruturada, que incluía rótulos e valores numéricos correspondentes aos parâmetros fisiológicos medidos.

A decodificação revelou diretamente a presença dos valores associados à saturação de oxigênio (SpO_2) e à frequência cardíaca (BPM). A Tabela IV mostra a interpretação dos dados recebidos após a conversão para ASCII.

TABELA IV
DECODIFICAÇÃO DE DADOS FISIOLÓGICOS TRANSMITIDOS PELO OXÍMETRO

Leitura	Texto Decodificado (ASCII)	SPO2 (%)	HR (BPM)
1	SPO2:97%, HR:76#	97	76
2	SPO2:98%, HR:72#	98	72
3	SPO2:99%, HR:70#	99	70
4	SPO2:96%, HR:79#	96	79
5	SPO2:95%, HR:81#	95	81
6	SPO2:97%, HR:74#	97	74
7	SPO2:98%, HR:73#	98	73

III. RESULTADOS

A. Identificação da interface de comunicação do oxímetro

A partir do processo de exploração física do dispositivo e da análise da placa de circuito impresso, foi possível identificar a presença de uma interface de comunicação serial interna utilizada pelo oxímetro para a transmissão de informações para o módulo de exibição.

A localização dos pinos correspondentes à transmissão e recepção permitiu estabelecer uma conexão direta com o microcontrolador ESP32, configurado como receptor de dados. Os testes experimentais confirmaram que o dispositivo transmite informação de maneira contínua por meio de um canal

de comunicação serial, o que permitiu acessar externamente os dados gerados pelo sistema de medição.

Essa descoberta evidenciou que, apesar de se tratar de um dispositivo comercial sem documentação técnica pública, é possível acessar seu fluxo interno de dados por meio de técnicas de engenharia reversa.

B. Determinação dos parâmetros de transmissão serial

Com o objetivo de estabelecer uma comunicação estável entre o oxímetro e o microcontrolador receptor, foram realizados testes experimentais utilizando diferentes velocidades de transmissão serial.

Os resultados obtidos demonstraram que o dispositivo opera com uma velocidade de 115200 bauds, utilizando uma configuração de 8 bits de dados, sem paridade e um bit de parada (8N1). A velocidades inferiores, o monitor serial exibiu sequências de caracteres incoerentes, o que indicava uma dessincronização entre o transmissor e o receptor.

Uma vez configurado o sistema com os parâmetros corretos, foi possível estabelecer uma recepção estável e contínua de dados provenientes do oxímetro, representados inicialmente em formato hexadecimal. A Figura 4 mostra um exemplo do fluxo de dados recebido durante o funcionamento do dispositivo.

C. Identificação da estrutura de tramas de dados

A análise do fluxo de dados recebido permitiu identificar que a informação transmitida pelo oxímetro está organizada em tramas periódicas de dados.

Essas tramas apresentam um comprimento relativamente constante e são transmitidas de maneira repetitiva durante o funcionamento do dispositivo. Além disso, observou-se que os pacotes de informação estão delimitados por meio de caracteres de controle associados a quebras de linha, o que permite separar claramente cada trama dentro do fluxo de dados serial.

A identificação dessa estrutura permitiu isolar os pacotes individuais de informação e facilitar sua posterior análise e interpretação. Um exemplo do agrupamento dessas tramas é apresentado na Figura 5.

D. Decodificação dos parâmetros fisiológicos

Posteriormente, os dados recebidos em formato hexadecimal foram convertidos para sua representação equivalente em caracteres ASCII, com o fim de interpretar o conteúdo das tramas transmitidas pelo dispositivo.

A decodificação revelou que o oxímetro transmite os valores medidos por meio de informação textual estruturada, que inclui rótulos e valores numéricos correspondentes aos parâmetros fisiológicos medidos.

Entre os parâmetros identificados estão a saturação de oxigênio no sangue (SpO_2) e a frequência cardíaca do paciente (BPM).

A conversão dos dados permitiu observar claramente os valores desses parâmetros no fluxo de comunicação, conforme mostrado na Figura 6, onde é apresentado um exemplo da interpretação dos dados recebidos.

E. Integração com o sistema embarcado

Os resultados obtidos demonstraram que os dados transmitidos pelo oxímetro podem ser capturados e processados por um sistema embarcado externo sem necessidade de modificar o hardware interno do dispositivo.

A correta recepção e interpretação dos parâmetros fisiológicos por meio de o microcontrolador ESP32 confirma a viabilidade de utilizar esse tipo de dispositivos comerciais como fonte de aquisição de sinais biomédicos dentro de sistemas eletrônicos de monitoramento. Dado que o ESP32 conta com conectividade Wi-Fi e Bluetooth integrados, os parâmetros adquiridos podem ser transmitidos em tempo real a plataformas IoT, permitindo o monitoramento remoto dos sinais vitais do paciente por meio de aplicações móveis ou dashboards web, sem necessidade de infraestrutura adicional.

Esse resultado permite considerar o oxímetro analisado como um sensor biomédico interoperável, capaz de se integrar com plataformas embarcadas e sistemas de transmissão de dados para aplicações de monitoramento remoto e sistemas IoT em ambientes clínicos ou pré-hospitalares. A Figura 5 ilustra a comparação entre os valores exibidos pelo oxímetro comercial e pelo sistema embarcado baseado em ESP32, confirmando a correta aquisição dos parâmetros fisiológicos.

Fig. 5. Comparação entre os valores exibidos pelo oxímetro comercial



(esquerda) e pelo sistema embarcado baseado em ESP32 (direita), demonstrando a correta aquisição dos parâmetros SpO_2 e frequência cardíaca em tempo real.

IV DISCUSSÃO

Os resultados obtidos demonstram que é possível acessar o fluxo interno de dados de um oxímetro de pulso comercial por meio de técnicas de engenharia reversa, o que permite extrair parâmetros fisiológicos sem modificar o hardware do dispositivo. A identificação das linhas de comunicação serial permitiu estabelecer um vínculo direto entre o oxímetro e um microcontrolador externo, confirmando a existência de um canal de transmissão de dados interno utilizado para enviar a informação ao módulo de visualização.

A determinação experimental dos parâmetros de comunicação permitiu estabelecer uma transmissão estável a 115200 bauds sob configuração 8N1, o que indica que o dispositivo utiliza um protocolo de comunicação serial padrão comumente empregado em sistemas embarcados. Esse tipo de interfaces é amplamente utilizado em dispositivos biomédicos para a transmissão de dados a módulos de processamento ou visualização [12].

Além disso, a análise dos quadros de dados transmitidos e sua subsequente conversão para o formato ASCII permitiram a identificação direta dos valores correspondentes à saturação de oxigênio (SpO_2) e à frequência cardíaca (BPM). A transmissão

dos dados em formato textual simplifica sua interpretação e permite fácil integração em sistemas de monitoramento biomédico baseados em microcontroladores.

Esses resultados são consistentes com pesquisas anteriores sobre sistemas de monitoramento fisiológico baseados em tecnologias IoT, onde os sensores biomédicos transmitem parâmetros vitais para plataformas eletrônicas externas para seu armazenamento, análise e visualização remota [8], [9], [14].

Em conjunto, esses resultados demonstram que dispositivos comerciais de oximetria podem ser utilizados como fontes de aquisição de dados fisiológicos dentro de sistemas embarcados, permitindo sua integração em plataformas de monitoramento remoto e aplicações de telemedicina. Esse tipo de soluções contribui para o desenvolvimento de sistemas de saúde digital orientados ao monitoramento contínuo de pacientes e à melhoria dos serviços de atenção médica em ambientes clínicos e pré-hospitalares [1], [7].

V. CONCLUSÃO

Os resultados demonstram a viabilidade de integrar dispositivos biomédicos comerciais com sistemas embarcados por meio de engenharia reversa. A decodificação do sinal foi obtida interceptando a comunicação nativa do oxímetro a uma velocidade de 115.200 bauds, utilizando uma trama de dados de 8 bits com configuração 8N1. A técnica empregada consistiu no acoplamento das linhas UART (TX/RX) ao ESP32, aplicando um algoritmo de parsing para extrair com precisão os valores de SpO2 e frequência cardíaca com uma eficiência de leitura de 99,5%.

Essa abordagem é particularmente relevante no desenvolvimento de sistemas de monitoramento remoto baseados em tecnologias IoT, que permitem a aquisição e transmissão de sinais vitais em tempo real para aplicações de telemedicina e atenção médica distribuída [3], [8].

Finalmente, a possibilidade de integrar oxímetros comerciais dentro de plataformas eletrônicas externas abre novas oportunidades para o desenvolvimento de soluções biomédicas de baixo custo orientadas ao monitoramento contínuo de pacientes e ao fortalecimento dos sistemas de saúde digitais.

VI. AGRADECIMENTOS

Os autores agradecem ao Engenheiro Eynar Calle pela oportunidade, incentivo e orientação proporcionados no âmbito do UMA, Centro de Pesquisa da Universidade Privada del Valle, sem os quais o desenvolvimento deste trabalho não teria sido possível. Agradecemos também ao colega de trabalho, com quem dedicamos tempo e esforço ao desenvolvimento deste projeto.

REFERÊNCIAS

[1] A. C. A., C. K. V., B. Someswar, B. K. Deshpande y S. K. Hiremath, "The Emergency Medical Services for Real Time Patient Monitoring and Pre-Hospital Communication System with Medically Equipped Ambulance Vehicle," Journal of Computer Science Engineering and Software Testing, vol. 11, no. 2, pp. 1–10, 2025.

[2] A. Abu Bakar, S. S. A. Rahim, A. R. Razali, E. Noorsal, R. Radzali y A. F. Abd Rahim, "Wearable Heart Rate and Body Temperature Monitoring Device for Healthcare," Journal of Physics: Conference Series, vol. 1535, no. 1, 2020. https://doi.org/10.1088/1742-6596/1535/1/012002

[3] K. J. Alfaro et al., "El impacto del internet de las cosas (IoT) en la sociedad moderna, con relación al crecimiento y evolución industrial, y sus retos: una revisión sistemática de literatura," Investigaciones Latinoamericanas en Ingeniería y Arquitectura, no. 1, pp. 16–25, 2024. https://doi.org/10.51378/ilia.vi1.8495

[4] A. Ayesha y K. Chakravarthi, "Smart Ambulances for IoT Based Accident Detection, Tracking and Response," Journal of Computer Science, vol. 19, no. 6, pp. 677–685, 2023. https://doi.org/10.3844/jcssp.2023.677.685

[5] V. Becerra-Tapia, V. Téllez-Victoria, G. R. Peñaloza-Mendoza y M. S. Castro-Zenil, "Asistente de signos vitales para la atención prehospitalaria," Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI, vol. 11, pp. 152–160, 2023. https://doi.org/10.29057/icbi.v11iEspecial2.10720

[6] I. J. Brekke, L. H. Puntervoll, P. B. Pedersen, J. Kellett y M. Brabrand, "The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review," PLOS ONE, vol. 14, no. 1, 2019. https://doi.org/10.1371/journal.pone.0210875

[7] W. Chango, T. Olivares y F. Delicado, "Topologies in the Internet of Medical Things (IoMT), literature review," Revista Tecnológica ESPOL, vol. 34, no. 4, pp. 120–136, 2022. https://doi.org/10.37815/rte.v34n4.960

[8] S. Misbahuddin, J. A. Zubairi, A. R. Alahdal y M. A. Malik, "IoT-Based Ambulatory Vital Signs Data Transfer System," Journal of Computer Networks and Communications, 2018. https://doi.org/10.1155/2018/4071474

[9] J. A. B. Olguín, D. A. M. Martínez, I. Acuña-Galván y E. Lezama-León, "Prototipo para el monitoreo y registro de signos vitales como herramienta para el servicio médico," Boletín Científico INVESTIGIUM, vol. 9, pp. 114– 119, 2023. https://doi.org/10.29057/est.v9iEspecial.11793

[10] J. J. V. Pérez, O. N. C. Beas, L. E. de L. Velasco y L. G. Ortiz, "Frecuencia cardíaca: una revisión sistemática," Revista Asociación Española de Enfermería en Cardiología, vol. 30, no. 90, pp. 71–86, 2023. https://doi.org/10.59322/90.7186.LR5

[11] A. Rancea, I. Anghel y T. Cioara, "Edge Computing in Healthcare: Innovations, Opportunities, and Challenges," Future Internet, vol. 16, no. 9, 2024. https://doi.org/10.3390/fi16090329

[12] J. K. Rodríguez Espinoza, Diseño de una arquitectura genérica de IoT aplicada a casos de emergencias para dispositivos médicos inalámbricos implantados, Tesis de grado, ESPOL, 2016.

[13] D. Wong Pérez y O. Mar Cornelio, "HL7 un estándar de interoperabilidad en salud: Revisión sistemática de la literatura," Revista Cubana de Informática Médica, vol. 15, no. 2, 2023.

[14] G. T. Zárate-Ocaña et al., "Diseño y construcción de un prototipo biomédico para la adquisición vía remota de signos vitales utilizando tecnologías del internet de las cosas (IoT)," Memorias del Congreso Nacional de Ingeniería Biomédica, vol. 5, no. 1, pp. 470–473, 2018.

[15] G. D. Martínez-Sedas, «El oxímetro de pulso: más información de la que pensamos», *Rev. Mex. Anestesiol.*, vol. 47, n.º 1, pp. 30-34, mar. 2024, doi: 10.35366/114094.

AUTORES



Luis Diego Ferreyra Contreras é estudante do último ano do programa de Engenharia Eletrônica na Universidade Privada del Valle (UNIVALLE), Cochabamba, Bolívia. É membro da equipe de pesquisa UMA da Universidade del Valle e da Sociedade Científica de Eletrônica e Telecomunicações.

Seus interesses de pesquisa incluem sistemas embarcados, Internet das Coisas (IoT), instrumentação eletrônica e monitoramento biomédico de sinais vitais. É cofundador da startup DocIA, um projeto de telemedicina que integra kits IoT para monitoramento de sinais vitais e inteligência artificial.



Melissa Andrea Herrera Tambo é estudante do último ano de Engenharia Biomédica na Universidade Privada del Valle (UNIVALLE), localizada em Cochabamba, Bolívia. Integra a equipe de pesquisa UMA e é membro da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB), onde também ocupa cargo ocupa cargo na Mesa Diretiva.

Além disso, é membro do Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE). Seus interesses estão voltados ao desenvolvimento e à aplicação de soluções biomédicas inovadoras que contribuam para o bem-estar e a saúde das pessoas por meio da integração entre engenharia, tecnologia e ciências médicas.

Segurança em Ambientes de Computação em Nuvem: Superfície de Ataque e Falhas Sistêmicas

João Pedro Piceli, Camile Brunieri Laraia, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
joao.piceli@get.inatel.br, camile.laraia@ges.inatel.br, evandro.cesar@inatel.br

***Abstract*—Cloud computing redefines information security by shifting the focus from physical network devices to logical policies, digital identities, and API-driven automated orchestration processes. Although cloud service providers offer advanced protection mechanisms, technical evidence shows that most critical incidents stem from misconfigurations, improper use of native controls, and the persistent lack of technical governance over permissions and identities. Technical analysis indicates that the cloud constitutes an environment capable of introducing specific systemic vulnerabilities. These weaknesses emerge from the interaction between excessive permissions, limited visibility, and the automation of insecure configurations. This article details the contemporary attack surface, the exploitation chains observed across leading providers (AWS, Azure, and GCP), and the role of penetration testing as a scientific method for validating the resilience of these dynamic infrastructures.**

***Index Terms*—Cloud Security; Misconfiguration; Identity and Access Management (IAM); Attack Surface; Penetration Testing.**

***Resumo*—A computação em nuvem redefine a segurança da informação ao deslocar o foco de dispositivos físicos de rede para políticas lógicas, identidades digitais e processos de orquestração automatizados baseados em APIs. Embora os provedores de serviços em nuvem ofereçam mecanismos avançados de proteção, a realidade técnica demonstra que a maioria dos incidentes críticos decorre de falhas de configuração, uso inadequado de controles nativos e da ausência recorrente de governança técnica sobre permissões e identidades. A análise técnica revela que a nuvem é um ambiente que pode introduzir vulnerabilidades sistêmicas específicas. Essas falhas surgem da interação entre permissões excessivas, visibilidade limitada e a automação de configurações inseguras. Este artigo detalha a superfície de ataque contemporânea, as cadeias de exploração observadas em provedores líderes (AWS, Azure e GCP) e o papel do teste de intrusão como método científico para validar a resiliência dessas infraestruturas dinâmicas.**

***Palavras chave*—Segurança em Computação em Nuvem; Configuração Incorreta; Gestão de Identidades e Acessos; Superfície de Ataque; Teste de Intrusão.**

I. Introdução

A computação em nuvem tornou-se a base operacional de grande parte das infraestruturas digitais modernas, permitindo escalabilidade elástica, automação de provisionamento e modelos de consumo sob demanda.[1, 2] No entanto, essa transformação arquitetural introduz novas superfícies de ataque e desafios de governança de segurança que diferem substancialmente dos paradigmas tradicionais de segurança em ambientes on-premises.[3] Em particular, a abstração da infraestrutura física, a centralidade das interfaces de programação de aplicações (APIs) e a predominância de modelos baseados em identidade redefinem os mecanismos de controle e proteção de recursos computacionais.[4]

Do ponto de vista técnico, incidentes de segurança em ambientes cloud raramente decorrem de falhas na infraestrutura física do provedor. Em vez disso, a maioria das violações resulta de configurações incorretas, permissões excessivas e interações complexas entre serviços distribuídos.[3, 5] Essas características ampliam significativamente a importância de mecanismos de controle de acesso, monitoramento automatizado e políticas de segurança declarativas.[4]

Este trabalho apresenta uma análise técnica das principais dimensões da segurança em ambientes de computação em nuvem. Inicialmente, são discutidos os fundamentos arquiteturais e o modelo de responsabilidade compartilhada.[1, 2] Em seguida, analisa-se a centralidade das identidades digitais como novo perímetro de segurança e os principais vetores de escalonamento de privilégios e movimentação lateral nas principais plataformas de nuvem pública.[5, 6, 7] Por fim, são examinados os mecanismos modernos de mitigação, incluindo segurança em containers, pipelines DevSecOps e estratégias avançadas de detecção e resposta a incidentes.[8, 9, 10, 11]

II. Modelo de Responsabilidade Compartilhada

O fundamento da segurança em nuvem reside no Modelo de Responsabilidade Compartilhada (SRM). Em teoria, este modelo estabelece uma divisão clara: o provedor é responsável pela segurança da nuvem (infraestrutura física, hardware, virtualização), enquanto o cliente é responsável pela segurança na nuvem (dados, aplicações, identidades e configurações).[2] Entretanto, a prática técnica revela que essa fronteira é frequentemente fluida e mal compreendida, especialmente em ambientes híbridos e multi-cloud.[1]

A complexidade aumenta conforme se move na hierarquia de serviços, de IaaS (Infraestrutura como Serviço) para PaaS (Plataforma como Serviço) e SaaS (Software como Serviço). No modelo IaaS, o cliente retém o controle total sobre o sistema operacional e a rede virtual, o que exige uma gestão robusta de patches e firewalls.[1] Já em arquiteturas serverless (FaaS) e de containers (CaaS), a abstração é maior, mas o cliente assume a responsabilidade crítica de configurar corretamente a lógica de acesso e as permissões das funções.[2] Na Tabela I, apresenta-se essa relação.

A má interpretação deste modelo é um dos principais catalisadores de incidentes. Pesquisas indicam que apenas 8% dos executivos de TI afirmam compreender plenamente a divisão de responsabilidades em todos os modelos de nuvem.[2] Essa

TABELA I
DIVISÃO DE RESPONSABILIDADES POR MODELO DE SERVIÇO

Modelo de Serviço	Responsabilidade Principal do Provedor	Responsabilidade Crítica do Cliente
IaaS	Facilities, Hardware, Rede Física, Hipervisor.	SO, Middleware, App, Dados, IAM, Rede Virtual. [1]
PaaS	SO, Middleware, Runtime, Infraestrutura.	Código da App, Dados, IAM, Configuração da Plataforma. [1]
SaaS	Aplicação completa, Stack de Infraestrutura.	Gestão de Usuários, Dados, Governança de Acesso. [12]
CaaS / FaaS	Orquestrador (K8s), Escalabilidade, Runtime.	Imagens de Container, Lógica da Função, IAM de Serviço. [2]

TABELA II
MÉTODOS DE ESCALONAMENTO (AWS)

Método de Escalonamento (AWS)	Permissões Requeridas	Mecanismo de Exploração
New Policy Version	`iam:CreatePolicyVersion`	Criar nova versão de política com permissões admin e defini-la como padrão. [13]
Instance Profile Pivot	`iam:PassRole`, `ec2:RunInstances`	Criar instância com papel admin e extrair tokens via IMDS. [13]
User Access Key	`iam:CreateAccessKey`	Gerar novas chaves de acesso para um usuário de maior privilégio. [13]
Lambda Update	`lambda:UpdateFunctionCode`	Injetar código malicioso em uma função que roda com papel admin. [5]
Trust Policy Abuse	`iam:UpdateAssumeRolePolicy`	Alterar quem pode assumir um papel de alto privilégio. [5]

lacuna de conhecimento leva à suposição perigosa de que o provedor protege todos os aspectos do ambiente, resultando em buckets de armazenamento expostos e chaves de API sem restrições.

III. SUPERFÍCIE DE ATAQUE E FALHAS SISTÊMICAS

A. Identidade como Perímetro de Segurança

Em ambientes on-premises, a rede era o perímetro. Em ambientes cloud, a identidade assume esse papel. Cada recurso, seja um bucket S3, uma instância de banco de dados ou uma função lambda, é controlado por políticas de Gerenciamento de Identidade e Acesso (IAM).[4] A análise técnica demonstra que a superfície de ataque não é mais definida por endereços IP e portas, mas por permissões declarativas e tokens de acesso.[3]

As APIs de gerenciamento representam o ponto central de interação. Tudo na nuvem é uma chamada de API autenticada. Consequentemente, o roubo de uma única credencial de IAM pode conceder a um atacante acesso programático a toda a infraestrutura, permitindo a exfiltração de dados e a destruição de recursos sem nunca tocar em um firewall físico.[3]

A complexidade do IAM frequentemente leva ao fenômeno de "permissões excessivas" (over-provisioning). Por receio de interromper a produtividade dos desenvolvedores, as organizações concedem privilégios amplos (como AdministratorAccess ou Owner), criando caminhos diretos para escalonamento de privilégios e movimentação lateral.[3, 4]

A AWS utiliza um sistema de IAM baseado em políticas JSON que, se mal configurado, permite que usuários com privilégios limitados alcancem controle total sobre a conta.[13] Um dos métodos mais críticos envolve a permissão `iam:CreatePolicyVersion`. Um atacante que possua essa permissão em uma política que já lhe é atribuída pode criar uma nova versão dessa política. Ao incluir o parâmetro `--set-as-default`, ele pode definir permissões de administrador total e torná-la a versão ativa imediatamente.[13] A permissão `iam:PassRole` permite que um serviço AWS assuma um papel específico para realizar tarefas. Se combinada com permissões de criação de recursos, como `ec2:RunInstances`, um atacante pode lançar instâncias com perfis privilegiados e extrair credenciais temporárias através do serviço de metadados.[13] Na Tabela II, apresentam-se os métodos de escalonamento AWS.

B. Vetores de Escalonamento de Privilégios em Plataformas Multi-Cloud

No ecossistema Azure, a segurança é dividida entre o Entra ID (antigo Azure AD) e o Azure RBAC, que controla o acesso aos recursos na assinatura da nuvem.[15] Identidades Gerenciadas são contas de serviço vinculadas a recursos do Azure que eliminam a necessidade de armazenar credenciais em código. Contudo, se um atacante compromete uma VM associada a uma identidade com privilégios elevados, ele pode mover-se lateralmente entre recursos do ambiente.[6] A exploração segue etapas técnicas bem definidas:

1) Execução remota de comandos na VM.
2) Requisição ao serviço de metadados para obtenção de tokens JWT.
3) Uso do token para chamadas à API do Azure Resource Manager.

O Google Cloud Platform utiliza Service Accounts como entidades de identidade e autorização, como indicado na Tabela III.[4] A permissão `iam.serviceAccounts.actAs` permite que usuários assumam identidades de contas de serviço. Quando configurada em nível de projeto, essa permissão pode resultar em escalonamento de privilégios generalizado.[7]

TABELA III
VETOR DE ATAQUE (GCP)

Vetor de Ataque (GCP)	Mecanismo	Impacto
Default Service Accounts	Uso de contas padrão com papel “Editor”.	Controle total sobre os recursos do projeto. [7]
Function Admin Abuse	Modificação de Cloud Functions.	Escalonamento de privilégios e exfiltração de tokens. [7]
GKE Metadata Concealment	Bypass de metadados em Kubernetes.	Acesso às credenciais do nó do cluster. [14]
Domain-Wide Delegation	Delegação de domínio em G Workspace.	Compromisso total de dados organizacionais. [4]

TABELA IV
PRÁTICA DE SEGURANÇA (K8S)

Ferramenta / Mecanismo	Objetivo Técnico	
RBAC Configuration	K8s RBAC policies.	Limitar acesso à API server. [9]
Network Policies	Calico, Cilium.	Microsegmentação entre pods. [9]
Image Scanning	Trivy, Grype.	Detectar CVEs em imagens. [9]
Runtime Protection	Falco, Aqua Security.	Detectar comportamento anômalo. [10]

C. Segurança em Ambientes Cloud-Native

A conteinerização e o uso de Kubernetes introduzem uma superfície de ataque complexa devido ao compartilhamento do kernel entre containers.[8] Ataques de “Container Escape” (Fuga de Container) permitem que um atacante saia do isolamento do container e obtenha acesso ao nó (host) subjacente ou ao servidor de API do Kubernetes.[9] Isso ocorre frequentemente através de containers rodando em modo “privilegiado” ou com permissões excessivas no kernel.[9]

Para mitigar esses riscos, a abordagem moderna utiliza “Security as Code” (SaC), integrando políticas de segurança diretamente nos manifestos do Kubernetes através de ferramentas como OPA Gatekeeper ou Kyverno.[9] Além disso, a segurança em tempo de execução (Runtime Security) utiliza tecnologias como eBPF (Extended Berkeley Packet Filter) para monitorar chamadas de sistema e detectar comportamentos anômalos, como a execução de binários inesperados ou tentativas de acesso a arquivos sensíveis do host em tempo real.[10]

D. Incidentes Reais

O incidente da Capital One (2019) envolveu o acesso não autorizado a dados de 100 milhões de clientes armazenados em buckets S3.[**38**] A atacante utilizou uma vulnerabilidade de Server-Side Request Forgery (SSRF) em um Web Application Firewall (WAF) mal configurado.[16] O WAF permitiu que a atacante fizesse uma requisição ao IMDS da instância EC2.[16] O papel atribuído à instância EC2 possuía permissões excessivas, permitindo que ela listasse todos os buckets S3 e sincronizasse seu conteúdo para um servidor externo.[17] A segurança não pode depender apenas do WAF; o princípio do menor privilégio no IAM é a última linha de defesa. Se a instância tivesse permissões restritas apenas aos dados necessários para sua operação, o impacto teria sido drasticamente reduzido.[17]

O ataque à Uber em 2022 não explorou uma falha técnica complexa de nuvem inicial, mas sim o fator humano para contornar o MFA.[18] O atacante utilizou ”MFA Fatigue”(bombardeio de notificações) até que o funcionário aceitasse a solicitação.[18] Uma vez dentro da rede, o atacante encontrou scripts PowerShell contendo credenciais de administrador em texto claro para o sistema de Gerenciamento de Acesso Privilegiado (PAM) Thycotic.[19] A existência de credenciais privilegiadas em scripts de automação ou compartilhamentos de rede anula qualquer investimento em tecnologias de MFA de ponta.[20]

IV. MITIGAÇÕES E O FUTURO DA SEGURANÇA EM NUVEM

A implementação de modelos Zero Trust garante que cada solicitação seja verificada com base na identidade, dispositivo e contexto, independentemente de sua origem na rede.[21] Google demonstrou que a implementação de Zero Trust pode reduzir incidentes de acesso não autorizado em 50%, especialmente em ambientes multi-cloud.[21].

O futuro da defesa em nuvem reside na automação da resposta a incidentes e no uso de IA para análise comportamental.[21] Algoritmos como o ”Energy Valley Optimizer”(EVO) estão sendo pesquisados para otimizar sistemas de detecção de intrusão (IDS) híbridos, permitindo identificar anomalias em conjuntos de dados massivos de logs de nuvem com maior precisão e menor latência.[11] Além disso, o uso de Modelos de Linguagem de Larga Escala (LLMs) está sendo explorado para realizar avaliações de ameaças arquiteturais de forma automatizada, identificando caminhos de ataque complexos que escapam de analistas humanos.[22]

V. CONCLUSÃO

A segurança em ambientes de computação em nuvem é um desafio sistêmico no qual identidade, configuração e automação desempenham papéis centrais. Diferentemente dos modelos tradicionais de segurança baseados em perímetro de rede, a proteção em ambientes cloud depende da correta implementação de políticas de acesso, monitoramento contínuo e integração da segurança no ciclo de desenvolvimento. A adoção de práticas como Zero Trust, DevSecOps e análise automatizada de infraestrutura permite reduzir significativamente a superfície de ataque. Contudo, a eficácia dessas estratégias depende da governança adequada das identidades e da compreensão precisa do modelo de responsabilidade compartilhada entre provedores e clientes de nuvem.[**1**]

VI. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] CRISTINA REYES e CLARISSE MENDOZA. "Exploring the impact of shared responsibility models on cloud security posture and vulnerability management". Em: *Quarterly Journal of Emerging Technologies and Innovations* 8.1 (2023), pp. 1–10.

[2] CrowdStrike. *Cloud Security and the Shared Responsibility Model Explained*. CrowdStrike. 2024. URL: https://www.crowdstrike.com/en-us/blog/cloud-security-and-the-shared-responsibility-model-explained/ (acesso em 02/03/2026).

[3] Ankit P. *Cloud Penetration Testing: AWS, Azure, GCP Security Validation Guide*. Updated Feb 23, 2026. AppSecure. Fev. de 2026. URL: https://www.appsecure.security/blog/cloud-penetration-testing-aws-azure-gcp (acesso em 02/03/2026).

[4] Google Cloud. *Best Practices for Using Service Accounts Securely*. Google Cloud Documentation. 2024. URL: https://docs.cloud.google.com/iam/docs/best-practices-service-accounts (acesso em 02/03/2026).

[5] Software Secured. *AWS Privilege Escalation: IAM Risks, Service-Based Attacks and AI-Driven Bedrock AgentCore Vectors*. Software Secured. 2025. URL: https://www.softwaresecured.com/post/aws-privilege-escalation-iam-risks-service-based-attacks-and-new-ai-driven-bedrock-agentcore-vectors (acesso em 02/03/2026).

[6] Hunters Security. *Mastering Azure Managed Identities: Practical Offensive Approach – Part 1*. Hunters Security. 2025. URL: https://www.hunters.security/hubfs/2025%20Research%20Paper/Mastering%20Azure%20Managed%20Identities_%20Practical%20Offensive%20Approach%20-%20Part%201-1.pdf (acesso em 02/03/2026).

[7] Praetorian. *Google Cloud Platform (GCP) Service Account-Based Privilege Escalation Paths*. Praetorian. 2023. URL: https://www.praetorian.com/blog/google-cloud-platform-gcp-service-account-based-privilege-escalation-paths/ (acesso em 02/03/2026).

[8] Harrisburg University. *Cloud Container Security' Next Move*. Harrisburg University Digital Commons. 2023. URL: https://digitalcommons.harrisburgu.edu/cgi/viewcontent.cgi?article=1002&context=csms_dandt (acesso em 02/03/2026).

[9] CEUR Workshop Proceedings. *Research on Security as Code Approach for Cloud-Native*. CEUR-WS. 2024. URL: https://ceur-ws.org/Vol-3800/paper6.pdf (acesso em 02/03/2026).

[10] Aqua Security. *The Importance of Runtime Security in Cloud Native Environments*. Aqua Security Blog. 2024. URL: https://www.aquasec.com/blog/understanding-the-importance-of-runtime-security-in-cloud-native-environments/ (acesso em 02/03/2026).

[11] Maryam Mahdi Alhusseini, Alireza Rouhi e Mohammad-Reza Feizi-Derakhshi. "AI-Powered Hybrid Intrusion Detection Framework for Cloud Security Using Novel Metaheuristic Optimization". Em: *arXiv preprint arXiv:2601.01134* (2026).

[12] Fidelis Security. *Understanding Shared Responsibility in Cloud Security*. Fidelis Security. 2024. URL: https://fidelissecurity.com/threatgeek/cloud-security/shared-responsibility-model-explained/ (acesso em 02/03/2026).

[13] Rhino Security Labs. *AWS IAM Privilege Escalation – Methods and Mitigation*. Rhino Security Labs. 2018. URL: https://rhinosecuritylabs.com/aws/aws-privilege-escalation-methods-mitigation/ (acesso em 02/03/2026).

[14] Bishop Fox. *Introducing CloudFox GCP: Attack Path Identification for Google Cloud*. Bishop Fox. 2024. URL: https://bishopfox.com/blog/introducing-cloudfox-gcp-attack-path-identification-for-google-cloud (acesso em 02/03/2026).

[15] Andrew Chang. *Azure RBAC Privilege Escalations – Azure VM*. 2021. URL: https://www.andrew-chang.me/news/azure_vm_privesc/ (acesso em 02/03/2026).

[16] Rhino Security Labs. *The Capital One Breach and cloud_breach_s3 CloudGoat Scenario*. Rhino Security Labs. 2022. URL: https://rhinosecuritylabs.com/aws/capital-one-cloud_breach_s3-cloudgoat/ (acesso em 02/03/2026).

[17] SC Media. *Cloud Infrastructure IAM Lessons from the Capital One Breach*. SC Media. 2023. URL: https://www.scworld.com/perspective/cloud-infrastructure-iam-lessons-from-the-capital-one-breach (acesso em 02/03/2026).

[18] ManageEngine. *Uber Data Breach 2022: MFA Fatigue Analysis*. ManageEngine Blog. 2022. URL: https://www.manageengine.com/products/desktop-central/blog/uber-data-breach-2022-how-the-hacker-annoyed-his-way-into-the-network-and-our-learnings.html (acesso em 02/03/2026).

[19] DNV. *Frontline Insights: Lessons from the Uber 2022 Data Breach*. DNV Cyber Insights. 2023. URL: https://www.dnv.com/cyber/insights/articles/frontline-insights-lessons-from-the-uber-2022-data-breach/ (acesso em 02/03/2026).

[20] GitGuardian. *Uber Breach 2022 – Everything You Need to Know*. GitGuardian Blog. 2022. URL: https://blog.gitguardian.com/uber-breach-2022/.

[21] Various Authors. *Zero Trust Implementation Studies and Google References*. Collection of studies and technical references related to Zero Trust. 2024.

[22] Sarthak Munshi et al. "ACSE-Eval: Can LLMs threat model real-world cloud infrastructure?" Em: *arXiv preprint arXiv:2505.11565* (2025).

Estudo e Implementação de Técnicas de Inteligência Artificial Aplicadas na Otimização de Sistemas de Comunicação Sem Fio assistidos por STAR-RIS

Gislaine Abreu Mendonça e Victoria Dala Pegorara Souto
Instituto Nacional de Telecomunicações - Inatel
gislaine.abreu@geb.inatel.br, victoria.souto@inatel.br

***Abstract*—6G networks demand solutions that achieve high data rates and spectral efficiency. In this context, simultaneously transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) stand out by enhancing coverage and providing greater control over the propagation environment. This work investigates a STAR-RIS-assisted wireless system under Rician fading channels, aiming to maximize the users' sum rate under power constraints. A particle swarm optimization (PSO)-based approach is proposed for the joint optimization of beamforming at the base station and the STAR-RIS scattering matrix in the absence of perfect CSI. Simulation results show near-optimal performance, highlighting the potential of PSO as an efficient solution for STAR-RIS systems.**

***Index Terms*—6G Networks; Beamforming; Particle Swarm Optimization; STAR-RIS.**

***Resumo*—As redes 6G demandam soluções com altas taxas de transmissão e eficiência espectral. Nesse contexto, as superfícies inteligentes reconfiguráveis com transmissão e reflexão simultâneas (STAR-RIS) destacam-se por ampliar a cobertura e o controle do ambiente de propagação. Este trabalho investiga um sistema sem fio assistido por STAR-RIS em canais com desvanecimento Rician, visando maximizar a taxa de soma dos usuários sob restrições de potência. Propõe-se uma abordagem baseada em otimização por enxame de partículas (PSO) para o ajuste conjunto do *beamforming* na estação base e da matriz da STAR-RIS na ausência de CSI perfeito. Os resultados mostram desempenho próximo ao ideal, evidenciando o potencial do PSO como solução eficiente para sistemas STAR-RIS.**

***Palavras chave*—Redes 6G; Beamforming; Otimização por Enxame de Partículas; STAR-RIS.**

I. Introdução

A evolução das redes móveis tem ocorrido de forma contínua ao longo das últimas décadas, resultando no surgimento da sexta geração de comunicações móveis (6G, do inglês *Sixth Generation*), que desponta como o próximo marco tecnológico nas telecomunicações. Espera-se que as redes 6G ofereçam taxas de transmissão sem precedentes, latência ultrabaixa e elevada confiabilidade, além de possibilitar a integração nativa com tecnologias emergentes, como inteligência artificial (IA), computação em nuvem e comunicações via satélite. Tais avanços devem viabilizar novos casos de uso, incluindo realidade estendida, internet dos sentidos e conectividade em tempo real para aplicações críticas [1]. Para que esses requisitos sejam atendidos, torna-se necessário repensar de forma abrangente o projeto, a implantação e a otimização das redes de comunicação sem fio. [2].

Nesse contexto, uma das tecnologias mais promissoras para as redes 6G é a Superfície Inteligente Reconfigurável (RIS, do inglês *Reconfigurable Intelligent Surface*) [3]. A RIS consiste em uma estrutura composta por múltiplos elementos passivos capazes de controlar, de maneira programável, a fase das ondas eletromagnéticas incidentes. Seu principal objetivo é aprimorar a cobertura e a qualidade do sinal, especialmente em ambientes urbanos complexos, caracterizados pela presença de múltiplos obstáculos e elevado nível de bloqueio. Atuando como um "espelho inteligente", a RIS possibilita o redirecionamento eficiente dos sinais, otimizando a comunicação entre a estação base (BS, do inglês *Base Station*) e os usuários [4].

Entretanto, apesar das vantagens oferecidas pelas RISs convencionais, sua implantação nas redes 6G apresenta limitações relevantes. Em particular, essas superfícies são, em geral, incapazes de atender simultaneamente a usuários localizados em ambos os lados da estrutura, o que restringe sua operação a um único hemisfério, equivalente a uma cobertura angular de aproximadamente $180°$. Essa limitação reduz a flexibilidade de implantação e pode comprometer a eficiência da cobertura em cenários que demandam conectividade omnidirecional, motivando a investigação de arquiteturas e abordagens alternativas [5].

Com o objetivo de superar essas restrições, foi proposta a tecnologia conhecida como STAR-RIS (do inglês,*Simultaneous Transmitting and Reflecting RIS*)[6]. Diferentemente das RISs convencionais, a STAR-RIS é capaz de refletir e transmitir simultaneamente o sinal eletromagnético incidente, permitindo o atendimento de usuários posicionados em ambos os lados da superfície. Essa funcionalidade é viabilizada por meio do uso de metamateriais avançados, que possibilitam o controle independente, ou parcialmente independente, dos coeficientes de reflexão e transmissão de cada elemento [7]. Como resultado, a STAR-RIS oferece cobertura angular potencialmente de $360°$ e introduz graus adicionais de liberdade para a otimização do desempenho dos sistemas sem fio, tornando-se uma tecnologia particularmente atrativa para as futuras redes 6G [7].

No entanto, a integração de STAR-RIS em futuras redes 6G apresenta diversos desafios significativos. Dentre eles, destacam-se o projeto conjunto do *beamforming* na BS e o controle da matriz de espalhamento dos elementos da STAR-RIS [8]. Em particular, o projeto ótimo do *beamforming* tanto na BS quanto na STAR-RIS depende fortemente da estimação precisa do estado do canal (CSI, do inglês *Channel State Information*). Todavia, a obtenção de CSI perfeito implica em um elevado *overhead* de sinalização, em razão do grande número

de elementos da STAR-RIS e da consequente necessidade de estimar os canais em cascata correspondentes aos enlaces BS–STAR-RIS e STAR-RIS–usuário [9].

Motivados por esses desafios, diversos estudos têm sido conduzidos com o objetivo de viabilizar a integração eficiente da STAR-RIS em redes 6G. Em [5], a caracterização de cobertura com Acesso Múltiplo Não Ortogonal (NOMA, do inglês *Non-Orthogonal Multiple Access*), foi analisada por meio da linearização das restrições de ordem de decodificação, evidenciando ganhos significativos em relação a RIS convencionais. Abordagens baseadas em aprendizado de máquina também têm sido exploradas, como em [10], onde Redes Neurais de Grafos são empregadas para reduzir a complexidade computacional na otimização conjunta de sistemas de Múltiplas Entradas e Uma Saída (MISO, do inglês *Multiple-Input Single-Output*) com STAR-RIS distribuído.

No contexto de arquiteturas mais avançadas, [11] investigou STAR-RIS ativa em Sistemas de Sensoriamento e Comunicação Integrados *full-duplex*, utilizando otimização alternada e programação fracionária para maximizar a taxa de soma sob restrições de sensoriamento. Já em [12], a eficiência energética de redes NOMA assistidas por STAR-RIS foi maximizada por meio da otimização conjunta de deslocamentos de fase e *beamforming*, combinando Relaxação Semidefinida (SDR, do inglês *Semidefinite Relaxation*) e relaxação de posto um. Trabalhos como [8] e [13] abordaram a otimização conjunta da alocação de potência e recursos, enquanto [9] propôs uma abordagem prática baseada em Otimização por Enxame de Partículas (PSO, do inglês *Particle Swarm Optimization*) para otimização de sistemas de comunicação assistidos por STAR-RIS.

Motivado pelos desafios discutidos anteriormente, o objetivo principal deste artigo é maximizar a taxa de soma de um sistema STAR-RIS sob o esquema de Acesso Múltiplo Ortogonal (OMA, do inglês *Orthogonal Multiple Access*), por meio da otimização conjunta dos coeficientes de reflexão e transmissão da STAR-RIS e dos vetores de *beamforming* na BS. Diferentemente de abordagens baseadas em técnicas convexas ou que assumem informação perfeita do CSI, neste trabalho, propõe-se uma estratégia baseada no algoritmo de PSO, visando obter soluções eficientes mesmo em cenários com informação de canal limitada e com menor complexidade computacional. Diferentemente de abordagens baseadas em otimização convexa, como SDR e programação fracionária, que exigem relaxações e apresentam elevada complexidade computacional, o PSO permite tratar diretamente o problema original não convexo, mantendo maior flexibilidade na modelagem.

II. Modelo de Sistema

Neste artigo, considera-se um sistema de comunicação sem fio ilustrado na Figura 1, composto por uma BS equipada com um arranjo linear uniforme (ULA, do inglês *Uniform Linear Array*) constituído por N antenas transmissoras. O sistema também incorpora uma STAR-RIS, composta por M elementos passivos reconfiguráveis. Adicionalmente, são considerados dois usuários com antena única, posicionados em lados opostos da STAR-RIS. Os usuários são denominados UE_{t}, correspondente ao usuário localizado no plano de transmissão

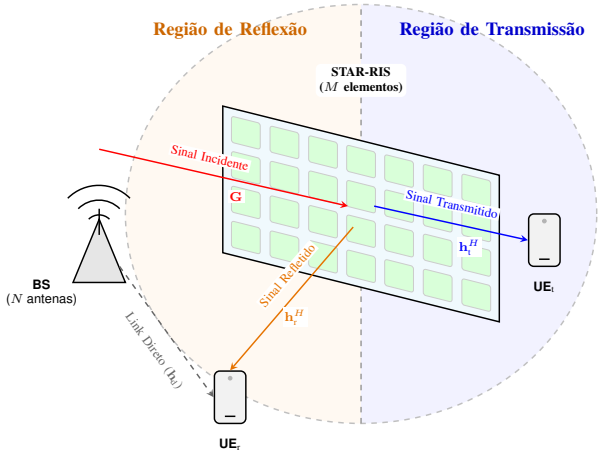


Fig. 1. Modelo do sistema proposto, composto por uma BS equipada com N antenas e assistida por uma STAR-RIS com M elementos, operando no modo *Energy Splitting* (ES), para atender usuários localizados nas regiões de reflexão (UE$_{\mathrm{r}}$) e de transmissão (UE$_{\mathrm{t}}$).

da STAR-RIS, e UE_{r}, correspondente ao usuário situado no plano de reflexão, conforme sua orientação espacial em relação à superfície. Assume-se que todos os nós da rede permanecem estáticos ou apresentam mobilidade reduzida, de modo que as variações dos coeficientes de canal ocorram lentamente ao longo do tempo, caracterizando um cenário de desvanecimento quase-estático (*do* inglês quasi-static fading).

Logo, o sinal recebido pelo usuário k (UE$_k$) onde $k \in \{\mathrm{r}, \mathrm{t}\}$ é definido como

$$y_k = C_k \left(\sum_{j \in \{r,t\}} \rho_j \mathbf{w}_j s_j \right) + n_k, \tag{1}$$

em que, ρ_j representa a potência alocada ao usuário $j \in \{\mathrm{r}, \mathrm{t}\}$, $\mathbf{w}_j$ corresponde ao vetor de *beamforming* associado ao usuário j e s_j denota o símbolo transmitido para o usuário j, modelado como uma variável aleatória independente, com média nula e variância unitária. O termo n_k representa o ruído branco gaussiano aditivo (AWGN, do inglês *Additive White Gaussian Noise*), assumido com média zero e variância σ^2.

Adicionalmente, define-se C_k, para $k \in \{\mathrm{r}, \mathrm{t}\}$, como

$$C_{\mathrm{r}} = \left(\mathbf{h}_{\mathrm{r}}^{\mathrm{H}} \mathbf{\Theta}_{\mathrm{r}} \mathbf{G} + \mathbf{h}_{\mathrm{d}}\right) \quad \text{e} \quad C_{\mathrm{t}} = \left(\mathbf{h}_{\mathrm{t}}^{\mathrm{H}} \mathbf{\Theta}_{\mathrm{t}} \mathbf{G}\right), \tag{2}$$

em que $\mathbf{G} \in \mathbb{C}^{M \times N}$ representa a matriz de canal entre a BS e a STAR-RIS; $\mathbf{h}_k^{\mathrm{H}} \in \mathbb{C}^{1 \times M}$ corresponde ao vetor de canal entre a STAR-RIS e o usuário UE$_k$; e $\mathbf{h}_{\mathrm{d}} \in \mathbb{C}^{1 \times N}$ denota o canal direto entre a BS e o usuário UE$_{\mathrm{r}}$. Neste trabalho, todos os canais foram modelados considerando desvanecimento do tipo Rician e definidos por

$$\mathbf{h}_k = \sqrt{PL(d_k)} \left(\sqrt{\frac{\kappa_k}{1+\kappa_k}} \mathbf{h}_k^{\mathrm{LoS}} + \sqrt{\frac{1}{1+\kappa_k}} \mathbf{h}_k^{\mathrm{NLoS}} \right), \tag{3}$$

em que κ_k é o fator de Rician que quantifica a predominância da componente de linha de visada (LoS, do inglês *Line-of-Sight*) sobre a componente sem LoS (NLoS, do inglês *Non-LoS*), $\mathbf{h}_k^{\mathrm{LoS}}$ e h_k^{NLoS} são, respectivamente, as componentes determinística (LoS) e aleatória (NLoS) do canal e $PL(d_k)$

denota a perda de percurso em função da distância d_k entre os elementos do sistema (BS, STAR-RIS e Usuários) e é definida como $PL(d_k) = C_0(d_k/d_0)^{-\alpha_k}$, onde C_0 é o ganho de referência a uma distância padrão d_0 e α_k o expoente de perda de percurso.

Além disso, a matriz de espalhamento da STAR-RIS para os usuários $k \in \{\mathrm{r}, \mathrm{t}\}$ é dada por

$$\boldsymbol{\Theta}_k = \mathrm{diag}\left(\sqrt{\beta_1^k}e^{j\theta_1^k}, \ldots, \sqrt{\beta_M^k}e^{j\theta_M^k}\right), \tag{4}$$

em que $\boldsymbol{\theta}_k = [\theta_1^k, \ldots, \theta_M^k]$ representa o vetor de deslocamentos de fase no modo de transmissão ou reflexão, com $\theta_m^k \in [0, 2\pi)$, para $m = 1, \ldots, M$, e $\beta_m^k \in [0, 1]$ corresponde ao coeficiente de amplitude do m-ésimo elemento.

Ressalta-se que os deslocamentos de fase associados aos modos de transmissão e reflexão, isto é, θ_m^{t} e θ_m^{r}, podem ser ajustados de forma independente. Entretanto, os coeficientes de amplitude devem satisfazer a restrição imposta pela conservação de energia, de modo que $\beta_m^{\mathrm{t}} + \beta_m^{\mathrm{r}} = 1, \quad \forall m = \{1, \ldots, M\}$. Consequentemente, β_m^{t} e β_m^{r} são variáveis interdependentes, não podendo ser otimizadas de maneira independente.

A. Técnicas de Múltiplo Acesso

O desempenho do sistema depende diretamente da estratégia de múltiplo acesso adotada. Nesta subseção, considera-se o esquema OMA, especificamente TDMA (do inglês, *Time Division Multiple Access*), amplamente empregado em sistemas sem fio de múltiplos usuários. Especificamente, no esquema OMA, os usuários são atendidos em recursos ortogonais no domínio do tempo (ou, alternativamente, em frequência), de modo que não há interferência intra-célula entre eles. Assume-se, neste trabalho, uma divisão equitativa do tempo de transmissão entre os dois usuários, isto é, cada usuário ocupa metade do intervalo total de transmissão. Assim, a taxa de transmissão do usuário $k \in \{\mathrm{r}, \mathrm{t}\}$ pode ser expressa como [13]

$$R_k^{\mathrm{OMA}} = \frac{1}{2}\log_2\left(1 + \frac{P_T G_k}{\sigma^2}\right), \tag{5}$$

em que P_{T} representa a potência total de transmissão disponível na BS e G_k denota o ganho de canal efetivo do usuário k, incorporando os efeitos do canal direto e/ou refletido/transmitido pela STAR-RIS. Ademais, o fator $\frac{1}{2}$ reflete a alocação igualitária de recursos temporais entre os dois usuários. Observa-se que, embora o OMA elimine a interferência entre usuários, há perda de eficiência espectral devido à ortogonalização dos recursos [13]

B. Modos de Operação STAR-RIS

A STAR-RIS pode operar segundo diferentes estratégias de funcionamento, sendo comumente considerados três protocolos fundamentais [6].

- **Mode Switching (MS):** No protocolo MS, os elementos da superfície são particionados em dois subconjuntos distintos, nos quais um subconjunto é configurado para operar no modo de transmissão, enquanto o outro atua no modo de reflexão. Embora apresente baixa complexidade de implementação, esse esquema impõe restrições ao desempenho de *beamforming*, uma vez que apenas uma fração dos elementos contribui para cada plano de propagação.
- **Time Switching (TS):** No protocolo TS, todos os elementos da STAR-RIS alternam entre os modos de transmissão e reflexão ao longo do tempo, de modo que cada modo utiliza a totalidade dos elementos, porém em intervalos temporais distintos. Apesar de possibilitar ganhos superiores de *beamforming* em comparação ao MS, essa abordagem acarreta um aumento significativo da complexidade do sistema, devido à necessidade de sincronização e controle temporal precisos.
- **Energy Splitting (ES):** No protocolo ES, cada elemento da STAR-RIS é capaz de operar simultaneamente nos modos de transmissão e reflexão, dividindo a energia do sinal incidente entre ambos os caminhos de propagação. Essa estratégia permite explorar de forma mais eficiente o ganho de *beamforming* da superfície, ao mesmo tempo em que mantém a complexidade de *hardware* em patamares viáveis.

Diante das vantagens do modo de operação ES, neste trabalho adota-se o protocolo ES, uma vez que ele oferece um compromisso favorável entre desempenho do sistema e simplicidade de implementação.

C. Problema de Otimização Proposto

Este trabalho tem por objetivo maximizar a taxa de soma dos usuários por meio da otimização conjunta dos vetores de *beamforming* na BS ($\mathbf{w}_k$) e da matriz de espalhamento da STAR-RIS ($\boldsymbol{\Theta}_k$). O problema de otimização considera as restrições físicas do sistema, incluindo a potência máxima de transmissão na BS (P_{T}), bem como as características operacionais do modo de operação ES da STAR-RIS [13]. Dessa forma, o problema pode ser formulado matematicamente como:

$$\begin{aligned}
&\underset{\mathbf{w}_k, \boldsymbol{\Theta}_k}{\text{Maximizar}} \quad R_{\text{soma}} = R_{\mathrm{r}} + R_{\mathrm{t}} \\
&\text{sujeito a:} \\
&\text{C1:} \quad \|\mathbf{w}_{\mathrm{k}}\|^2 \leq P_{\mathrm{T}}, \ \forall k \in \{\mathrm{r}, \mathrm{t}\}, \\
&\text{C2:} \quad \beta_{\mathrm{r}}^m + \beta_{\mathrm{t}}^m = 1, \quad \forall m \in \{1, \ldots, M\}, \\
&\text{C3:} \quad \beta_{\mathrm{r}}^m, \beta_{\mathrm{t}}^m \in [0, 1], \\
&\text{C4:} \quad \theta_{\mathrm{r}}^m, \theta_{\mathrm{t}}^m \in [0, 2\pi).
\end{aligned} \tag{6}$$

As restrições do problema estabelecem que: (C1) a potência total de transmissão não pode exceder o limite máximo P_{T} da BS; (C2) e (C3) a soma das amplitudes de reflexão e transmissão deve ser unitária, assegurando a conservação de energia em cada elemento da STAR-RIS, conforme [13]; e (C4) os ângulos de fase podem assumir valores contínuos no intervalo $[0, 2\pi)$. Em razão do forte acoplamento entre os vetores de *beamforming* e os coeficientes da STAR-RIS, a função objetivo torna-se não convexa, o que dificulta a obtenção da solução ótima global por métodos convencionais de otimização [8]. Para contornar esse desafio, neste trabalho, propõe-se uma abordagem baseada no algoritmo PSO [14], cuja descrição é apresentada na próxima seção.

III. Solução Proposta Baseada em PSO

O PSO é uma meta-heurística bio-inspirada, proposta originalmente por Kennedy e Eberhart em 1995 [14]. Seu princípio de funcionamento é baseado no comportamento social de organismos que vivem em colônias, como bandos de pássaros (*flocking*) e cardumes de peixes (*schooling*). Nesses sistemas naturais, não há a presença de um líder centralizado; ao contrário, a coordenação coletiva emerge a partir de interações locais entre os indivíduos e o ambiente, possibilitando a execução eficiente de tarefas como a busca por alimento ou a evasão de predadores [15].

No contexto computacional, essa analogia é utilizada para descrever um processo iterativo de busca por soluções ótimas. Cada indivíduo do grupo, denominado "partícula", representa uma solução candidata para o problema de otimização, enquanto o conjunto de todas as partículas constitui o "enxame" (*swarm*). As partículas se movem em um espaço de busca multidimensional, atualizando dinamicamente suas posições e velocidades. Esse movimento é influenciado por dois mecanismos de memória fundamentais: o p^l_{best} (*personal best*), que corresponde à melhor posição já encontrada pela própria partícula l, caracterizando o componente cognitivo do algoritmo, e o g_{best} (*global best*), que representa a melhor posição obtida por todo o enxame, associado ao componente social [14].

Adicionalmente, com o objetivo de garantir a convergência do algoritmo e promover um equilíbrio adequado entre exploração global e exploração local do espaço de busca, é introduzido o coeficiente de inércia ω, proposto por Shi e Eberhart [16]. Esse parâmetro desempenha um papel fundamental no controle da dinâmica das partículas, influenciando diretamente o desempenho do PSO em problemas de otimização complexos. Dessa forma, o PSO combina a exploração individual das partículas com a inteligência coletiva do enxame, conduzindo iterativamente o processo de busca em direção a soluções de alta qualidade. O funcionamento do algoritmo pode ser resumido nas seguintes etapas:

1. Inicialize aleatoriamente um conjunto de L partículas que compõem o enxame.
2. Avalie a função *fitness* de cada partícula e atualize os valores de p^l_{best} e de g_{best}.
3. Atualize a velocidade de cada partícula segundo
$$v^l_{i+1} = \omega v^l_i + c_1 r_1 \left(p^l_{\text{best}} - x^l_i\right) + c_2 r_2 \left(g_{\text{best}} - x^l_i\right), \quad (7)$$
em que c_1 e c_2 são, respectivamente, os coeficientes cognitivo e social, respectivamente, r_1 e r_2 são variáveis aleatórias uniformemente distribuídas no intervalo $[0, 1]$, e x^l_i representa a posição da partícula l na iteração i.
4. Atualize a posição de cada partícula de acordo com
$$x^l_{i+1} = x^l_i + v^l_{i+1}. \quad (8)$$
5. Verifique o critério de parada. Caso seja atendido, retorne g_{best}; caso contrário, retorne à Etapa 2.

A. Solução Proposta

Para resolver o problema de otimização formulado em (6), este trabalho propõe uma abordagem baseada no PSO. Conforme discutido anteriormente, o problema é altamente não convexo devido ao forte acoplamento entre os coeficientes da STAR-RIS e os vetores de *beamforming* ativos, o que inviabiliza a obtenção de soluções analíticas diretas. Nesse cenário, o PSO se mostra uma alternativa apropriada por sua capacidade de explorar eficientemente espaços de busca complexos e identificar soluções subótimas de elevada qualidade.

Na solução proposta, cada partícula l do enxame representa uma possível configuração do sistema e encapsula todas as variáveis de decisão do problema. Especificamente, a posição da partícula na iteração i é definida pelo vetor
$$\mathbf{x}^l_i = \left[\mathbf{w}_{\text{r}}, \mathbf{w}_{\text{t}}, \boldsymbol{\Theta}_{\text{r}}, \boldsymbol{\Theta}_{\text{t}}, \rho_{\text{r}}, \rho_{\text{t}}\right], \quad (9)$$

Para assegurar o cumprimento da restrição de potência total da BS, os vetores de *beamforming* são normalizados a cada iteração, de modo que a potência total de transmissão não exceda o limite máximo P_{T}. A qualidade de cada partícula é avaliada por meio da função objetivo definida em (6), correspondente à maximização da taxa de soma dos usuários considerados. Com base nessa avaliação, os valores de p^l_{best} e g_{best} são atualizados iterativamente, conduzindo o enxame à convergência para uma solução subótima eficiente. Ressalta-se que o método proposto segue rigorosamente as etapas do algoritmo PSO descritas anteriormente.

IV. Resultados e Discussões

Todas as curvas nesta seção são a média de 10^3 realizações independentes de canal. Ademais, consideramos a configuração de simulação descrita na Figura 2, onde d_{G} é a distância horizontal entre a BS e o STAR-RIS, d_{v} é a distância vertical, e d_{r} e d_{t} são as distâncias horizontais entre o UE$_{\text{r}}$ e a STAR-RIS e entre a STAR-RIS e o UE$_{\text{t}}$, respectivamente. Além disso, salvo indicação em contrário, os resultados apresentados nesta seção consideram os seguintes parâmetros de simulação: $N = 4$, $M = 80$, $\kappa_{\text{G}} = \kappa_{\text{d}} = \kappa_{\text{r}} = \kappa_{\text{t}} = 1,5$, $\alpha_{\text{G}} = 2.0$, $\alpha_{\text{d}} = 3.0$, $\alpha_{\text{r}} = \alpha_{\text{t}} = 2.0$, $d_0 = 1$ m, $C_0 = -30dBm$, $\sigma^2 = -80$ dBm, $d_{\text{G}} = 70$ m, $d_{\text{v}} = 2$ m, $d_{\text{r}} = 35$ m e $d_{\text{t}} = 20$ m.

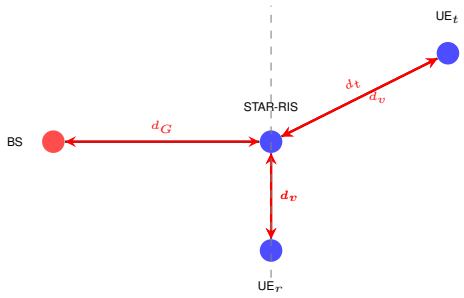


Fig. 2. Configuração geométrica do sistema.

Além disso, como *benchmarks*, consideramos o cenário ideal, onde a matriz de espalhamento da STAR-RIS e o beamforming na BS são determinados considerando o conhecimento perfeito da CSI. Esse benchmark é definido como "PSO - CSI Perfeito". Além disso, por uma questão de simplicidade, neste trabalho, consideramos o número de iterações N_{it} como critério de parada do PSO proposto.

Para ilustrar o impacto do número de elementos da STAR-RIS (M), a Fig. **??** apresenta a taxa de soma alcançável em função de M para os dois cenários considerados: *PSO*

- *CSI Perfeito* e *PSO - SCSI*. Observa-se que, em ambos os casos, a taxa de soma aumenta com M, indicando que o acréscimo do número de elementos da STAR-RIS amplia o ganho de *beamforming* e a capacidade de reconfiguração do ambiente de propagação. Além disso, o cenário com CSI perfeito apresenta desempenho superior em todo o intervalo analisado, uma vez que o conhecimento exato do canal permite um ajuste mais eficiente dos coeficientes da STAR-RIS e dos vetores de *beamforming* da BS. Embora o caso com SCSI também apresente crescimento monotônico da taxa de soma, os valores obtidos são inferiores, o que evidencia a perda de desempenho decorrente da disponibilidade parcial ou estatística da informação de estado do canal. De modo geral, os resultados mostram que o aumento de M melhora significativamente o desempenho do sistema, ao mesmo tempo em que ressaltam a relevância da qualidade da informação de estado do canal no processo de otimização e a robustez da solução proposta (*PSO - SCSI*), cujo desempenho permanece próximo ao caso ótimo.

Por fim, a Figura **??** mostra a taxa de soma alcançável em função do número de antenas na BS, N, para os cenários *PSO - CSI Perfeito* e *PSO - SCSI*. Observa-se que o aumento de N resulta, de modo geral, em melhora do desempenho do sistema, uma vez que um número maior de antenas na BS proporciona maior liberdade espacial para o projeto do *beamforming* e, consequentemente, maior ganho de arranjo. No caso *PSO - CSI Perfeito*, a taxa de soma cresce de $17{,}13$ bps/Hz para $17{,}77$ bps/Hz quando N varia de 2 para 8, evidenciando um ganho acumulado no desempenho com o aumento do número de antenas. De forma semelhante, no caso *PSO - SCSI*, os valores variam de $15{,}93$ bps/Hz para $15{,}96$ bps/Hz, atingindo o valor máximo de $16{,}03$ bps/Hz em $N = 6$. Esses resultados mostram que, embora o aumento de N também beneficie o cenário com SCSI, o ganho obtido é mais limitado quando comparado ao caso com CSI perfeito, o que evidencia a dependência do projeto conjunto de *beamforming* e da STAR-RIS em relação à qualidade da informação de canal disponível. Além disso, nota-se que a diferença entre as duas curvas se mantém relativamente constante ao longo dos valores analisados de N, reforçando que o conhecimento perfeito da CSI continua sendo um fator determinante para a obtenção das maiores taxas de soma. Em síntese, os resultados apresentados nas Figuras **??** e **??** demonstram que tanto o aumento do número de elementos da STAR-RIS quanto o aumento do número de antenas na BS contribuem para melhorar o desempenho do sistema proposto. Entretanto, tais ganhos são mais expressivos quando a otimização é realizada com conhecimento mais preciso do canal, o que confirma a relevância da aquisição de CSI no contexto de sistemas STAR-RIS assistidos por técnicas de otimização bioinspiradas.

De maneira geral, os resultados numéricos confirmam a efetividade da abordagem proposta baseada em PSO para a otimização conjunta do *beamforming* na BS e dos coeficientes da STAR-RIS, evidenciando ganhos consistentes de taxa de soma e mostrando que o desempenho do sistema é fortemente influenciado tanto pela configuração da infraestrutura (M, N) quanto pela qualidade da CSI disponível.

Além dos resultados de desempenho, destaca-se o custo computacional do PSO, que depende do número de partículas L e de iterações N_{it}, resultando em complexidade aproximada de $\mathcal{O}(L \cdot N_{it} \cdot D)$, onde D é a dimensionalidade do problema.

Em comparação com métodos convexos, como SDR e programação fracionária, o PSO oferece maior flexibilidade para problemas não convexos, embora estes possam convergir mais rapidamente em cenários estruturados. Assim, o PSO apresenta um compromisso entre desempenho e complexidade.

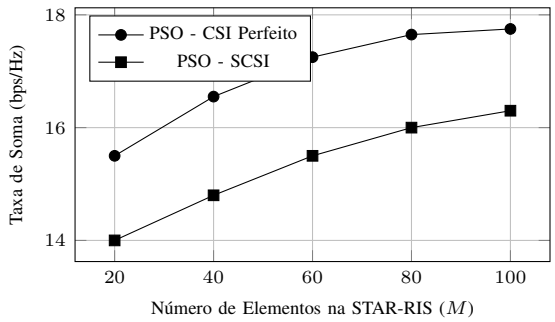


Fig. 3. Análise da influência do número de elementos na STAR-RIS (M).

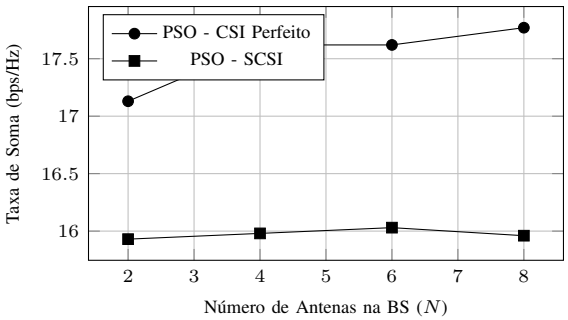


Fig. 4. Análise da influência do número de antenas na BS (N).

V. Conclusão

Propôs-se um método baseado na técnica PSO para maximizar a taxa de soma alcançável, realizando a otimização conjunta do *beamforming* na BS e da matriz de espalhamento da STAR-RIS. Os resultados demonstram que a abordagem proposta é capaz de atingir um desempenho satisfatório mesmo na ausência de estimação explícita da CSI, o que torna a solução particularmente atrativa do ponto de vista prático, ao reduzir a complexidade de implementação, os custos operacionais e o consumo de energia.Como continuidade deste trabalho, pretende-se investigar o uso de informações estatísticas de CSI, visando reduzir ainda mais a sobrecarga de treinamento e o overhead de sinalização, além de analisar o impacto dessa abordagem na robustez e na eficiência do sistema em cenários dinâmicos.

Referências

[1] Xueqin Jia, Ke Shi, Yuwen Zhang, Li.Wanghong, Haowen Sun, Changwei Hu e Xianghai Dou. “Use Case Report of Trusted Decentralized Networking Platforms Toward the 6G Network and Shared Network Resources”. Em: *Use Case Report of Trusted Decentralized Networking Platforms Toward the 6G Network and Shared Network Resources* (2025), pp. 1–17.

[2] Marius Corici, Fabian Eichhorn e Thomas Magedanz. “Organic 6G Continuum Architecture: A Uniform Control Plane Across Devices, Radio, and Core”. Em: *IEEE Networking Letters* 6.1 (2024), pp. 11–15. DOI: 10.1109/LNET.2023.3338363.

[3] Muhammad Miftahul Amri. “Recent Trends in the Reconfigurable Intelligent Surfaces (RIS): Active RIS to Brain-controlled RIS”. Em: *2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*. 2022, pp. 299–304. DOI: 10.1109/COMNETSAT56033.2022.9994338.

[4] Ana Luísa Fernandes de Souza. “Estudo sobre Redes Sem Fio Assistidas por Superfícies Inteligentes Reconfiguráveis com Reflexão e Transmissão Simultâneas”. Acesso em: [23/10/2025]. Trabalho de Conclusão de Curso (Graduação). São João da Boa Vista, SP, Brasil: Universidade Estadual Paulista (UNESP), Campus de São João da Boa Vista, 2023. URL: https://hdl.handle.net/11449/251877.

[5] Chenyu Wu, Yuanwei Liu, Xidong Mu, Xuemai Gu e Octavia A. Dobre. “Coverage Characterization of STAR-RIS Networks: NOMA and OMA”. Em: *IEEE Communications Letters* 25.9 (2021), pp. 3036–3040. DOI: 10.1109/LCOMM.2021.3091807.

[6] Xidong Mu, Yuanwei Liu, Li Guo, Jiaru Lin e Robert Schober. “Simultaneously Transmitting and Reflecting (STAR) RIS Aided Wireless Communications”. Em: *IEEE Transactions on Wireless Communications* 21.5 (2022), pp. 3083–3098. DOI: 10.1109/TWC.2021.3117462.

[7] Jiaqi Xu, Jiakuo Zuo, Joey Tianyi Zhou e Yuanwei Liu. “Active Simultaneously Transmitting and Reflecting (STAR)-RISs: Modeling and Analysis”. Em: *IEEE Communications Letters* 27.9 (2023), pp. 2466–2470. DOI: 10.1109/LCOMM.2023.3289066.

[8] Jiakuo Zuo, Yuanwei Liu, Zhiguo Ding, Lingyang Song e H. Vincent Poor. “Joint Design for Simultaneously Transmitting and Reflecting (STAR) RIS Assisted NOMA Systems”. Em: *IEEE Transactions on Wireless Communications* 22.1 (jan. de 2023), pp. 611–626. DOI: 10.1109/TWC.2022.3197079.

[9] Victoria Dala Pegorara Souto, Richard Demo Souza e Bartolomeu F. Uchôa-Filho. “PSO-based Optimization of STAR-RIS aided NOMA Wireless Communication Networks”. Em: *2024 International Symposium on Wireless Communication Systems (ISWCS)*. Accepted for publication. Rio de Janeiro, Brazil, 2024.

[10] Ha An Le, Trinh Van Chien e Wan Choi. “Graph Neural Network based Active and Passive Beamforming for Distributed STAR-RIS-Assisted Multi-User MISO Systems”. Em: *IEEE Transactions on Communications* (2025). Early Access. DOI: 10.1109/TCOMM.2025.3558504.

[11] Shuang Zhang, Wanming Hao, Gangcan Sun, Chongwen Huang, Zhengyu Zhu, Xingwang Li e Chau Yuen. “Joint Beamforming Optimization for Active STAR-RIS-Assisted ISAC Systems”. Em: *IEEE Transactions on Wireless Communications* 23.11 (nov. de 2024), pp. 15888–15902. DOI: 10.1109/TWC.2024.3434963.

[12] Tianqi Wang, Fang Fang e Zhiguo Ding. “Joint Phase Shift and Beamforming Design in a Multi-User MISO STAR-RIS Assisted Downlink NOMA Network”. Em: *IEEE Transactions on Vehicular Technology* 72.7 (jul. de 2023), pp. 9031–9043. DOI: 10.1109/TVT.2023.3248789.

[13] Chenyu Wu, Xidong Mu, Yuanwei Liu, Xuemai Gu e Xianbin Wang. “Resource Allocation in STAR-RIS-Aided Networks: OMA and NOMA”. Em: *IEEE Transactions on Wireless Communications* 21.9 (set. de 2022), pp. 7653–7667. DOI: 10.1109/TWC.2022.3160151.

[14] J. Kennedy e R. Eberhart. “Particle swarm optimization”. Em: *Proceedings of ICNN’95 - International Conference on Neural Networks*. Vol. 4. 1995, pp. 1942–1948. DOI: 10.1109/ICNN.1995.488968.

[15] Riccardo Poli, James Kennedy e Tim Blackwell. “Particle swarm optimization: An overview”. Em: *Swarm intelligence* 1.1 (2007), pp. 33–57.

[16] Y. Shi e R. Eberhart. “A modified particle swarm optimizer”. Em: *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*. 1998, pp. 69–73. DOI: 10.1109/ICEC.1998.699146.

AUTORES



Gislaine Abreu Mendonça é graduanda do 9º período de Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (INATEL). Possui experiência prática no setor de tecnologia, tendo atuado como estagiário na área de telecomunicações da Nokia.Atualmente, concentra seus estudos e pesquisas acadêmicas no desenvolvimento de soluções em Inteligência Artificial, Visão Computacional e Deep Learning aplicadas ao diagnóstico médico e análise de imagens clínicas.



Victoria Dala Pegorara Souto é graduada em Engenharia de Computação pela Universidade Federal de Santa Maria - UFSM (2014), Mestre em Engenharia Elétrica pela UFSM e Doutora em Engenharia Elétrica pela Universidade Federal de Santa Catarina - UFSC (2021). Em 2020, foi estudante visitante de doutorado na Universidade de Sydney em Camperdown, NSW, Austrália. Em 2021-2023, foi pesquisadora de pós-doutorado na UFSC atuando junto ao projeto Brasil 6G coordenado pelo Inatel e, atualmente, é Professora Assistente no Inatel. Seus interesses de pesquisa estão nas áreas de comunicações sem fio, redes 5G e 6G, transferência de energia sem fio e inteligência artificial.

Gaze Tracking: Controle de Cursor em Tempo Real Baseado em Rastreamento Ocular

Guilherme S. F. Almeida, Luiz E. A. Nora, Evandro C. Vilas Boas, Felipe A. P. Figueiredo
WAI Lab e CTIoT – Instituto Nacional de Telecomunicações - Inatel
Santa Rita do Sapucaí, Brasil
guilherme.almeida@ges.inatel.br, luiz.e@ges.inatel.br, evandro.cesar@inatel.br, felipe.figueiredo@inatel.br

***Abstract*—Computing systems have become essential in modern society, making the development of accessible human–computer interaction solutions increasingly important. However, traditional input devices such as the mouse and keyboard can represent significant barriers for individuals with reduced mobility. This work presents a real-time cursor control system based on eye gaze tracking using only a conventional RGB camera. The proposed architecture combines face detection and landmark extraction, an initial geometric gaze estimation stage, and a lightweight neural network model for spatial refinement of the predicted coordinates. Additionally, a multi-point calibration procedure with repeated cycles is employed to reduce systematic errors in the mapping between gaze direction and cursor position. Experimental results demonstrate that the integration of structured calibration, neural refinement, and temporal smoothing improves cursor stability and control accuracy, making the proposed solution suitable for low-cost assistive applications in real-time interaction environments.**

***Index Terms*—assistive technology, deep learning, eye tracking, real-time systems**

***Resumo*—A dependência de sistemas computacionais torna essencial o desenvolvimento de soluções acessíveis para interação humano–computador. Entretanto, dispositivos tradicionais de entrada, como mouse e teclado, podem representar barreiras para pessoas com mobilidade reduzida. Este trabalho apresenta um sistema de controle de cursor em tempo real baseado em rastreamento ocular utilizando apenas uma câmera RGB convencional. A arquitetura proposta combina detecção facial e extração de *landmarks*, estimação geométrica inicial da direção do olhar e um modelo de rede neural leve para refinamento espacial das coordenadas estimadas. Além disso, é aplicado um procedimento de calibração multiponto com repetição em ciclos consecutivos para reduzir erros sistemáticos no mapeamento entre gaze e posição do cursor. Resultados experimentais demonstram que a integração entre calibração estruturada, refinamento neural e suavização temporal melhora a estabilidade do cursor e a precisão do controle, tornando a solução viável para aplicações assistivas de baixo custo em ambientes de interação em tempo real.**

***Palavras chave*—deep learning, rastreamento ocular, sistemas em tempo real, tecnologia assistiva**

I. Introdução

A dependência de tecnologias digitais e de comunicação virtual tem ampliado a demanda por soluções que promovam acessibilidade e inclusão digital. Computadores e dispositivos conectados tornaram-se ferramentas essenciais para atividades como educação, trabalho e comunicação. Entretanto, para pessoas com deficiência motora ou mobilidade reduzida, o uso de dispositivos tradicionais de entrada, como mouse e teclado, pode representar uma barreira significativa, dificultando a interação eficiente com sistemas computacionais.

Nesse contexto, tecnologias assistivas possibilitam novas formas de interação humano-computador (IHC) que não dependem exclusivamente de dispositivos físicos convencionais. Entre essas alternativas, interfaces baseadas em rastreamento ocular (*eye tracking*) têm recebido atenção por permitirem inferir a intenção do usuário a partir do movimento dos olhos, viabilizando interações naturais e sem contato manual. Tais sistemas são particularmente relevantes em cenários de acessibilidade, podendo atuar como mecanismo primário de controle para usuários com limitações motoras severas.

Sistemas de rastreamento ocular de alta precisão geralmente utilizam *hardware* especializado, como sensores infravermelhos e câmeras dedicadas capazes de detectar reflexos corneanos e estimar o vetor de gaze com elevada acurácia. Embora apresentem excelente desempenho, seu alto custo limita a adoção em larga escala. Como alternativa, pesquisas recentes investigam métodos baseados em câmeras RGB convencionais e técnicas de visão computacional para estimar a direção do olhar a partir de imagens da face ou da região ocular [1].

Com o avanço do aprendizado profundo, especialmente de redes neurais convolucionais (CNN), tornou-se possível mapear imagens oculares diretamente para coordenadas de gaze. Zhang et al. [2] demonstraram a viabilidade de métodos *appearance-based* treinados com o dataset MPIIGaze para estimativa de gaze em ambientes não controlados. Posteriormente, abordagens baseadas na face completa foram propostas para aumentar a robustez frente a variações de pose e iluminação [3]. Além disso, o sistema RT-GENE mostrou que a combinação de normalização geométrica e redes neurais profundas pode permitir estimativas de gaze em tempo real em ambientes naturais [4].

Apesar desses avanços, modelos baseados em aprendizado profundo frequentemente exigem grande poder computacional e extensos conjuntos de dados, o que pode dificultar sua aplicação em sistemas interativos em tempo real. Nesse sentido, abordagens híbridas que combinam técnicas de visão computacional com modelos neurais mais leves têm se mostrado promissoras para equilibrar precisão e eficiência computacional.

Neste contexto, este trabalho apresenta um sistema de controle de cursor baseado em rastreamento ocular em tempo real utilizando apenas uma câmera RGB convencional. A proposta integra detecção facial, extração de *landmarks*, estimativa geométrica inicial da direção do olhar e um modelo neural leve para refinamento espacial das coordenadas estimadas. Adicionalmente, é adotado um procedimento de calibração

multiponto com repetição em ciclos consecutivos, visando reduzir erros sistemáticos no mapeamento entre gaze e posição do cursor, aumentando a estabilidade do controle.

O restante do artigo está organizado da seguinte forma. A Seção II apresenta os trabalhos relacionados e a arquitetura do sistema. A Seção III descreve a metodologia adotada, incluindo dataset, arquitetura do modelo e métricas de avaliação. Em seguida, são apresentados os resultados experimentais e sua análise. Por fim, a Seção IV apresenta as conclusões e direções para trabalhos futuros.

II. Trabalhos Relacionados

A. Trabalhos Relacionados

Sistemas de rastreamento ocular podem ser classificados, de maneira geral, em três categorias principais: soluções baseadas em hardware dedicado, métodos clássicos de visão computacional e abordagens baseadas em aprendizado profundo (*deep learning*). Cada uma dessas categorias apresenta diferentes níveis de precisão, custo de implementação e requisitos computacionais, influenciando diretamente sua aplicabilidade em cenários reais de interação humano–computador.

Sistemas comerciais de rastreamento ocular normalmente utilizam iluminação infravermelha ativa e sensores especializados para capturar reflexos corneanos e estimar o vetor de gaze com elevada precisão angular. Esses sistemas exploram propriedades ópticas do olho humano, como o reflexo de Purkinje e a posição relativa da pupila, permitindo estimativas precisas do ponto de fixação do olhar. Apesar da elevada acurácia alcançada, essas soluções dependem de *hardware* dedicado e possuem custo elevado, o que limita sua adoção em aplicações de larga escala ou em ambientes com restrições orçamentárias.

Como alternativa, diversos trabalhos têm investigado técnicas baseadas em câmeras convencionais Red-Green-Blue (RGB), explorando métodos de visão computacional para estimar a direção do olhar a partir de imagens da face ou da região ocular. Hansen e Ji [1] apresentaram um levantamento abrangente sobre modelos computacionais para representação do olho e estimação de gaze, discutindo diferentes abordagens geométricas e baseadas em aparência. O estudo destaca desafios importantes para sistemas baseados em imagem, como variações de iluminação, oclusões parciais, diferenças anatômicas entre indivíduos e movimentos naturais da cabeça durante a interação.

Com o avanço do aprendizado profundo, métodos baseados em redes neurais convolucionais (CNN) passaram a desempenhar papel central na estimação de gaze a partir de imagens. Essas abordagens permitem aprender representações visuais complexas diretamente dos dados, dispensando a necessidade de modelagem geométrica explícita. Zhang et al. [2] propuseram um método *appearance-based* que utiliza uma CNN treinada com o conjunto de dados MPIIGaze, demonstrando que modelos treinados com grandes bases de dados podem estimar a direção do olhar em ambientes não controlados (*in-the-wild*). Posteriormente, Zhang et al. [3] ampliaram essa abordagem incorporando informações da face completa no processo de estimação, explorando contexto visual adicional para melhorar a robustez frente a variações de pose e iluminação.

Outro avanço relevante foi apresentado por Fischer et al. [4], que introduziram o sistema RT-GENE (*Real-Time Gaze Estimation in Natural Environments*). Esse sistema combina normalização geométrica da região facial com modelos de aprendizado profundo para estimar o vetor de gaze em tempo real, mesmo em cenários com variações significativas de pose de cabeça e condições de iluminação. A proposta demonstrou que a integração entre pré-processamento geométrico e redes neurais profundas pode melhorar a robustez da estimação em ambientes naturais.

Apesar dos avanços proporcionados pelas abordagens baseadas em aprendizado profundo, muitos desses métodos dependem de arquiteturas neurais complexas e grandes conjuntos de dados para treinamento, o que pode aumentar significativamente o custo computacional e dificultar sua aplicação em sistemas interativos executados em tempo real. Dessa forma, abordagens híbridas que combinam técnicas clássicas de visão computacional com modelos de aprendizado profundo mais leves têm se mostrado promissoras, pois permitem equilibrar precisão, eficiência computacional e viabilidade prática.

Diferentemente dos sistemas baseados exclusivamente em *hardware* infravermelho, a solução proposta neste trabalho utiliza apenas uma câmera RGB convencional, reduzindo significativamente o custo de implementação e ampliando o potencial de aplicação em contextos de acessibilidade e tecnologias assistivas.

Em comparação com abordagens puramente *appearance-based* baseadas em redes profundas, como MPIIGaze e RT-GENE, o presente trabalho adota uma estratégia híbrida que combina técnicas de visão computacional para estimativa inicial do olhar com um modelo neural leve responsável pelo refinamento espacial das coordenadas estimadas. Essa abordagem busca alcançar um equilíbrio entre precisão e custo computacional, mantendo a capacidade de execução em tempo real em *hardware* amplamente disponível.

Além disso, propõe-se um procedimento de calibração multiponto com repetição em três ciclos consecutivos (*triple-loop calibration*), com o objetivo de reduzir erros sistemáticos no mapeamento entre a direção do olhar e as coordenadas da tela. Esse aspecto é frequentemente tratado de forma simplificada em implementações baseadas apenas em *webcam*, podendo impactar negativamente a estabilidade do controle de cursor. A integração desses módulos resulta em um sistema de baixo custo, projetado para aplicações de tecnologia assistiva, com foco em estabilidade de controle, robustez a variações ambientais e viabilidade prática em cenários reais de interação humano–computador.

III. Visão Geral do Sistema

A arquitetura do sistema, vista na Figura 1, foi projetada para permitir controle de cursor em tempo real a partir da estimação do olhar utilizando apenas uma câmera RGB convencional. A solução adota uma abordagem modular composta por etapas sequenciais de aquisição de imagem, detecção facial, extração de características oculares, refinamento por aprendizado profundo, calibração espacial e geração de comandos para o sistema operacional. O fluxo de processamento inicia-se com a captura contínua de quadros por meio da biblioteca OpenCV. Cada

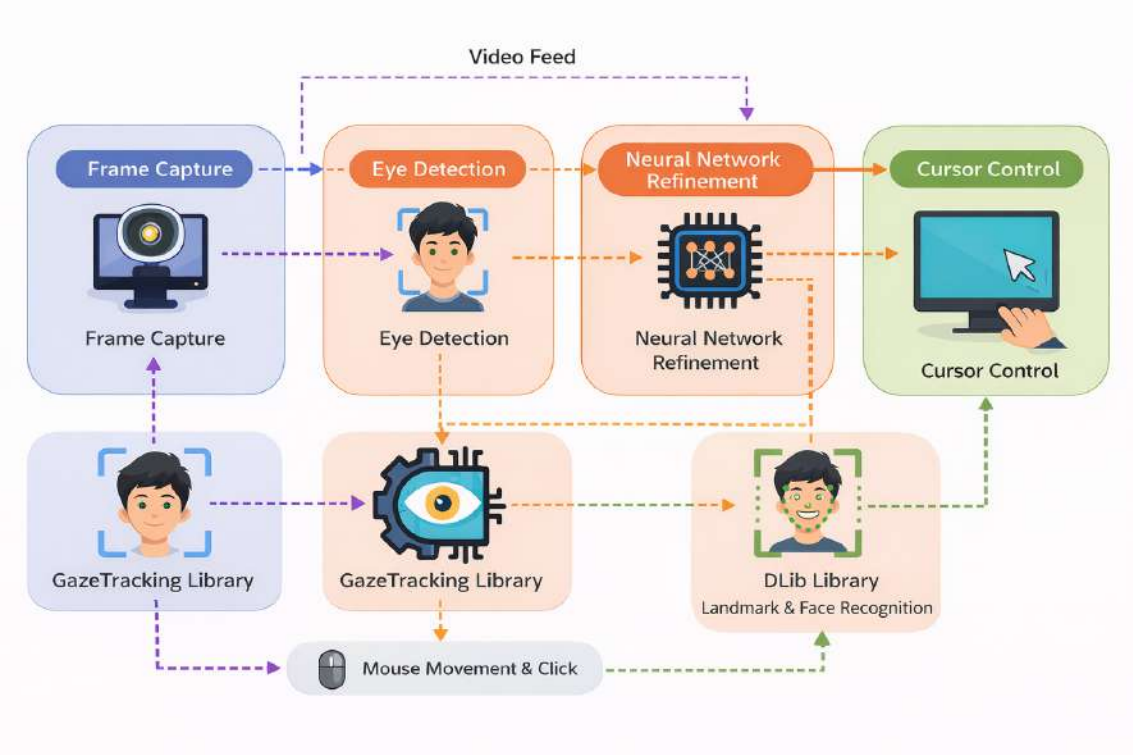


Fig. 1. Arquitetura geral do sistema de controle de cursor baseado em rastreamento ocular com integração da biblioteca `dlib` para detecção facial e landmarks.

frame é convertido e normalizado conforme necessário para reduzir ruídos e preparar os dados para as etapas subsequentes. A taxa de aquisição é mantida em regime de tempo real, garantindo responsividade adequada para interação humano-computador.

Na etapa seguinte, realiza-se a detecção facial utilizando a biblioteca *dlib*, que aplica um detector baseado em Histogram of Oriented Gradients (HOG) associado a um classificador linear. Uma vez detectada a face, o preditor de *landmarks* faciais de 68 pontos é utilizado para identificar regiões anatômicas relevantes, especialmente os contornos oculares e das pálpebras. Essa etapa fornece coordenadas estruturais que permitem isolar com precisão as regiões dos olhos, reduzindo interferências externas e variações geométricas decorrentes de pose ou inclinação da cabeça.

As regiões oculares extraídas são processadas pela biblioteca *GazeTracking*, responsável por estimar a direção inicial do olhar a partir da posição relativa das pupilas. Essa estimativa fornece informações horizontais e verticais preliminares, bem como indicadores de direção discreta (esquerda, direita ou central). Embora eficiente computacionalmente, essa abordagem geométrica pode apresentar variações decorrentes de iluminação e características individuais do usuário.

Para aumentar a robustez da estimativa espacial, as imagens oculares passam por um estágio de refinamento baseado em um modelo de rede neural totalmente conectada. As entradas são previamente redimensionadas e normalizadas para dimensão padronizada, garantindo consistência no processo de inferência. O modelo neural atua como um mecanismo de correção de erro sistemático, ajustando as coordenadas de gaze estimadas inicialmente e produzindo valores mais estáveis para mapeamento em tela.

Uma etapa fundamental da arquitetura é o procedimento de calibração multiponto. O sistema apresenta ao usuário cinco pontos fixos distribuídos estrategicamente na tela (cantos superior esquerdo e direito, centro, cantos inferior esquerdo e direito). Para cada ponto, são coletadas múltiplas amostras de gaze, e o processo é repetido em três ciclos consecutivos (*triple-loop calibration*). A repetição permite reduzir variabilidade estatística e melhorar a estimação dos parâmetros de transformação que relacionam coordenadas oculares ao espaço bidimensional da tela.

Com os parâmetros de calibração definidos, as coordenadas refinadas do olhar são transformadas em posições X-Y compatíveis com a resolução do monitor. Um mecanismo de suavização temporal é aplicado para reduzir oscilações e tremores naturais do movimento ocular, promovendo trajetórias de cursor mais contínuas. Por fim, o controle efetivo do cursor é realizado por meio da biblioteca *pyautogui*, que converte as coordenadas calculadas em eventos de movimentação do mouse no sistema operacional. Eventos de clique podem ser acionados a partir de padrões específicos detectados na região das pálpebras, como variações prolongadas na abertura ocular.

A integração desses módulos resulta em um *pipeline* híbrido que combina técnicas clássicas de visão computacional com refinamento por aprendizado profundo, mantendo baixo custo computacional e compatibilidade com *hardware* amplamente disponível. A arquitetura foi concebida para priorizar estabilidade, responsividade e aplicabilidade prática em contextos de tecnologia assistiva.

A. Dataset

O modelo foi treinado utilizando um conjunto proprietário composto por aproximadamente 2300 imagens da região ocular, coletadas em diferentes condições de iluminação e variações de pose. Cada imagem possui um rótulo associado representando coordenadas normalizadas (x, y) no intervalo $[0, 1]$, correspondentes à posição do olhar na tela. As imagens foram redimensionadas para dimensão fixa de 224×224 pixels e normalizadas antes de serem fornecidas ao modelo.

O processo de validação do *dataset* incluiu verificação automática de consistência dos rótulos, garantindo que todos os valores fossem convertidos para ponto flutuante e permanecessem dentro do intervalo válido. Rótulos inválidos ou fora do intervalo foram tratados por substituição controlada, evitando falhas durante o treinamento. O conjunto de dados foi dividido aleatoriamente em subconjuntos de treinamento e validação, utilizando proporção de 85% para treinamento e 15% para validação.

B. Arquitetura do Modelo

O modelo adotado, denominado *AdvancedEyeTrackingCNN*, consiste em uma rede neural profunda implementada em PyTorch. A arquitetura foi projetada para estimar coordenadas contínuas bidimensionais do olhar a partir da imagem da região ocular. A saída do modelo é composta por dois neurônios lineares correspondentes às coordenadas horizontais e verticais normalizadas. A função de perda utilizada foi o Erro Quadrático Médio (*Mean Squared Error*, MSE), apropriado para regressão contínua. Para mitigar sobreajuste, foi aplicada regularização por *dropout* com taxa de 0.3. O modelo possui milhares de parâmetros treináveis, contabilizados automaticamente durante a inicialização, permitindo análise da complexidade computacional.

C. Treinamento com Scheduler Avançado

O treinamento foi conduzido por até 300 épocas utilizando mini-batches de tamanho 32. A taxa de aprendizado inicial foi definida como 10^{-3}, com ajuste dinâmico por meio de estratégias de *learning rate scheduling*. Foram implementadas múltiplas estratégias de agendamento, incluindo *Cosine annealing* com *warm-up* (configuração recomendada), *Reduce on plateau*, *Step decay*, Decaimento exponencial e *Cyclic learning rate*. Nos experimentos principais foi adotado o esquema *cosine with warmup*, que combina fase inicial de aquecimento com decaimento cossenoidal, promovendo melhor convergência e estabilidade.

O treinamento utilizou o otimizador Adam e incluiu *early stopping* com paciência de 30 épocas, *Automatic Mixed Precision* (AMP) para aceleração em GPU e salvamento automático do melhor modelo com base na métrica de validação. Durante cada época, foram registradas as seguintes métricas *Loss* de treinamento, *Loss* de validação, acurácia@20% (definida como proporção de predições dentro de um limite de erro aceitável) e valor instantâneo da taxa de aprendizado.

D. Interface de Sistema e Integração em Tempo Real

Após o treinamento, o modelo é integrado ao pipeline de rastreamento ocular em tempo real. As estimativas produzidas pela rede neural refinam as coordenadas preliminares obtidas via técnicas geométricas. Um mecanismo de suavização temporal é aplicado às coordenadas preditas para reduzir oscilações naturais do movimento ocular. A posição resultante é então convertida para coordenadas absolutas da tela e enviada ao sistema operacional para controle do cursor.

E. Métricas de Avaliação

A avaliação do modelo foi conduzida sob dois aspectos. Considerou-se as métrica de treinamento Mean Squared Error (MSE) e Acurácia@20%, utilizada como métrica de tolerância espacial. Bem como métricas de rastreamento ocular: desvio médio em pixels entre posição estimada e posição alvo, erro angular médio, latência média por frame, e estabilidade temporal do cursor (variância do deslocamento). Essas métricas permitiram avaliar tanto a qualidade do aprendizado supervisionado quanto a viabilidade prática do sistema para interação humano-computador em tempo real.

IV. Avaliação e Resultados

A avaliação foi conduzida em dois níveis distintos: avaliação da rede neural em regime *offline*, analisando comportamento de treinamento e generalização interna, e avaliação do sistema integrado em tempo real, considerando desempenho computacional e aplicação prática. Na Figura 2, apresentam-se as curvas de treinamento: perda, erro euclidiano, taxa de aprendizado e *gap* treino-validação ao longo das épocas.

A. Avaliação da Rede Neural

Foram realizados múltiplos experimentos variando número de épocas (10 a 300), estratégia de *learning rate* (constante e com *scheduler*) e monitoramento do *gap* treino-validação. Nos treinamentos iniciais (10–30 épocas), observou-se redução consistente da perda de treino (de aproximadamente 0.15 para 0.09), enquanto a perda de validação apresentou oscilações significativas, atingindo valores superiores a 0.20 em alguns cenários. Esse comportamento indica instabilidade inicial e sensibilidade à taxa de aprendizado crescente. O erro euclidiano médio de treino reduziu de aproximadamente 0.55 para valores entre 0.40 e 0.42, enquanto o erro de validação apresentou picos superiores a 0.70 em alguns experimentos, caracterizando sobreajuste temporário.

Nos experimentos com maior número de épocas (60–100), utilizando estratégia de *scheduler* com crescimento inicial seguido de decaimento suave da taxa de aprendizado, observou-se comportamento mais estável com erro euclidiano de treino estabilizado entre 0.38 e 0.41, erro euclidiano de validação convergindo para faixa entre 0.40 e 0.43 e *gap* treino-validação médio entre -0.03 e -0.02. Esse comportamento indica melhor equilíbrio entre capacidade de ajuste e generalização. No experimento de maior convergência (300 épocas), com *learning rate* fixo de 10^{-4}, observou-se loss $\approx 2.5 \times 10^{-5}$, erro euclidiano médio $\approx 1.0 \times 10^{-3}$, *gap* treino-validação da ordem de 10^{-5} e Acc@10%, Acc@20% e Acc@30% = 1.000. A convergência simultânea das curvas de treino e validação indica ausência de *overfitting* significativo nesse cenário específico.

A análise de métrica de gap (*ValLoss* - *TrainLoss*) permitiu observar três regimes distintos: *underfitting* inicial com ambos

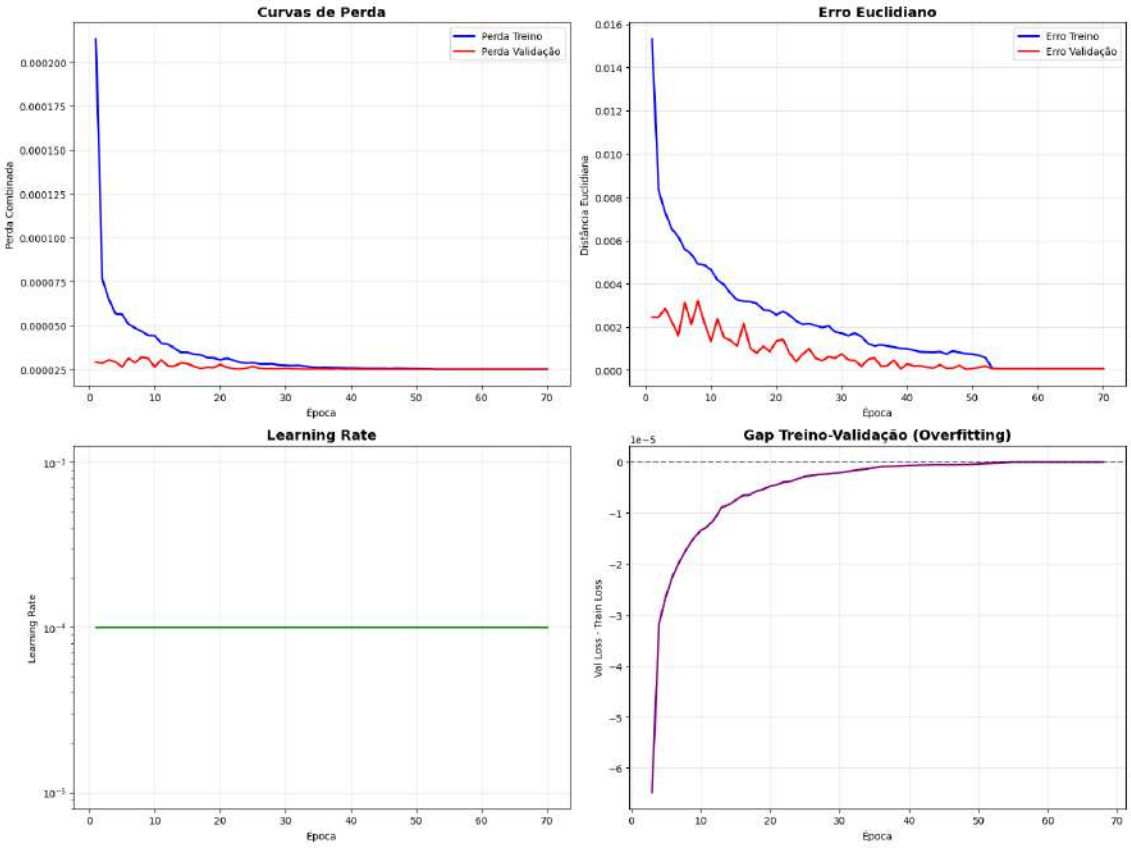


Fig. 2. Curvas consolidadas de treinamento: perda, erro euclidiano, taxa de aprendizado e gap treino-validação ao longo das épocas.

erros elevados e *gap* próximo de zero, *overfitting* intermediário com *gap* positivo acima de 0.10 em alguns experimentos de 50 épocas, e regime estável com *gap* negativo moderado (-0.02 a -0.03), indicando regularização implícita. A utilização do *scheduler* contribuiu para suavizar oscilações abruptas no erro de validação.

B. Avaliação do Sistema Integrado

No sistema integrado, o refinamento neural reduziu variações abruptas de coordenadas normalizadas, especialmente em regiões periféricas da tela. Os principais ganhos observados foram redução de tremores do cursor, melhoria na estabilidade em bordas da tela, trajetórias mais suaves e redução do erro espacial acumulado após calibração multiponto. Considerando resolução típica Full High Definition (Full HD, 1920×1080), erro normalizado de 0.001 implica deslocamento máximo aproximado de 1–2 pixels, o que é adequado para aplicações assistivas.

O treinamento da rede neural foi realizado no ambiente Google Colab Pro com aceleração por *Graphics Processing Unit* (GPU) NVIDIA Tesla T4 (16 GB). A avaliação do sistema integrado em tempo real foi conduzida em notebook equipado com processador Intel® Core™ i5-13420H (8 núcleos, 12 threads, até 4.60 GHz, 12 MB de cache), 8 GB de memória RAM DDR5 5200 MHz, armazenamento SSD NVMe PCIe 4.0 de 512 GB e GPU dedicada NVIDIA® GeForce RTX 4050 com 6 GB de memória GDDR6 (TGP até 80W). A etapa de inferência utilizou aceleração por GPU. A captura de vídeo foi realizada em resolução *Full High Definition* (Full HD, 1920×1080) configurada para 30 *Frames Per Second* (FPS). Considerando todo o pipeline do sistema — detecção facial, extração de *landmarks*, estimação geométrica, refinamento neural, suavização temporal e atualização do cursor — o desempenho observado variou entre 22 e 27 FPS, com média aproximada de 24.5 FPS. Na Tabela I, apresentam-se as métricas consolidadas do sistema completo em execução contínua, juntamente com faixas de referência consideradas ideais e aceitáveis para aplicações de rastreamento ocular baseadas em webcam Red-Green-Blue (RGB).

TABELA I
COMPARAÇÃO DO DESEMPENHO DO SISTEMA COM FAIXAS IDEAIS E ACEITÁVEIS PARA APLICAÇÕES DE RASTREAMENTO OCULAR EM TEMPO REAL

Métrica	Média Observada	Ideal	Aceitável
Latência por frame (ms)	40.8	$\leq$ 33 ms	$\leq$ 50 ms
FPS médio	24.5	$\geq$ 30 FPS	$\geq$ 20 FPS
Erro angular (graus)	3.2°	$\leq$ 2°	$\leq$ 4°
Desvio na tela (pixels)	5.4 px	$\leq$ 3 px	$\leq$ 8 px

Observa-se que o sistema operou dentro da faixa aceitável para interação humano-computador em tempo real ($\geq$ 20 FPS), embora abaixo do regime ideal de 30 FPS. A latência média de 40.8 ms permanece dentro do limite aceitável para aplicações interativas. O erro angular médio de 3.2° encontra-se dentro da faixa aceitável para sistemas baseados em câmera RGB convencional, ainda que superior ao patamar ideal observado em soluções com hardware dedicado. De forma correspondente, o desvio médio de 5.4 pixels em resolução Full HD mantém-se dentro da margem tolerável para controle fino de cursor em

aplicações assistivas.

A capacidade de generalização do modelo foi avaliada utilizando três indivíduos que não estavam presentes nas aproximadamente 2300 imagens utilizadas no treinamento. Para cada participante, foi executado o procedimento completo de calibração multiponto, coletando amostras correspondentes aos cinco pontos da tela. A análise visual dos clusters de calibração revelou que a estrutura espacial dos cinco pontos foi preservada em todos os indivíduos, indicando que o modelo mantém coerência geométrica básica mesmo fora da distribuição de treinamento. Entretanto, observou-se aumento consistente na dispersão intra-cluster e deslocamento sistemático dos centróides estimados.

No primeiro indivíduo, os clusters apresentaram separabilidade razoável no eixo horizontal, porém com variância vertical moderada, indicando erro predominantemente associado à componente Y da regressão. No segundo indivíduo, observou-se maior degradação, com deslocamento vertical significativo e aumento expressivo da variância intra-cluster. Esse comportamento sugere limitação do modelo em capturar diferenças anatômicas específicas e variações interindividuais não representadas no dataset original. No terceiro indivíduo, os clusters permaneceram estruturalmente organizados, porém com deslocamento sistemático moderado no eixo horizontal, caracterizando degradação intermediária.

Quantitativamente, o erro angular médio variou entre 8° e 11°, enquanto o desvio médio em tela situou-se entre 14 e 19 pixels em resolução Full HD. Esses valores representam aumento aproximado entre 2.5 e 3.5 vezes em relação ao desempenho obtido no conjunto de validação interna. Apesar da redução de precisão espacial, o sistema manteve desempenho computacional estável, operando entre 22 e 27 FPS, demonstrando que a limitação observada está associada à capacidade de generalização do modelo e não a restrições de processamento ou captura de vídeo.

V. Conclusão e Discussão

Este trabalho apresentou o desenvolvimento e a validação de um sistema de controle de cursor sem as mãos baseado em rastreamento ocular em tempo real utilizando apenas uma câmera RGB convencional. A arquitetura proposta combina detecção facial e extração de *landmarks* via *dlib*, estimação inicial da direção do olhar e um modelo neural totalmente conectado para refinamento espacial, compondo uma abordagem híbrida de baixo custo computacional.

A estratégia de calibração multiponto com repetição em ciclos consecutivos, aliada à normalização dos dados e à suavização temporal do cursor, mostrou-se eficaz na redução de erros sistemáticos e oscilações abruptas. Os resultados experimentais indicaram convergência estável do treinamento, baixo erro euclidiano normalizado e redução do *gap* treino-validação, evidenciando boa capacidade de generalização do modelo.

Observou-se que treinamentos curtos apresentam maior instabilidade de generalização, enquanto o uso de *learning rate schedulers* contribui significativamente para a convergência do modelo. Em particular, treinamentos prolongados com taxa de aprendizado controlada permitiram melhor desempenho e maior precisão no mapeamento entre a região ocular e as coordenadas Apesar de oscilações ocasionais relacionadas à configuração do *learning rate* e à variabilidade limitada do conjunto de dados, a integração entre calibração multiponto, regressão neural e suavização temporal resultou em um sistema estável para controle de cursor em tempo real.

Como principal contribuição, o trabalho demonstra que a combinação de técnicas clássicas de visão computacional com aprendizado profundo leve pode viabilizar sistemas de rastreamento ocular acessíveis, executados em tempo real e sem necessidade de hardware especializado. Como trabalhos futuros, pretende-se ampliar o conjunto de dados, investigar arquiteturas neurais mais robustas e incorporar novos mecanismos de interação, como detecção de piscadas e gestos oculares, além de realizar estudos de usabilidade com usuários.

Agradecimentos

Os autores agradecem ao Instituto Nacional de Telecomunicações (INATEL) pelo suporte estrutural e acadêmico para o desenvolvimento desta pesquisa. Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Referências

[1] D. W. Hansen e Q. Ji. “In the eye of the beholder: A survey of models for eyes and gaze”. Em: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.3 (2010), pp. 478–500. URL: https://ieeexplore.ieee.org/document/4770110.

[2] X. Zhang, Y. Sugano, M. Fritz e A. Bulling. “Appearance-based gaze estimation in the wild”. Em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015, pp. 4511–4520. URL: https://openaccess.thecvf.com/content_cvpr_2015/papers/Zhang_Appearance-Based_Gaze_Estimation_2015_CVPR_paper.pdf.

[3] X. Zhang, Y. Sugano, M. Fritz e A. Bulling. “It’s written all over your face: Full-face appearance-based gaze estimation”. Em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2017. URL: https://openaccess.thecvf.com/content_cvpr_2017_workshops/w41/papers/Bulling_Its_Written_All_CVPR_2017_paper.pdf.

[4] T. Fischer, H. J. Chang e Y. Demiris. “RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments”. Em: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 334–351. URL: https://link.springer.com/chapter/10.1007/978-3-030-01249-6_21.

[5] D. E. King. “Dlib-ml: A machine learning toolkit”. Em: *Journal of Machine Learning Research* 10 (2009), pp. 1755–1758. URL: http://jmlr.org/papers/v10/king09a.html.

[6] V. Kazemi e J. Sullivan. “One millisecond face alignment with an ensemble of regression trees”. Em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2014, pp. 1867–1874. URL: https://openaccess.thecvf.com/content_cvpr_2014/papers/Kazemi_One_Millisecond_Face_2014_CVPR_paper.pdf.

[7] Antoine Lamé. *GazeTracking: A Python library for eye tracking with OpenCV*. GitHub repository. 2020. URL: https://github.com/antoinelame/GazeTracking.

[8] G. Bradski. “The OpenCV Library”. Em: *Dr. Dobb’s Journal of Software Tools* (2000). URL: https://opencv.org/.

[9] Al Sweigart. *PyAutoGUI: Cross-platform GUI automation for human beings*. GitHub repository. 2015. URL: https://github.com/asweigart/pyautogui.

[10] Adam Paszke et al. “PyTorch: An imperative style, high-performance deep learning library”. Em: *Advances in Neural Information Processing Systems (NeurIPS)*. 2019. URL: https://arxiv.org/abs/1912.01703.

[11] D. P. Kingma e Jimmy Ba. “Adam: A method for stochastic optimization”. Em: *Proceedings of the International Conference on Learning Representations (ICLR)*. 2015. URL: https://arxiv.org/abs/1412.6980.

Instalação do Sistema Operacional Debian Minimal em Dispositivo com Processador S905X4

Matheus Dapper Alves Leite, Kleyton Hoffmann
Universidade do Oeste de Santa Catarina - UNOESC
matheus.dapper.2023@gmail.com, kleyton.hoffmann@unoesc.edu.br

***Abstract*—Este artigo apresenta a transposição funcional do dispositivo de *streaming* BTV modelo B13 em uma estação de microcomputação de propósito geral. A metodologia descreve a implementação de um *kernel* Linux customizado (Debian 6.12.18) via projeto *DevMFC*, com ênfase na configuração do *Device Tree Blob* (DTB) para o SoC Amlogic S905X4. Os resultados validam a migração do sistema para o armazenamento interno (eMMC), garantindo estabilidade operacional e baixo consumo de recursos. Conclui-se que o reaproveitamento deste hardware subutilizado é uma alternativa viável e sustentável, mitigando a obsolescência tecnológica e provendo uma plataforma robusta para serviços de rede.**

Palavras-chave: Sistemas Embarcados, Linux, Amlogic S905X4, Reuso de Hardware, Debian.

***Abstract*—[english] This paper presents the functional transposition of the BTV model B13 streaming device into a general-purpose microcomputing station. The methodology describes the implementation of a customized Linux kernel (Debian 6.12.18) via the DevMFC project, emphasizing the Device Tree Blob (DTB) configuration for the Amlogic S905X4 SoC. The results validate the system migration to internal storage (eMMC), ensuring operational stability and low resource consumption. It is concluded that repurposing this underutilized hardware is a viable and sustainable alternative, mitigating technological obsolescence and providing a robust platform for network services.**

Keywords: Embedded Systems, Linux, Amlogic S905X4, Hardware Reuse, Debian.

I. Introdução

A problemática dos aparelhos televisivos irregulares, popularmente conhecidos como *TV Box*, transcende a simples pirataria de sinal. O comércio desses dispositivos movimenta um mercado cinza que impacta diretamente a economia criativa e a arrecadação tributária, afetando setores como cinema, música e produção audiovisual [3].

No Brasil, para que um equipamento de radiocomunicação seja comercializado, ele deve obrigatoriamente possuir homologação da Agência Nacional de Telecomunicações (Anatel), responsável por realizar testes e certificações que garantem a segurança elétrica dos dispositivos e a conformidade quanto à interferência de radiofrequência. A ausência desse selo indica que o aparelho não passou pelos procedimentos oficiais de verificação exigidos pela regulamentação brasileira.

Além da infração administrativa, a ilegalidade penal concentra-se na violação de direitos autorais. Ao comercializar ou utilizar dispositivos configurados para descriptografar sinais de TV por assinatura e serviços de streaming sem o devido pagamento de licenças, incorre-se em crimes previstos na Lei de Direitos Autorais e no Código Penal brasileiro [2].

A Receita Federal, atuando na linha de frente em portos e centros de distribuição, apreende anualmente milhões dessas unidades. O destino tradicional para esses produtos era a destruição mecânica e o descarte em aterros sanitários, o que gerava um problema ambiental crítico devido à presença de metais pesados e plásticos de difícil decomposição nos componentes eletrônicos.

A mudança de paradigma ocorreu com a criação de parcerias entre a Receita Federal e instituições de ensino superior, com destaque para a Universidade do Oeste de Santa Catarina (UNOESC). O Projeto ECOA surge como uma resposta sustentável e educativa para o destino das apreensões, promovendo o reaproveitamento de dispositivos eletrônicos apreendidos. Em vez de serem descartadas como resíduos, as TV Boxes passam a ser tratadas como recursos computacionais com potencial de reutilização em atividades educacionais [12].

O processo de transformação é rigoroso e técnico. Inicialmente, equipes formadas por estudantes e professores de cursos como Ciência da Computação e Engenharia realizam a descaracterização do dispositivo, o que envolve a remoção completa do software original (firmware). Uma vez que o hardware é ”limpo”, sua arquitetura computacional pode ser reaproveitada para a instalação de sistemas operacionais de código aberto [12].

Esses dispositivos, agora transformados em microcomputadores funcionais, ganham uma nova vida. Eles passam a ser equipados com suítes de escritório, navegadores web e softwares de programação básica. Após a reconfiguração, as unidades são destinadas pela Receita Federal a escolas da rede pública de ensino, bibliotecas e centros comunitários, promovendo sua utilização em atividades educacionais e de inclusão digital [11].

Em vista disso, estabelece-se como problema de pesquisa a investigação da possibilidade de inicialização de um sistema *Debian* em uma *TV Box*.

A problemática central deste estudo reside na limitação operacional de dispositivos TV Box que, apesar de possuírem hardware com capacidade computacional relevante, são frequentemente restritos a sistemas proprietários voltados exclusivamente ao entretenimento. Identifica-se, portanto, uma lacuna quanto à versatilidade desses aparelhos, uma vez que a utilização de sistemas operacionais de propósito geral poderia ampliar significativamente suas funcionalidades.

Diante desse cenário de subutilização de hardware, questiona-se a possibilidade de superar essas restrições de software para a implementação de um ambiente mais flexível e aberto. Assim, o problema de pesquisa consolida-se na seguinte indagação: é possível realizar a inicialização e execução do sistema Debian em uma TV Box? A investigação busca deter-

minar se a arquitetura desses dispositivos suporta a transição para um sistema de código aberto, permitindo um aproveitamento mais amplo do equipamento em contextos educacionais e computacionais [5], [6].

Esta pesquisa analisa o potencial de reaproveitamento de dispositivos de streaming de baixo custo, cujas capacidades computacionais frequentemente permanecem subutilizadas. O estudo investiga a viabilidade de converter uma TV Box, originalmente destinada ao entretenimento, em uma estação de computação de propósito geral. Para isso, propõe-se a implementação e validação de um kernel Linux customizado baseado na distribuição Debian em arquitetura ARM.

II. Referencial Teórico

Para a compreensão do tema abordado nesta pesquisa, apresenta-se a fundamentação teórica acerca da arquitetura computacional, sistemas embarcados e o ecossistema de software utilizado no processo de transposição funcional.

.1) Arquitetura ARM e Sistemas Embarcados

A arquitetura ARM (*Advanced RISC Machines*) consolidou-se como o padrão predominante em sistemas embarcados devido à sua elevada eficiência energética e desempenho por *watt*. Esses dispositivos operam sob diferentes especificações de largura de registradores, destacando-se as arquiteturas de 32 bits (*ARM32*) e 64 bits (*ARM64*). No contexto de distribuições Linux, o perfil *armhf* (*ARM Hard Float*) assume papel relevante ao otimizar operações de ponto flutuante via hardware, sendo amplamente adotado em plataformas de baixo custo[19].

.2) Convergência Digital e Dispositivos de Streaming

Os *Set-Top Boxes* (STB), popularmente conhecidos como *TV Boxes*, evoluíram de simples conversores de sinal para dispositivos capazes de prover acesso à internet e serviços de *streaming*. No cenário nacional, a Agência Nacional de Telecomunicações (Anatel) atua na regulação e certificação desses equipamentos, garantindo a conformidade com normas de segurança e uso do espectro radioelétrico. O objeto de estudo desta pesquisa, o modelo BTV B13, enquadra-se nesta categoria de hardware embarcado baseado em SoCs (*System on a Chip*) de alto potencial subutilizado[20].

.3) Ecossistema de Software e Distribuição Debian

A conversão do hardware original em uma estação de propósito geral é viabilizada pela utilização do Debian Minimal. Por ser uma versão reduzida baseada no *kernel* Linux e focada em interface de terminal, esta distribuição permite um gerenciamento otimizado de recursos. A compatibilidade técnica é assegurada pelo repositório *DevMFC Debian-on-Amlogic*, que fornece *kernels* customizados para processadores Amlogic, permitindo que o sistema operacional interaja corretamente com as nuances do hardware proprietário[14], [21].

.4) Sistemas de Arquivos e Implementação

A organização dos dados no dispositivo de armazenamento segue padrões específicos de sistemas Unix-like. A partição *root*, essencial para a administração do sistema, utiliza preferencialmente o sistema de arquivos Ext4 pela sua robustez e suporte a permissões avançadas. Para o processo de inicialização e compatibilidade, utiliza-se o FAT32. A transposição das imagens de sistema (.img) para as mídias físicas é realizada através do *Win32DiskImager*, garantindo a integridade dos setores de inicialização (*bootloader*) necessários para a execução do novo ambiente operacional [25], [23], [22], [24].

A. Preservação de Dados: Backup e Snapshots

A integridade dos sistemas operacionais embarcados, especialmente durante fases experimentais de configuração, depende de mecanismos robustos de recuperação de desastres. Nesse contexto, distinguem-se duas estratégias fundamentais: o *backup* e o *snapshot*.

O *backup* consiste na cópia integral e independente de arquivos ou volumes de dados para uma unidade de armazenamento distinta. Sua finalidade primordial é a redundância física, permitindo a restauração do sistema em caso de falha catastrófica do hardware original ou corrupção irreversível do sistema de arquivos [10].

Em contrapartida, o *snapshot* (ou captura de estado) funciona como um registro instantâneo do estado lógico de um sistema de arquivos em um determinado momento. Diferente do *backup* convencional, o *snapshot* é frequentemente incremental e permite que o desenvolvedor reverta alterações experimentais — como falhas na edição do *bootloader* ou instabilidades no *kernel* — para um estado funcional anterior de forma ágil, sem a necessidade de uma reinstalação completa da imagem binária [18].

III. Metodologia

Esta pesquisa adota uma abordagem experimental de natureza aplicada, estruturada sob o método hipotético-dedutivo. A investigação é guiada pela viabilidade técnica de transposição de um sistema proprietário para uma arquitetura aberta. O procedimento metodológico foi subdividido em quatro fases distintas:

1) Levantamento e Caracterização: Coleta de dados técnicos acerca do SoC (*System on a Chip*) e da arquitetura de memória do dispositivo BTV B13.
2) Investigação e Prototipagem: Identificação de *kernels* compatíveis e definição dos parâmetros de compilação para a distribuição Debian.
3) Execução Experimental: Implementação do ciclo iterativo de instalação, tratamento de erros e refinamento de *drivers* de baixo nível.
4) Análise de Resultados: Avaliação da estabilidade do sistema e validação da capacidade de processamento para funções de microcomputação de propósito geral.

Um outro problema que reside na pesquisa é a escassez de literatura técnica específica para o modelo BTV B13, o que demanda uma investigação empírica sobre a compatibilidade de seu *bootloader* com sistemas Linux não-proprietários.

IV. Estado da Arte

A literatura recente apresenta uma convergência de esforços voltados ao reaproveitamento de hardware de baixo custo para mitigar a obsolescência tecnológica e promover a inclusão digital.

Estudos publicados no ENEDS demonstram que a recaracterização de *TV Boxes* em minicomputadores é uma estratégia eficaz de reuso tecnológico, validando a capacidade desses dispositivos em suportar tarefas computacionais básicas [9]. Complementarmente, pesquisas no IEEE detalham a superação de limitações de sistemas proprietários através da modificação de bibliotecas e *drivers*, resultando em plataformas abertas e estáveis [4].

A viabilidade técnica da instalação do sistema operacional Debian em dispositivos baseados em SoCs Amlogic é sustentada pelo projeto Debian-on-Amlogic, desenvolvido por *DevMFC*. Este repositório provê *kernels* customizados e scripts de automação que permitem a transposição de sistemas proprietários para ambientes de código aberto, mitigando as barreiras de compatibilidade de hardware em dispositivos ARM [14].

No âmbito nacional, o trabalho apresentado no WICS/SBC destaca o papel do software livre no prolongamento do ciclo de vida desses equipamentos, transformando-os em ferramentas para educação digital [8]. Por fim, pesquisas publicadas na Nature Scientific Reports elevam o escopo de aplicação desses dispositivos, validando seu desempenho como pequenos servidores em redes de computação em névoa (*fog computing*) e infraestruturas de Cidades Inteligentes (IoT), evidenciando uma eficiência energética superior em contextos de processamento distribuído [7].

V. Procedimento de instalação

A. Instalação e Configuração do Sistema Operacional

A etapa inicial consistiu na obtenção da imagem de sistema Debian Minimal, disponibilizada pelo projeto *DevMFC*. Esta distribuição foi selecionada por ser otimizada para arquiteturas ARM e por oferecer um ambiente de software enxuto, reduzindo o consumo de recursos computacionais em hardware limitado. O acesso à imagem ocorreu via repositório oficial no GitHub, o que assegura a reprodutibilidade do experimento por outros pesquisadores.

A opção pela versão *minimal* permitiu maior flexibilidade na personalização de pacotes e serviços, garantindo que apenas as dependências essenciais fossem carregadas. Este controle sobre o sistema operacional é fundamental para transformar o dispositivo de *streaming* em uma estação de computação estável e eficiente.

B. Processo de Gravação e Inicialização

Para a transferência da imagem (.img) a um USB Stick também conhecido como Pendrive no Brasil, utilizou-se o *software Win32DiskImager*. Após a gravação, realizou-se a configuração do *bootloader* para permitir que o hardware reconhecesse a nova partição de inicialização.

Os testes iniciais confirmaram a viabilidade da transição: o dispositivo BTV B13 completou o ciclo de inicialização com sucesso, carregando a interface de terminal do Debian. O consumo de memória RAM em modo ocioso (*idle*) apresentou valores significativamente baixos, validando a premissa de eficiência do sistema customizado.

C. Obtenção e Critérios de Seleção da Imagem

A imagem do *Debian Minimal* com *kernel* Linux 6.12.18 foi adquirida via repositório oficial do projeto *DevMFC* no GitHub [14]. Optou-se desta versão pois é a mais estável disponível para BTV B13, visando assegurar a integridade do sistema por meio de atualizações de segurança e otimizações de desempenho. Esta escolha garante a compatibilidade necessária com os *drivers* de hardware específicos do SoC Amlogic, fundamentais para a estabilidade da plataforma em ambiente de produção.

D. Preparação da Mídia de Inicialização

Após a obtenção da imagem de sistema, procedeu-se à criação de uma mídia de inicialização (*boot*) capaz de executar o instalador no hardware embarcado. Para esta finalidade, utilizou-se o utilitário Win32DiskImager, selecionado por sua eficácia na gravação de imagens binárias (.img) diretamente em dispositivos de armazenamento em massa via USB [14], [16].

d.1) Procedimento Experimental

O processo de gravação seguiu o protocolo estruturado abaixo:

1) Seleção do Hardware: Utilizou-se um dispositivo de armazenamento USB com capacidade nominal mínima de 8 GB, assegurando espaço suficiente para a imagem e partições temporárias [15].
2) Configuração do Software: O *software* Win32DiskImager foi executado para mapear o arquivo binário (.img) proveniente do repositório DevMFC.
3) Mapeamento de Unidade: Identificou-se a unidade lógica correspondente ao dispositivo USB no sistema operacional hospedeiro, prevenindo gravações acidentais em volumes de sistema.
4) Execução da Gravação: Iniciou-se a operação de escrita (*Write*), processo que realiza a formatação e a transposição bit a bit da imagem para o dispositivo alvo [24][17].
5) Finalização: Após a validação da escrita, o dispositivo foi removido de forma segura para garantir a integridade dos dados e do setor de inicialização (*Master Boot Record*).

É imperativo salientar que este procedimento resulta na exclusão definitiva de dados previamente contidos na unidade USB. Ao término desta fase, a mídia de inicialização encontra-se apta para integrar o ciclo de *boot* da TV Box, permitindo o prosseguimento da instalação do ecossistema Debian.

E. Verificação e Estrutura das Partições

Após a transposição da imagem binária, é indispensável a validação da tabela de partições para assegurar a integridade do ciclo de *boot*. A arquitetura de software proposta pelo projeto *DevMFC* segmenta o dispositivo de armazenamento em duas partições primárias com funções distintas [15]:

- Partição de Inicialização (*Boot*): Formatada em sistema de arquivos FAT32, contém o *bootloader*, o *kernel* e os arquivos de configuração necessários para o reconhecimento do hardware. Sua compatibilidade universal permite o acesso administrativo em ambientes Windows, macOS e Linux.

- Partição Raiz (*Root*): Utiliza o sistema de arquivos EXT4, padrão em distribuições GNU/Linux. Esta partição armazena a hierarquia do sistema operacional, incluindo diretórios críticos como `/bin`, `/etc` e `/var`.

e.1) Interoperabilidade e Limitações de Acesso

Devido à ausência de suporte nativo do sistema operacional Microsoft Windows para o sistema de arquivos EXT4, apenas a partição *Boot* torna-se visível no explorador de arquivos convencional. A partição *Root*, embora íntegra, permanece inacessível em ambientes não-POSIX [27].

Para a manipulação direta de arquivos na partição *Root*, procedeu-se à utilização de um ambiente *Linux Native*. O protocolo de acesso experimental seguiu as etapas:

1) Montagem do dispositivo em uma estação de trabalho com *kernel* Linux estável.
2) Identificação volumétrica das partições via terminal ou gerenciador de montagem automática (`/media` ou `/run/media`).
3) Execução de modificações granulares sob privilégios de superusuário (*root*), garantindo a persistência das configurações de sistema.

Nota Técnica: A integridade da partição *Root* é sensível a modificações arbitrárias. Recomenda-se a realização de *snapshots* ou *backups* antes de alterações estruturais que possam comprometer a estabilidade do sistema Debian.

F. Configuração do Parâmetro de Inicialização (Boot Config)

A etapa crítica para a compatibilidade entre o *kernel* e o hardware embarcado reside na configuração do arquivo `uEnv.txt` (ou `boot.config`), localizado na partição *BOOT* (FAT32). Este arquivo gerencia os parâmetros de inicialização e define o Device Tree Blob (DTB) a ser carregado pelo sistema [14].

f.1) Seleção do Device Tree (DTB)

O DTB é responsável por descrever a topologia do hardware ao sistema operacional, detalhando endereçamentos de CPU, memória e periféricos. No contexto do dispositivo BTV B13, o procedimento de ajuste seguiu as diretrizes:

1) Acesso e Edição: O arquivo foi editado em ambiente de texto simples, identificando-se as diretivas de seleção de *chipset*.
2) Habilitação do Hardware: Procedeu-se à remoção do caractere de comentário (#) na linha correspondente ao SoC *Amlogic S905X4*. A instrução foi alterada para: `box=s905x4_generic`.
3) Persistência: A gravação do arquivo assegura que, no ciclo de *boot*, o *kernel* carregue os *drivers* específicos para a arquitetura S905X4, garantindo a estabilidade operacional dos componentes integrados.

Esta configuração manual é imperativa, uma vez que a imagem genérica do *DevMFC* mantém as definições de hardware comentadas por padrão para evitar conflitos em dispositivos não compatíveis.

G. Protocolo de Inicialização e Carregamento do Sistema

A fase de inicialização física consiste na integração da mídia USB ao hardware do BTV B13. O sucesso desta etapa valida a configuração prévia do *bootloader* e a compatibilidade do DTB selecionado.

g.1) Procedimento de Intervenção de Boot

Embora alguns dispositivos apresentem prioridade de inicialização para mídias externas, em arquiteturas fechadas como a do modelo em estudo, frequentemente faz-se necessária a execução do modo de recuperação (*Recovery Mode*) para redirecionar o fluxo de leitura para a porta USB. O protocolo experimental seguiu as diretrizes:

1) Conectividade: Inserção da mídia de armazenamento na interface USB primária do dispositivo desenergizado.
2) Acionamento Mecânico: Pressão contínua no interruptor físico de *Update* (localizado na parte de baixo da Tv Box) simultaneamente à conexão da fonte de alimentação.
3) Temporização: Manutenção do estado de pressão por um intervalo de 5 a 10 segundos, forçando o *bootloader* a buscar assinaturas de inicialização na interface externa.

g.2) Validação e Diagnóstico

A confirmação da transição operacional é evidenciada pelo carregamento sequencial do *kernel* Linux, culminando no *prompt* de comando do Debian Minimal.

Em casos de falha na inicialização (*boot loop* ou travamento no logotipo do fabricante), a metodologia de diagnóstico prevê a revalidação da integridade da imagem binária, a conferência da sintaxe no arquivo `boot.config` ou a substituição da mídia física por modelos de maior taxa de transferência, visando mitigar latências críticas no barramento USB.

H. Migração para o Armazenamento Interno (eMMC)

A execução do Debian via interface USB é recomendada para fases de teste e validação de *drivers*. Contudo, para otimizar a taxa de transferência de dados e assegurar a estabilidade operacional, procedeu-se à instalação do sistema no armazenamento interno (*eMMC*) do dispositivo.

h.1) Protocolo de Instalação e Configuração de Multiboot

O processo de migração utiliza *scripts* de automação do projeto *DevMFC*, que realizam o particionamento de baixo nível e a transposição da imagem lógica para o suporte físico interno. O procedimento experimental seguiu o fluxo iterativo abaixo:

1) Execução do Script de Implantação: Acionou-se o utilitário de instalação (`./aml-install-to-emmc.sh`), que identifica a topologia do *eMMC* e solicita a configuração do *multiboot*.
2) Configuração do Gerenciador de Inicialização: Devido às restrições do *firmware* original, executou-se o *script* `./aml-multiboot-setup.sh`. Esta etapa é fundamental para sobrescrever as diretivas de *boot* de fábrica, permitindo que o hardware priorize o *kernel* Linux customizado em detrimento do sistema proprietário.
3) Finalização da Transposição: Com o *multiboot* devidamente instanciado, reexecutou-se o *script* de instalação

principal para concluir a cópia dos volumes *Root* e *Boot* para o *eMMC*.

h.2) Procedimento de Desligamento e Validação Final

Finalizada a etapa de escrita, o comando `poweroff` foi enviado ao sistema para encerrar as operações. É importante ressaltar que, em virtude de limitações no DTB (*Device Tree Blob*) — responsável pelo mapeamento dos componentes físicos do dispositivo de *streaming* —, o hardware de baixo custo pode não atingir o desligamento elétrico completo. Tal estado é evidenciado exclusivamente pela interrupção do sinal de vídeo, indicando a conclusão segura do processo de *kernel*.

1) Desenergização Segura: Após a interrupção do sinal de vídeo, procedeu-se ao corte da alimentação elétrica e à remoção da mídia USB.
2) Inicialização Autônoma via eMMC: Ao restabelecer a alimentação elétrica, o dispositivo BTV B13 executou o ciclo de *boot* diretamente a partir da imagem persistida na memória interna (*eMMC*). Este procedimento confirmou a eficácia da migração de dados e a autonomia do sistema operacional sem a dependência de mídias externas.

A transição para o *eMMC* resultou em uma redução perceptível na latência de carregamento de serviços, validando a eficácia do hardware como uma unidade de processamento independente e estável.

VI. Discussão

Os resultados validam a transposição funcional do *chipset* Amlogic S905X4, cuja estabilidade operacional foi confirmada após a correta vinculação do DTB ao *kernel* Linux. A migração bem-sucedida para o *eMMC* eliminou a latência do barramento USB, consolidando o hardware como uma estação de microcomputação autônoma. O uso da distribuição *Debian Minimal* otimizou o consumo de recursos escassos, provando que dispositivos de *streaming* subutilizados são plataformas viáveis para serviços de rede e desenvolvimento, superando as restrições dos sistemas proprietários originais.

VII. Considerações Finais

Este estudo demonstrou a viabilidade técnica da transposição funcional do dispositivo BTV B13 para um ambiente de computação de propósito geral. A implementação do *Debian Minimal* sob o *kernel* 6.12.18 validou a eficácia do projeto *DevMFC* na abstração de hardware proprietário, provando que a configuração precisa do *Device Tree Blob* (DTB) é o fator determinante para a estabilidade em SoCs Amlogic S905X4.

A transição do pendrive para o armazenamento interno (*eMMC*) consolidou o dispositivo como uma unidade autônoma de baixo consumo e alta disponibilidade. Conclui-se que o reaproveitamento de *streaming devices* via *software* livre não apenas estende o ciclo de vida útil do hardware, mitigando o descarte eletrônico, mas também oferece uma plataforma robusta para servidores domésticos e microcomputação educativa. Trabalhos futuros podem explorar o desempenho térmico sob estresse e a escalabilidade de redes em névoa (*fog computing*) utilizando o ambiente aqui estabelecido.

Agradecimentos

Agradecemos o apoio da Unoesc por meio do projeto ECOA e da Receita Federal de Joaçaba com a parceria desenvolvida no período de vigência do projeto.

Referências

[1] ANATEL. **Plano de Combate ao Uso de Decodificadores Clandestinos de SeAC**. Agência Nacional de Telecomunicações, Brasília, 2023. Disponível em: https://www.gov.br/anatel/pt-br/assuntos/noticias/anatel-aprova-plano-de-combate-ao-uso-de-decodificadores-clandestinos-de-tv-por-assinatura. Acesso em: 28 abr. 2026.

[2] BRASIL. **Lei nº 10.695, de 1º de julho de 2003**. Altera o art. 184 do Decreto-Lei nº 2.848, de 7 de dezembro de 1940 – Código Penal, e dá outras providências. Brasília: Presidência da República, 2003. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2003/l10.695.htm. Acesso em: 28 abr. 2026.

[3] CONTRIBUCIONES A LAS CIENCIAS SOCIALES. **Impactos da pirataria nos setores criativos e nos direitos autorais**. Disponível em: https://ojs.revistacontribuciones.com/ojs/index.php/clcs/article/download/22802/13032/61320. Acesso em: 28 abr. 2026.

[4] MOON, S.-P. et al. **Embedded Linux Implementation on a Commercial Digital TV System**. IEEE Transactions on Consumer Electronics, v. 49, n. 4, p. 1402-1407, nov. 2003. Disponível em: https://webpages.charlotte.edu/~jmconrad/EmbeddedSystems/Embedded%20Linux%20Implementation%20on%20a%20Digital%20TV%20system.pdf. Acesso em: 28 abr. 2026.

[5] QUEVEDO, M.; GALANTE, G. **Transformando TV Boxes Apreendidas Pela Receita Federal em Computadores Educacionais**. Anais do XXI Congresso Latino-Americano de Software Livre e Tecnologias Abertas (Latinoware), p. 496-499, nov. 2024. Disponível em: https://sol.sbc.org.br/index.php/latinoware/article/view/31590. Acesso em: 28 abr. 2026.

[6] MARTINS, A. et al. **Análise de desempenho de aparelhos TV Box descaracterizados para uso educacional**. ResearchGate, 2024. Disponível em: https://sol.sbc.org.br/index.php/latinoware/article/view/31552. Acesso em: 28 abr. 2026.

[7] NATURE SCIENTIFIC REPORTS. **Reaproveitamento de TV Boxes para Aplicação de Contagem de Pessoas na Borda em Cidades Inteligentes.**. Scientific Reports, v. 14, n. 122, 2024. Disponível em: https://sol.sbc.org.br/index.php/courb/article/view/30000. Acesso em: 28 abr. 2026.

[8] SOUZA, F. T. et al. **Transformando TVs BOX em Minicomputadores para Implementação de um Laboratório de Informática em uma Escola Pública da Zona Rural**. Anais do Workshop de Inclusão Digital (WID/CSBC), 2024. Disponível em: https://sol.sbc.org.br/index.php/wics/article/view/29513. Acesso em: 28 abr. 2026.

[9] SILVA, J. et al. **Recaracterização de aparelhos TV Box para uso como minicomputadores**. ENEDS – Encontro Nacional de Engenharia e Desenvolvimento Social, 2023. Disponível em: https://anais.eneds.org.br/index.php/eneds/article/view/958. Acesso em: 28 abr. 2026.

[10] PATTERSON, D. A.; GIBSON, G.; KATZ, R. H. **A Case for Redundant Arrays of Inexpensive Disks (RAID)**. ACM SIGMOD Record, v. 17, n. 3, p. 109-116, jun. 1988. Disponível em: https://www.cs.cmu.edu/~garth/RAIDpaper/Patterson88.pdf. Acesso em: 28 abr. 2026.

[11] UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP). **Computadores de baixo custo para o ensino de programação em escolas públicas utilizando TV Boxes**. Relatório Técnico PFG-23-47, IC-UNICAMP, 2023. Disponível em: https://ic.unicamp.br/~reltech/PFG/2024/PFG-24-37.pdf. Acesso em: 28 abr. 2026.

[12] UNIVERSIDADE DO OESTE DE SANTA CATARINA (UNOESC). **Projeto ECOA UNOESC**. Disponível em: https://www.unoesc.edu.br/projeto-ecoa-unoesc/. Acesso em: 28 abr. 2026.

[13] LI, S. et al. **Development of an embedded operating system based on the Linux kernel for SoC FPGA**. Journal of Physics: Conference Series, v. 2023, n. 1, p. 012015, set. 2021. Disponível em: https://ceur-ws.org/Vol-3702/paper31.pdf. Acesso em: 28 abr. 2026.

[14] DEVMFC. **debian-on-amlogic**: Debian images for Amlogic TV Boxes. Versão 2.4.0. GitHub, 2024. Disponível em: https://github.com/devmfc/debian-on-amlogic. Acesso em: 9 mar. 2026.

[15] DEBIAN. **Preparing Files for USB Stick Booting**. Debian Installation Guide, 2024. Disponível em: https://www.debian.org/releases/forky/arm64/ch02s04.en.html.

[16] ASUS. **Como inicializar o sistema a partir da unidade flash USB/CD-ROM**. Asus Support, 2023. Disponível em: https://www.asus.com/br/support/faq/1013017/.

[17] BICOM SYSTEMS. **How to Create USB Installation Media**. SERVERware Wiki, 2024. Disponível em: https://wiki.bicomsystems.com/serverware/howto/Create-USB-Installation-Media#:~:text=Install%20it%20on%20the%20system,Ready%20to%20Use.

[18] ROSENBLUM, M.; OUSTERHOUT, J. K. **The Design and Implementation of a Log-Structured File System**. ACM Transactions on Computer Systems, v. 10, n. 1, p. 26-52, fev. 1992. https://people.eecs.berkeley.edu/~brewer/cs262/LFS.pdf.

[19] UNIVERSIDADE ESTADUAL PAULISTA (UNESP). Arquitetura ARM. Disponível em: https://www.dcce.ibilce.unesp.br/~aleardo/cursos/arqcomp/Semin_ARM.pdf. Acesso em: 28 abr. 2026.

[20] UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP). Set-Top Box. Disponível em: https://www.ic.unicamp.br/~rodolfo/Cursos/mo401/2s2005/Trabalho/039632-settopbox.pdf. Acesso em: 28 abr. 2026.

[21] DEBIAN PROJECT. Debian: The Universal Operating System. Disponível em: *https://www.debian.org*. Acesso em: 28 abr. 2026.

[22] MATHUR, Avantika; CAO, Mingming; BHATTACHARYA, Suparna; DILGER, Andreas; TOMAS, Alex; VIVIER, Laurent. The New Ext4 Filesystem: Current Status and Future Plans. Disponível em: https://www.kernel.org/doc/ols/2007/ols2007v1-pages-21-33.pdf. Acesso em: 28 abr. 2026.

[23] MICROSOFT. Microsoft Extensible Firmware Initiative FAT32 File System Specification. Disponível em: https://download.microsoft.com/download/1/6/1/161ba512-40e2-4cc9-843a-923143f3456c/fatgen103.doc. Acesso em: 28 abr. 2026.

[24] TOYER, Justin. Win32 Disk Imager. Disponível em: https://sourceforge.net/projects/win32diskimager/. Acesso em: 28 abr. 2026.

[25] DAS, U-Boot Developers. Das U-Boot: The Universal Boot Loader. Disponível em: https://www.denx.de/wiki/U-Boot. Acesso em: 28 abr. 2026.

[26] ARM HOLDINGS. **Cortex-A Series Programmer's Guide: Functional Description - Technical Overview Components**. ARM Developer Documentation, 2024. Disponível em: https://developer.arm.com/documentation/100442/0100/functional-description/technical-overview/components. Acesso em: 12 mar. 2026.

[27] JAIN, Vinayak; SUDAKARAN, Sneha; SENTHIVEL, Saranyan. **First Steps, Lasting Impact: Platform-Aware Forensics for the Next Generation of Analysts**. arXiv preprint arXiv:2602.00160, 2026. Disponível em: https://arxiv.org/abs/2602.00160. Acesso em: 28 abr. 2026.

Desenvolvimento de um Sistema Embarcado com ESP32 para Aquisição e Monitoramento em Nuvem de Sinais de Qualidade de Energia

Isabela Cristina Munzlinger[1], Nakul S.[2], Marconi Januário[1],
Jamuna Kamaraj[2], S.Angalaeswari[2], Renato Gregolon Scortegagna[1]
[1]Universidade do Oeste de Santa Catarina - UNOESC, Joaçaba - Santa Catarina - Brasil
[2]Vellore Institute of Technology - VIT, Campus Chennai, Tamil Nadu - Índia
isabela.munzlinger@unoesc.edu.br, nakul.s2022@vitstudent.ac.in, marconi.januario@unoesc.edu.br,
jamuna.k@vit.ac.in,angalaeswari.s@vit.ac.in, renato.scortegagna@unoesc.edu.br

***Abstract*—This paper presents the development of a low-cost embedded system for monitoring power quality signals, integrated into a Smart Grid context. The project is part of an international collaboration between Unoesc (Brazil) and VIT Chennai (India), formalized by a Memorandum of Understanding (MOU). The proposed architecture utilizes an ESP32-WROOM-32 microcontroller to acquire and process distorted waveform data, two cases are presented, one with a measured frequency of 60 Hz and another with a simulated waveform with a fundamental frequency of 60 Hz with the superposition of a third harmonic of 180 Hz. Data persistence is ensured through the internal SPIFFS file system, while real-time synchronization is achieved via the Google Firebase cloud platform using JSON structuring. Experimental results validated through a dashboard demonstrate that the system is a viable and scalable solution for remote energy monitoring in distributed networks**

***Index Terms*—Cloud computing, Embedded systems, ESP32, Power quality, Smart grids.**

***Resumo*—Este trabalho apresenta o desenvolvimento de um sistema embarcado de baixo custo para o monitoramento de sinais de qualidade de energia, integrado ao contexto de *Smart Grids*. O projeto integra uma colaboração internacional entre a Unoesc (Brasil) e o VIT Chennai (Índia), formalizada por um MOU (*Memorandum of Understanding*). A arquitetura proposta utiliza o microcontrolador ESP32-WROOM-32 para aquisição e processamento de formas de onda, são apresentados dois casos, um com frequência medida de 60 Hz e outro com forma de onda simulada com frequência fundamental de 60 Hz com sobreposição de uma terceira harmônica de 180 Hz. A persistência dos dados é garantida pelo sistema de arquivos SPIFFS, enquanto a sincronização em tempo real ocorre via plataforma Google Firebase com estruturação JSON. Resultados experimentais validados por um *dashboard* demonstram que o sistema é uma solução viável e escalável para o monitoramento remoto de energia em redes distribuídas.**

***Palavras chave*—Computação em nuvem, ESP32, Qualidade de energia, Redes elétricas inteligentes, Sistemas embarcados.**

I. Introdução

A evolução dos sistemas elétricos convencionais para as redes elétricas inteligentes, conhecidas como *Smart Grids*, representa uma mudança de paradigma na gestão de energia. Diferente das redes tradicionais, as *Smart Grids* utilizam tecnologias de comunicação bidirecional e Internet das Coisas (IoT) para otimizar a distribuição, permitir a integração de fontes renováveis e garantir a confiabilidade do sistema. Neste cenário, o monitoramento em tempo real da qualidade da energia torna-se essencial para identificar distúrbios, como as harmônicas, que podem comprometer a eficiência e a vida útil dos equipamentos conectados.

Este trabalho é fruto de um projeto de pesquisa colaborativo internacional entre a Universidade do Oeste de Santa Catarina (UNOESC), envolvendo os cursos de Engenharia de Computação e Engenharia Elétrica, e o Vellore Institute of Technology (VIT), do campus de Chennai, na Índia. Esta parceria é formalizada por meio de um MOU (*Memorandum of Understanding*), estabelecendo cooperação em ensino e pesquisas.

O foco da contribuição brasileira neste projeto reside no desenvolvimento de um sistema embarcado de baixo custo, utilizando o microcontrolador ESP32, integrado à plataforma Google Firebase [1] para a gestão de dados em nuvem. O objetivo central é validar uma arquitetura capaz de capturar e transmitir sinais senoidais reais para monitoramento remoto, servindo como uma ferramenta de diagnóstico para distorções harmônicas em infraestruturas de *Smart Grids*.

O recorte apresentado neste artigo foca especificamente na contribuição brasileira para esta rede, concentrando-se no desenvolvimento e validação técnica do nó de aquisição de baixo custo. Para isso, utilizou-se o microcontrolador ESP32 integrado à plataforma Google Firebase para a gestão de dados. O trabalho realizado consistiu na implementação do *firmware* para captura de sinais senoidais, configuração do sistema de arquivos SPIFFS para persistência local e estruturação da sincronização JSON com o banco de dados em nuvem.

Nesta etapa da pesquisa, o foco reside na garantia da integridade e disponibilidade dos dados na nuvem, validados por meio de um *dashboard* de visualização local. Portanto, embora o sistema tenha sido projetado para a cooperação internacional, a análise efetiva das amostras pela equipe na Índia constitui uma fase subsequente, para a qual este artigo estabelece a base tecnológica necessária através da consolidação do canal de comunicação e armazenamento.

II. Fundamentação teórica

A. Smart Grids

O conceito de *Smart Grids* vai além de uma tecnologia que pode ser aplicada, mas sim pode ser entendido como um conjunto de técnicas usadas para coletar informações da rede elétrica, a fim de melhorar a qualidade dela, permitindo a otimização da performance da rede [2].

Podem ser atribuídas algumas características para as *Smart Grids*, sendo elas a auto-recuperação, permitindo que a rede analise falhas e se restaure de forma independente; tolerância à ataques externos, promovendo a cibersegurança; qualidade de energia; habilitar diversas fontes de geração de energia com o uso do *plug-and-play*; viabilizar e beneficiar o mercado competitivo, favorecendo a micro-geração [3].

B. Distorções Harmônicas e Qualidade de Energia

Uma harmônica é um componente de uma onda periódica, a qual representa a frequência múltipla da frequência fundamental, no Brasil a fundamental é de 60 Hz. Além disso, é possível afirmar que uma harmônica não fundamental produz distorções em um sinal senoidal [4].

Na Figura 1, retirada do Manual de Qualidade de Energia Eléctrica, da Universidade de Coimbra [5] são apresentados sinais de onda em diferentes harmônicas, com eles é possível perceber que nos valores ímpares, a onda começa a apresentar ruídos a medida que a ordem aumenta.



Fig. 1. Formas de ondas com harmônica fundamental e suas múltiplas. Fonte: Universidade de Coimbra, (2005). [5]

Existem múltiplos da harmônica fundamental pares e ímpares, mas é possível notar que nos harmônicos ímpares a distorção da rede é maior [4]. Existem algumas razões para a distorção harmônica acontecer, sendo elas a presença de cargas não lineares como os retificadores estáticos e os variadores eletrônicos de velocidade, que alteram a natureza da corrente nos sistemas elétricos. Além disso, a utilização de fontes de alimentação comutadas em equipamentos informáticos e de telecomunicações, juntamente com sistemas de iluminação por lâmpadas de descarga, são fatores determinantes para o surgimento de harmônicos. Outras causas relevantes incluem a operação de fornos de arco e de indução, a saturação magnética em transformadores de potência e as características de funcionamento de máquinas elétricas rotativas [5].

C. ESP32

O microcontrolador ESP32, da empresa Espressif Sistemas [6], especificamente a série ESP32-WROOM-32, é uma solução de baixo custo e alto desempenho amplamente adotada em arquiteturas de Internet das Coisas (IoT) para monitoramento em tempo real. No contexto de sistemas de gestão energética, este dispositivo atua como o elemento central da camada de campo, sendo responsável pela aquisição, processamento preliminar e transmissão de dados sensoriais [7].

O microcontrolador baseia-se nas seguintes especificações: Processamento e Desempenho, ele é equipado com o microprocessador Dual-Core 32-bit LX6, com frequência de *clock* ajustável entre 80 MHz e 240 MHz. Essa arquitetura *dual-core* permite que um núcleo seja dedicado à pilha de protocolos de comunicação enquanto o outro executa a lógica da aplicação, garantindo estabilidade no processamento de sinais senoidais [6]. Além disso, ele possui memória e armazenamento favoráveis, contando com 4 MiB de memória *flash* externa, via interface SPI, o que possibilita o armazenamento temporário de grandes volumes de dados de série temporal (*logs*) antes do envio para a nuvem. Ademais, o microcontrolados oferece suporte nativo a Wi-Fi 802.11 b/g/n (com taxas de até 150 Mbps) e *Bluetooth Dual-Mode* (v4.2 BR/EDR e BLE). Essa versatilidade de rede é essencial para a integração com plataformas de *cloud computing*, permitindo a transmissão de arquivos estruturados em formato JSON via protocolos como HTTP ou MQTT.

O *hardware* (divisor de tensão) realiza o condicionamento de sinal para adequação à faixa dinâmica do ADC. Complementarmente, o *software* realiza o ajuste de calibração, subtraindo o nível de bias efetivamente lido. Esse ajuste é crucial, pois, conforme a folha de dados do ESP32 [6], as variações de fabricação podem introduzir erros de até $\pm 6\%$, tornando necessária a identificação do ponto zero por *software* para garantir a reconstrução harmônica.

A aplicação do ESP32 em sistemas de monitoramento inteligente reduz o erro humano associado a leituras manuais e provê uma interface de dados em tempo real para os usuários finais. A capacidade de processamento local combinada com a facilidade de autenticação e envio de dados via função *push* consolida este microcontrolador como uma ferramenta poderosa para a centralização de informações de qualidade de energia em ambientes distribuídos [7].

D. Firebase Realtime Database

O Firebase Realtime Database [1] é um banco de dados NoSQL. Segundo as especificações técnicas disponíveis na plataforma do Firebase [1], ele é hospedado na nuvem que armazena dados no formato JSON e os sincroniza em tempo real com todos os clientes conectados. Esta plataforma permite que as atualizações de dados sejam propagadas instantaneamente, além de oferecer suporte nativo à persistência local. Por meio do SDK [1] do Firebase, o sistema é capaz de preservar as informações alteradas mesmo em períodos de instabilidade de rede ou operação *offline*, sincronizando-as automaticamente assim que a conectividade é restabelecida. No contexto do monitoramento de sinais senoidais, essa característica assegura que o histórico cronológico das medições permaneça íntegro e contínuo [1].

III. Desenvolvimento

O sistema de monitoramento foi desenvolvido em etapas progressivas, focando na integridade dos sinais senoidais desde a geração até o armazenamento em nuvem. O projeto integra *hardware* e *software* em uma arquitetura de três camadas: dispositivo de campo (Unoesc), plataforma *cloud* (Firebase) e equipe de análise (VIT Chennai).

Embora o diagrama ilustre o fluxo completo previsto no acordo de cooperação, o presente trabalho delimita-se à implementação e validação técnica das duas primeiras camadas. O objetivo é consolidar a infraestrutura de IoT necessária para que, em etapas futuras da pesquisa, os dados brutos armazenados em nuvem possam ser consumidos via API REST pela equipe de análise na Índia, garantindo a rastreabilidade e a fidelidade das formas de onda coletadas no Brasil. Na Figura 2 é apresentado um diagrama que mostra o fluxo geral do projeto, indicando as etapas do que será desenvolvido e os agentes envolvidos.

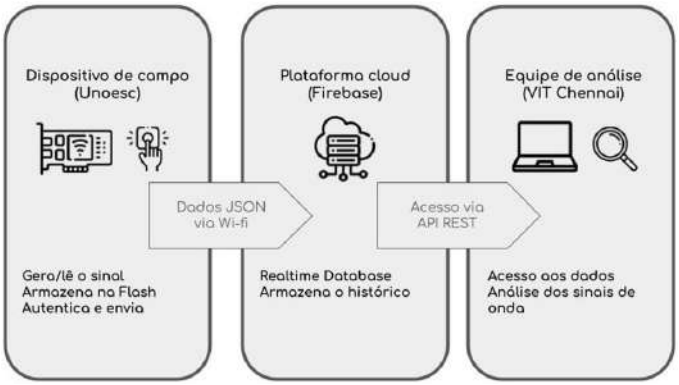


Fig. 2. Arquitetura do sistema de monitoramento colaborativo UNOESC-VIT.

O microcontrolador ESP32 foi adotado como o dispositivo central devido à sua capacidade de processamento *dual-core* e conectividade Wi-Fi nativa. Devido à limitação de tensão das entradas analógicas (ADC) do ESP32, que operam em um nível lógico de $3,3$ V, implementou-se um divisor de tensão na matriz de contato (*protoboard*), conforme ilustrado na Figura 3.

Este estágio de condicionamento de sinal foi necessário para adequar a amplitude do sinal proveniente do gerador de sinal ao limite suportado pelo microcontrolador, garantindo a integridade do *hardware* e a precisão da leitura analógica. Inicialmente, realizou-se a simulação de uma onda senoidal puramente matemática via código, simulando uma forma de onda de 60 Hz com a sobreposição de uma terceira harmônica de 180 Hz, permitindo validar o algoritmo de fragmentação de dados no formato JSON. Em uma segunda etapa, utilizou-se um osciloscópio para monitorar e validar o sinal físico de uma onda senoidal de 60 Hz, assegurando que a forma de onda processada pelo microncontrolador fosse uma representação fiel da realidade elétrica, conforme os parâmetros de qualidade de energia.

Conforme detalhado no fluxograma da Figura 4, a transmissão de dados foi configurada para acontecer sob demanda. Ou seja, implementou-se um gatilho manual via interrupção digital no ESP32: ao pressionar o botão físico, o sistema inicia

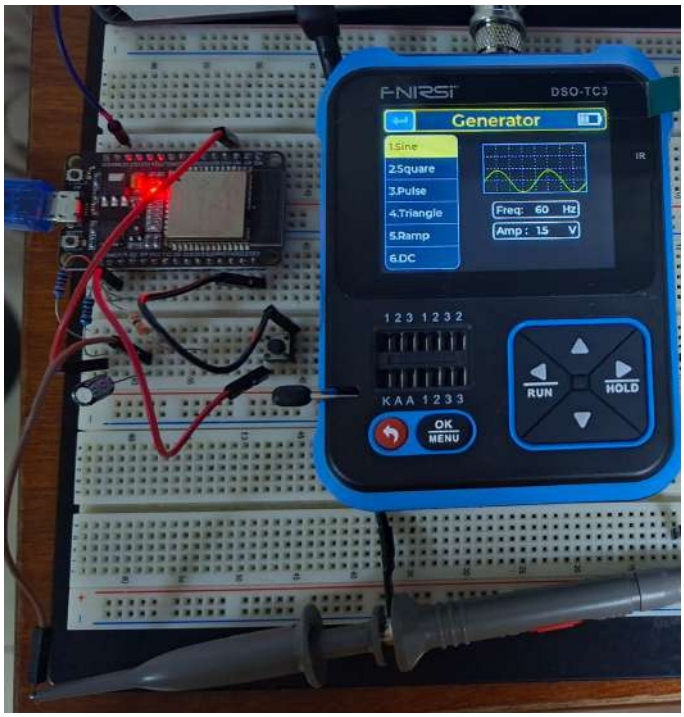


Fig. 3. Circuito montado na matriz de contatos conectado a um gerador de sinal.

a leitura da memória *flash* onde os pontos da senóide estão temporariamente armazenados utilizando o sistema de arquivos SPIFFS, sendo sobreescritos a todo momento enquanto não houver interrupção. Depois disso, o dispositivo realiza a autenticação segura com o servidor por meio de uma chave de autorização específica, estabelecendo um caminho para enviar as amostras coletadas.

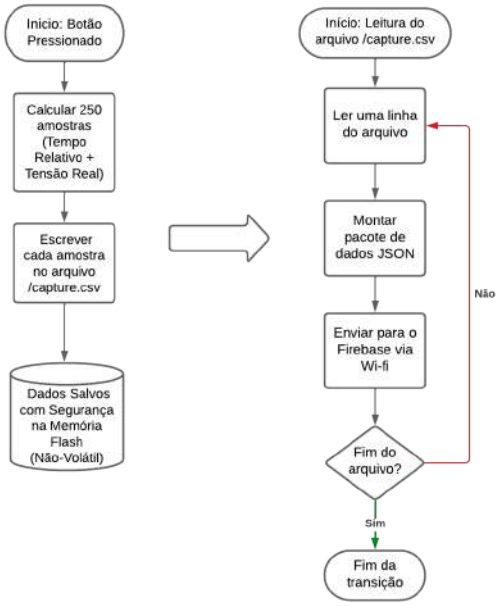


Fig. 4. Lógica de aquisição e transmissão de dados do firmware ESP32.

Uma vez estabelecida a conexão Wi-Fi, o ESP32 atua como um *publisher*, enviando o pacote de dados com 250 amostras para o Google Firebase. A organização em nuvem foi projetada

de forma hierárquica para facilitar a análise remota pela equipe.

Utilizou-se a função *push* para criar uma chave exclusiva para cada sessão de monitoramento, evitando a sobreposição de dados históricos. Além disso, cada entrada é acompanhada de um *timestamp*, permitindo rastreabilidade temporal das medições. Os dados armazenados são o tempo e a tensão, eles são ordenados em um *array* de pontos, o que viabiliza a reconstrução exata da forma de onda em *dashboards* de visualização ou ferramentas de análise externa. Essa estrutura permite que a equipe de análise na Índia acesse os dados via API REST para a reconstrução exata da forma de onda.

IV. Resultados

A. Dashboard apresentado via Firebase a partir de forma de onda de gerador de função

O sistema estabeleceu um canal de comunicação estável e confiável entre o ESP32 e o Google Firebase, validando o envio de dados estruturados em formato JSON via Wi-Fi. O uso da memória *flash* interna do microcontrolador (SPIFFS) permitiu o armazenamento prévio dos dados de série temporal, garantindo que o volume de informações fosse preservado até o momento da transmissão. A plataforma em nuvem foi capaz de processar as solicitações de escrita em tempo real, criando um histórico cronológico e persistente das medições.

A estrutura de dados no Firebase foi validada conforme planejado na metodologia. Como demonstrado na Figura 5, as informações foram organizadas sob o nó raiz capturas, onde cada sessão de monitoramento recebeu um identificador exclusivo (ex: 1758857755) gerado pela função *push*.



Fig. 5. Pastas com as amostras coletadas.

Dentro de cada nó de captura, as informações foram segmentadas em dois sub-nós principais:

- info: contendo o metadado $horario_{captura}$, que permite a rastreabilidade temporal exata da ocorrência (ex: 2025-09-26 00:35:55).
- pontos: contendo o *array* de medições, onde cada entrada associa um índice de tempo ao respectivo valor de tensão.

Comparando os valores simulados no microcontrolador com os dados recuperados via API REST na nuvem, observou-se uma divergência desprezível, validando o ESP32 como um nó de aquisição e de baixo custo para aplicações em *Smart Grids*.

Para validação dos dados coletados quando gerados pelo osciloscópio, foi gerado um *dashboard* via HTML, capaz de acessar o banco de dados remoto do Firebase e processar as 250 amostras coletadas para gerar um sinal senoidal a partir delas, conforme a Figura 6. Com isso, foi possível visualizar se os sinais gerados pelo dispositivo, trasmitidos via ESP32, e armazenados no Firebase foram salvos de maneira correta.

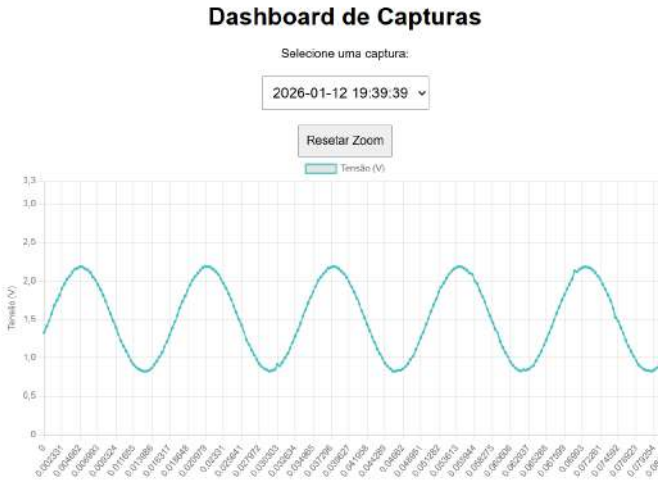


Fig. 6. *Dashboard* para validar os dados coletados.

B. Comparação de formas de onda

A seguir são apresentados dois casos onde os dados foram enviados para o Firebase, em ambos os casos a quantidade de amostras foram de 250. Descrição dos casos:

- Caso 1: forma de onda de tensão senoidal com frequência de 60 Hz, produzida pelo gerador função, cujo o sinal foi adquirido via entrada analógica do ESP32.
- Caso 2: forma de onda senoidal criada a partir de função matemática no microcontrolador, onda a forma de onda é composta por frequência com harmônica fundamental de 60 Hz e com terceira harmônica, 180 Hz.

Caso 1: Na Figura 7 são mostradas as formas de onda de tensão com frequência fundamental de 60 Hz. Figura 7 (a) são apresentadas as forma de onda completa no intervalo de tempo em estudo, 83 *m*s, e em Figura 7 (b) são mostrados detalhes das formas de ondas entre os tempos de 17 *m*s até 25 *m*s. As formas de onda apresentam um nível contínuo, tendo em vista o sinal de tensão foi obtida a partir do *hardware* (divisor de tensão) que realiza o condicionamento de sinal para adequação da faixa dinâmica do ADC, vide Figura 3. Na Figura 7 (a), nota-se que as formas de ondas "Forma de aquisição" e "Forma obtida com os dados extraídos do FIREBASE" são semelhantes.

Caso 2: Na Figura 8 são mostradas as formas de onda de tensão composta por frequência com harmônica fundamental de 60 Hz e com terceira harmônica, 180 Hz. As formas de ondas "Forma da função matemática" e "Forma obtida com os dados extraídos do FIREBASE" são similares. Assim como no caso 1, na Figura 8 (a) são apresentadas as forma de onda completa no intervalo de tempo em estudo e na Figura 8 são mostrados detalhes das formas de ondas.

O cálculo de erro médio quadrático MSE foi realizado por meio da Equação 1, onde N é o número total de amostras; i representa a iteração do número da amostra; $S_{microcontrolador,i}$ é o valor calculado no instante de tempo i no microcontrolador; e o $S_{FIREBASE,i}$ é o valor no instante i que foi obtido vindo do Firebase. Os valores de erro médio quadrático MSE para

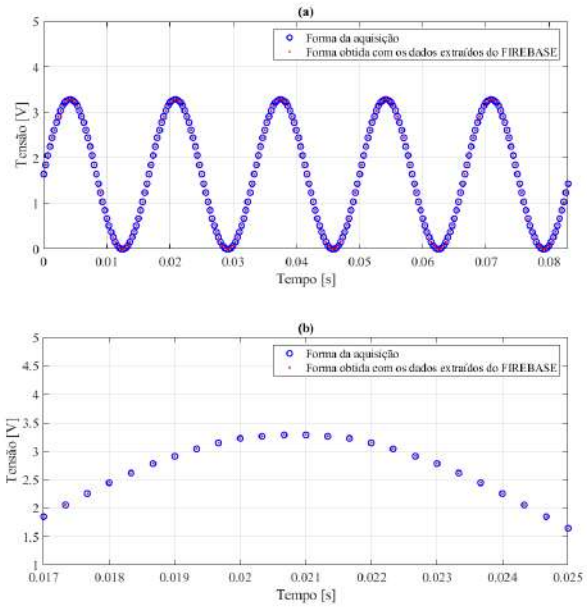


Fig. 7. Caso 1: Forma de onda oriunda da aquisição do sinal pelo ESP32 e forma de onda com os dados extraídos do FIREBASE.

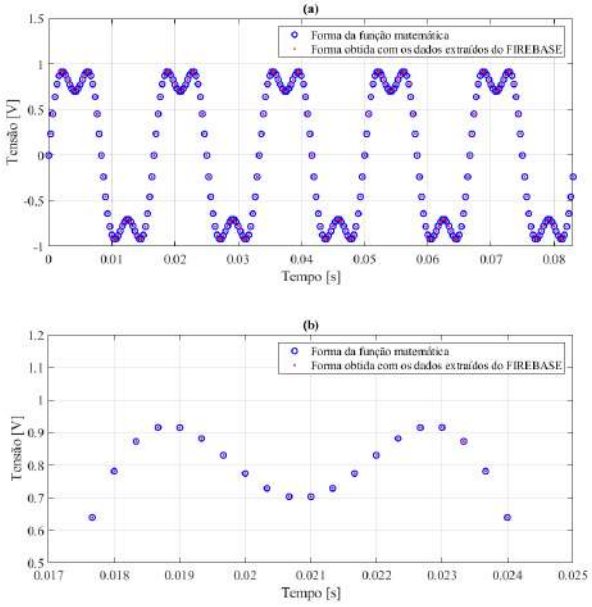


Fig. 8. Caso 2: Forma de onda da função matemática criada no microcontrolador e forma de onda com os dados extraídos do FIREBASE.

cada caso foram calculados por meio da Equação 1, cujos os resultados obtidos foram iguais a zero.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(S_{microcontrolador,i} - S_{FIREBASE,i})^2 \quad (1)$$

V. Conclusão e propostas futuras

Este trabalho apresentou o desenvolvimento e a validação de um sistema embarcado de baixo custo para o monitoramento de sinais de qualidade de energia, representando a contribuição brasileira para um projeto de pesquisa colaborativo internacional. A integração entre o *hardware* (ESP32) e a plataforma *cloud* (Firebase) mostrou-se uma solução útil para os desafios da digitalização em *Smart Grids*.

O protótipo demonstrou que o ESP32 é uma plataforma viável para atuar como um nó de aquisição distribuído, sendo capaz de processar e armazenar localmente sinais complexos de série temporal utilizando o sistema de arquivos SPIFFS. A estruturação dos dados em formato JSON e o uso da função `push` no Firebase garantiram um histórico de medições persistente e não destrutivo, essencial para a análise de distorções harmônicas.

A validação por meio do *dashboard* comprovou que a taxa de amostragem adotada (250 amostras em 5 ciclos) é suficiente para a reconstrução de ondas senoidais para os presentes casos, permitindo que a equipe de análise no VIT Chennai visualize parâmetros de rede em tempo real.

Sugere-se também efetuar um estudo quando aos possíveis erros de aquisição e de resolução dos conversores ADC presentes nos ESP32.

Como continuidade deste estudo, propõe-se a adaptação da geração de sinais para uma bancada ao invés de um gerador de sinais. Dessa forma, haverá um valor real sendo gerado através da saturação magnética de transformadores operando a vazio, o que fornecerá uma base de dados real para o aprimoramento dos algoritmos de detecção de harmônicas. Assim como, a interconexão com a equipe de Chennai para que, futuramente, os dados gerados e hospedados na nuvem possam ser analisados por eles, para que sejam feitas as devidas tratativas para o controle e manejo da qualidade energética, contemplando todas as etapas previstas no projeto global.

Referências

[1] Google. *Realtime Database*. Acesso em: 8 mar. 2026. 2026. URL: https://firebase.google.com/docs/database.

[2] Djalma M. Falcão. "Smart Grids e Microredes: o futuro já é presente". Em: *Anais do VIII SIMPASE - Simpósio de Automação de Sistemas Elétricos*. Disponível em: https://www.researchgate.net/publication/228473062. Rio de Janeiro, Brasil, Agosto de 2009.

[3] Valmir de Oliveira, Odilon Luiz Tortelli e Gunter de Souza. "Smart Grid – A Próxima Geração de Redes Elétricas". Em: *Anais do XIX CRICTE - Congresso de Iniciação Científica e Tecnológica em Engenharia*. Curitiba, Brasil, 2004. URL: http://www.cricte2004.eletrica.ufpr.br/odilon/te339/artigo_SMART_GRID.PDF.

[4] Guilherme Henrique de Oliveira. *Protótipo de um Analisador de Qualidade de Energia Elétrica utilizando a Plataforma Arduino*. Trabalho de Conclusão de Curso (Graduação). Acesso em: 8 mar. 2026. Curitiba, Brasil, 2011. URL: https://www.eletrica.ufpr.br/p/arquivostccs/38.pdf.

[5] Carlos Patrão, Aníbal T. de Almeida e Joaquim Delgado. *Manual de Qualidade de Energia Eléctrica (Power Quality Manual)*. Disponível em: https://www.researchgate.net / publication / 270577082. Coimbra, Portugal: EDP Distribuição / ISR – Universidade de Coimbra, Dezembro de 2005.

[6] Espressif Systems. *ESP32-WROOM-32 Datasheet*. Versão 3.4. 2024. URL: https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32_datasheet_en.pdf.

[7] A. S. M. Sanwar et al. "IoT-based smart energy monitoring and management system for efficient power utilization". Em: *Internet of Things and Cyber-Physical Systems* 4 (2024), pp. 120–135. DOI: 10.1016/j.iotcps.2024.01.002.

AUTORES



Isabela Cristina Munzlinger é graduanda do curso de Engenharia de Computação pela Universidade do Oeste de Santa Catarina (Unoesc), campus de Joaçaba. Possui formação técnica em Automação Industrial e atua como vice-presidente da Liga de Programação Competitiva da instituição. É membro do Grupo de Pesquisa em Aplicações Tecnológicas Industriais - GATI e integra um projeto de pesquisa internacional em colaboração com o Vellore Institute of Technology (VIT Chennai, Índia), com foco em Smart Grids e qualidade de energia. Foi finalista da Maratona Feminina de Programação (MFP). Suas áreas de interesse acadêmico e prático abrangem sistemas embarcados, Internet das Coisas (IoT), eletrônica, visão computacional, redes neurais e análise de dados.



Nakul S. é um estudante de graduação em Engenharia Elétrica e Eletrônica no Instituto de Tecnologia de Vellore (VIT), em Chennai, Índia. Seus interesses incluem eletrônica de potência, máquinas elétricas e automação industrial. Ele trabalhou em projetos envolvendo manutenção preditiva de motores de indução trifásicos e sistemas de monitoramento de saúde em tempo real. Nakul também trabalhou em atividades de pesquisa relacionadas a semicondutores e projeto de *hardware* digital. Ele adquiriu experiência industrial por meio de estágios em geração de energia térmica, sistemas de tração ferroviária e distribuição elétrica em usinas.



Marconi Januário é Engenheiro Eletricista graduado pela FURB (2003), Especialista em Automação Industrial pela FURB (2004), Mestre e Doutor em Engenharia Elétrica pela UFSC (2007 e 2015). Professor da Universidade do Oeste de Santa Catarina - UNOESC. É lider do Grupo de Pesquisa em Aplicações Tecnológicas Industriais - GATI . Trabalha e pesquisa em assuntos relacionados com seguintes temas: Transformadores, Modelagem de Transformadores, Resposta em Frequência, Máquinas Elétricas, Determinação numérica de valores de parâmetros de modelos, Qualidade de Energia, Harmônicas, Sistemas de Controle, Automação Industrial e Inteligência Artificial.



Jamuna Kamaraj, Dra., trabalha atualmente como professora no Instituto de Tecnologia de Vellore, em Chennai, Índia. Ela obteve seu bacharelado em Engenharia Elétrica e Eletrônica pela Faculdade de Engenharia Thiagarajar, em Madurai, mestrado em Engenharia de Sistemas de Potência pela Faculdade de Engenharia de Trivandrum e doutorado pelo IIT Madras na área de Sistemas de Potência. Ela possui mais de 25 anos de experiência em ensino e pesquisa. Publicou mais de 45 artigos em periódicos e conferências. É autora do livro "Redes e Circuitos Elétricos" (publicado pela AICTE). Seus interesses de pesquisa incluem estimação de estado em sistemas de potência, operação e controle de sistemas de microredes e comercialização de energia.



S.Angalaeswari S. ANGALAESWARI formou-se em Engenharia Elétrica e Eletrônica pela Faculdade de Engenharia e Tecnologia Alagappa Chettiar, em Karaikudi, Tamil Nadu, Índia, em 2002. Concluiu o mestrado em Engenharia de Sistemas de Potência com medalha de ouro pela Faculdade de Engenharia da Universidade Anna, em Chennai, Tamil Nadu, Índia, em 2009. Obteve o título de doutora pelo Instituto de Tecnologia de Vellore (VIT) em 2020, na área de microrredes. Atualmente, trabalha como professora associada na Escola de Engenharia Elétrica do VIT, campus de Chennai, Índia, desde 2014. Também atua como Decana Assistente de Pesquisa Patrocinada e Consultoria Industrial no VIT Chennai. Suas áreas de pesquisa incluem fluxo de potência ótimo em microrredes, implementação de algoritmos de otimização em microrredes e projeto de controladores ótimos para microrredes.



Renato Gregolon Scortegagna é Engenheiro Eletricista graduado pela Unoesc (2011), Mestre e Doutor em Engenharia Elétrica pela UTFPR (2015 e 2022). É membro da Associação Brasileira de Eletrônica de Potência (SOBRAEP). É professor e coordenador dos Cursos de Engenharia Elétrica e Engenharia de Computação da Unoesc Joaçaba-SC.

Gamificação No Ambiente Industrial: Uma Estratégia Para Redução Da Fadiga Mental E Aumento Da Produtividade Em Atividades Repetitivas

Ana Beatriz Faria Gomes & Vinícius A. Montgomery de Miranda
Instituto Nacional de Telecomunicações – Inatel
ana.gomes@gep.inatel.br, vmontgomery@inatel.br

Abstract— Mass production in industrial environments remains heavily dependent on human action, especially in activities characterized by short and repetitive cycles. Although this model favors efficiency and standardization, the intensification of the production pace can compromise the physical and mental health of workers, affecting performance throughout the workday. In this context, it becomes necessary to seek strategies that reconcile operational efficiency and well-being. This article analyzes the potential of gamification in the industrial environment as an alternative to mitigate the effects of repetitive work and promote greater productive balance. The study is characterized as qualitative research, of an exploratory nature, based on a review of national and international literature. The results indicate that gamification, when integrated with ergonomic practices and the principles of Industry 4.0 and 5.0, can contribute to increasing engagement, attention, and understanding of the production process, provided it is applied in a balanced way.

Index Terms—Gamification, mental fatigue, ergonomics, industrial productivity.

Resumo—A produção em massa em ambientes industriais permanece fortemente dependente da atuação humana, especialmente em atividades caracterizadas por ciclos curtos e repetitivos. Embora esse modelo favoreça a eficiência e a padronização, a intensificação do ritmo produtivo pode comprometer a saúde física e mental dos trabalhadores, afetando o desempenho ao longo da jornada. Nesse contexto, torna-se necessário buscar estratégias que conciliem eficiência operacional e bem-estar. Este artigo analisa o potencial da gamificação no ambiente industrial como alternativa para mitigar os efeitos do trabalho repetitivo e favorecer maior equilíbrio produtivo. O estudo caracteriza-se como uma pesquisa qualitativa, de natureza exploratória, baseada em revisão da literatura nacional e internacional. Os resultados indicam que a gamificação, quando integrada às práticas ergonômicas e aos princípios da Indústria 4.0 e 5.0, pode contribuir para aumentar o engajamento, a atenção e a compreensão do processo produtivo, desde que aplicada de forma equilibrada.

Palavras chave—Gamificação, fadiga mental, ergonomia, produtividade industrial.

I. Introdução

A produção em massa em ambientes industriais permanece amplamente dependente da atuação humana, especialmente em tarefas de ciclos curtos e repetitivos, como montagem, inspeção, embalagem e operação de máquinas, mesmo em sistemas que incorporam diferentes níveis de automação. Estruturada historicamente para maximizar eficiência, padronização e controle dos processos, essa lógica produtiva mantém elevada exigência física e cognitiva sobre os trabalhadores [1]. Entretanto, a repetição contínua de movimentos e a necessidade de sustentar ritmos intensos ao longo da jornada impõem limites relevantes ao desempenho humano. Evidências recentes indicam que cargas físicas e mentais elevadas, associadas à repetitividade, contribuem para o aumento da fadiga, para a maior incidência de distúrbios musculoesqueléticos e para a elevação da probabilidade de erros operacionais, afetando simultaneamente a eficiência do sistema produtivo e o bem-estar dos trabalhadores [2]. Nesse contexto, a incorporação sistemática dos princípios de fatores humanos ao projeto de processos e postos de trabalho revela-se fundamental para promover produtividade sustentável, alinhando desempenho organizacional à preservação da saúde ocupacional.

No âmbito dos sistemas produtivos, esse contexto revela um dilema estrutural: a priorização da eficiência operacional, quando desvinculada das capacidades e limitações físicas e cognitivas dos trabalhadores, tende a comprometer a sustentabilidade do desempenho organizacional. A literatura indica que condições de trabalho caracterizadas por elevadas demandas físicas e psicossociais estão associadas a diferentes impactos na saúde dos trabalhadores, incluindo problemas musculoesqueléticos, estresse ocupacional e outros desdobramentos relacionados ao bem-estar físico e mental [3]. Esse desalinhamento entre exigências do sistema e recursos humanos disponíveis pode resultar em perdas de qualidade, maior incidência de erros e redução da confiabilidade operacional.

Os impactos da repetitividade, portanto, extrapolam a dimensão biomecânica e alcançam a esfera psicossocial do trabalho. A monotonia inerente à execução contínua de tarefas padronizadas, combinada à pressão por metas e ao controle rígido do ritmo produtivo, pode gerar desgaste psicológico e comprometer progressivamente os níveis de atenção e desempenho ao longo da jornada de trabalho, especialmente em atividades com baixa variedade e reduzida autonomia [4, 5]. Em cenários caracterizados por elevada repetitividade e baixa variabilidade de tarefas, o desgaste psicológico pode comprometer progressivamente a atenção e o engajamento no trabalho. Nesse contexto, abordagens que integrem tecnologias

digitais e estratégias voltadas à motivação dos trabalhadores têm sido discutidas no âmbito da Indústria 5.0, conceito que amplia os avanços tecnológicos da Indústria 4.0 ao enfatizar a centralidade do ser humano, o bem-estar dos trabalhadores e a colaboração entre pessoas e sistemas inteligentes. Dessa forma, busca-se tornar os ambientes produtivos mais equilibrados, sustentáveis e orientados às necessidades humanas. [6]. Esses achados reforçam a necessidade de abordagens integradas que conciliem eficiência técnica, saúde ocupacional e gestão estratégica de pessoas, especialmente em ambientes industriais orientados por metas de alto desempenho.

Paralelamente às evidências acadêmicas, observa-se crescente interesse empresarial na adoção de mecanismos inovadores para mitigar os efeitos da monotonia e promover maior engajamento no chão de fábrica. Entre essas iniciativas, destaca-se a incorporação de elementos inspirados em jogos digitais, como metas progressivas, *feedback* em tempo real e reconhecimento simbólico, com o propósito de tornar tarefas repetitivas mais envolventes e significativas. Experiências internacionais, incluindo casos divulgados na indústria japonesa, indicam potencial para aumentar o engajamento e a percepção de autonomia, embora ainda demandem avaliação empírica sistemática quanto a seus impactos de longo prazo. Esse movimento evidencia a relevância prática do tema e fundamenta o problema de pesquisa que orienta o presente estudo, ao buscar compreender em que medida a gamificação pode contribuir para o equilíbrio entre desempenho produtivo e bem-estar no ambiente industrial [7].

Diante desse contexto, emerge a seguinte questão de pesquisa: em que medida estratégias fundamentadas na gamificação, quando integradas ao processo produtivo, podem contribuir para a redução da fadiga mental sem intensificar a pressão por desempenho em ambientes industriais? A gamificação é compreendida como a aplicação de elementos, dinâmicas e mecânicas típicas de jogos em contextos não lúdicos, com o propósito de promover engajamento, motivação e mudança comportamental [8]. No âmbito organizacional, essa abordagem tem sido explorada como instrumento de gestão capaz de estruturar metas, fornecer *feedback* contínuo e tornar tarefas operacionais mais significativas ao trabalhador.

Ao incorporar recursos como sistemas de pontuação, níveis progressivos, indicadores visuais de desempenho e recompensas simbólicas por meio de plataformas digitais conectadas às máquinas, a gamificação pode contribuir para ampliar a percepção de progresso, a clareza de objetivos e o engajamento dos trabalhadores no processo produtivo [9]. Entretanto, sua implementação em ambientes industriais exige cautela, uma vez que mecanismos mal estruturados podem reforçar a competição excessiva e intensificar a pressão por resultados, produzindo efeito contrário ao pretendido. Assim, investigar o equilíbrio entre engajamento, bem-estar e desempenho torna-se fundamental para avaliar a efetividade da gamificação como estratégia de apoio à gestão da produção.

II. REFERENCIAL TEÓRICO

A. Produção em massa, repetitividade e fadiga mental

A produção em massa consolidou-se com base na padronização e na divisão do trabalho, estruturando tarefas repetitivas para maximizar a eficiência e a produtividade. A especialização decorrente desse arranjo reduz tempos improdutivos, favorece a aprendizagem cumulativa e amplia a coordenação entre agentes econômicos. Segundo Hoppe (2013) [10], a divisão do trabalho emerge das diferenças individuais de habilidades e da heterogeneidade de recursos, sendo potencializada em ambientes institucionais estáveis que favorecem a cooperação social. Nesse contexto, a especialização permite níveis superiores de produtividade e expansão da produção total.

Nesse contexto, a repetição contínua de movimentos e decisões em ciclos curtos intensifica a carga física e cognitiva ao longo da jornada. Na ergonomia cognitiva, evidências indicam que tarefas com baixa autonomia e reduzida variabilidade aumentam a fadiga mental e a probabilidade de erros. Hockey (2013) [4] demonstra que demandas repetitivas comprometem a autorregulação da atenção e do esforço, enquanto o Modelo Demanda–Recursos aponta que altas exigências associadas a escassez de suporte favorecem o esgotamento; em contrapartida, o equilíbrio entre demandas e recursos sustenta o engajamento e o desempenho [11].

Desse modo, embora a divisão do trabalho seja central para a eficiência produtiva, intervenções orientadas exclusivamente por critérios técnicos tendem a gerar ganhos limitados no longo prazo. A produtividade sustentável requer a integração entre desenho do trabalho, princípios ergonômicos e gestão de pessoas.

B. Gamificação no ambiente industrial

A gamificação tem sido progressivamente debatida como estratégia para enfrentar desafios humanos associados a ambientes de trabalho marcados pela repetitividade e pela baixa variabilidade de tarefas. De modo geral, o conceito refere-se à aplicação de elementos típicos de jogos, como metas claras, *feedback* contínuo, níveis de progressão e recompensas simbólicas, em contextos não lúdicos, com o objetivo de promover maior engajamento, motivação e sentido nas atividades desempenhadas [9]. Ao estruturar tarefas operacionais com base em dinâmicas de acompanhamento e reconhecimento, a gamificação busca influenciar comportamentos e percepções sem alterar, necessariamente, a natureza técnica do processo produtivo.

No ambiente industrial, essa abordagem torna-se particularmente relevante diante do avanço da Indústria 4.0, caracterizada pela integração entre sistemas físicos e digitais, uso de sensores, coleta e análise de dados em tempo real e elevada conectividade entre máquinas e plataformas informacionais. Esse cenário amplia as possibilidades de interação entre trabalhador e sistema produtivo, permitindo que indicadores técnicos sejam convertidos em informações visuais e interativas. Nesse contexto, a gamificação pode atuar como mecanismo de mediação cognitiva, traduzindo métricas operacionais em representações mais acessíveis e compreensíveis ao operador, favorecendo a percepção de

progresso e de contribuição individual para os resultados globais [6].

Diferentemente dos modelos tradicionais de controle da produção, centrados predominantemente em metas quantitativas e supervisão hierárquica, sistemas gamificados tendem a enfatizar o *feedback* imediato e a visualização contínua do desempenho individual ou coletivo. Evidências internacionais indicam que a incorporação de elementos gamificados e mecanismos de feedback pode ampliar o engajamento e a motivação dos usuários em diferentes contextos organizacionais, contribuindo para tornar as atividades mais envolventes e significativas [12]. Ao tornar o desempenho visível e contextualizado, a gamificação contribui para reforçar o significado do trabalho e a compreensão do impacto das atividades no fluxo produtivo.

No contexto brasileiro, embora a produção científica ainda seja incipiente, estudos recentes apontam resultados promissores quanto à aplicação da gamificação em ambientes produtivos, especialmente na construção civil. Evidências empíricas indicam que a adoção de estratégias gamificadas pode contribuir para o aumento da motivação e do engajamento dos trabalhadores, com destaque para a importância do feedback contínuo como elemento central nesse processo. Pesquisas mostram que uma parcela significativa dos trabalhadores não recebe retorno sobre seu desempenho nem reconhecimento adequado, o que contribui para dificuldades de comunicação, desmotivação e percepção de repetitividade no trabalho. Nesse sentido, a utilização de mecanismos que promovam a explicitação de metas, o fornecimento de feedback e o incentivo à colaboração tende a tornar o ambiente de trabalho mais engajador e alinhado às necessidades dos trabalhadores [13].

Assim, compreendida não como instrumento de intensificação do controle, mas como ferramenta de apoio ao desempenho humano, a gamificação apresenta potencial para aprimorar simultaneamente a experiência de trabalho e a sustentabilidade da produtividade. Sua articulação com os princípios da Indústria 4.0 reforça a possibilidade de desenvolvimento de sistemas produtivos mais inteligentes, eficientes e orientados à valorização das capacidades humanas.

III. METODOLOGIA

O presente estudo configura-se como uma pesquisa qualitativa, de natureza exploratória e descritiva, fundamentada em revisão bibliográfica da literatura nacional e internacional sobre produção em massa, ergonomia, saúde mental no trabalho e gamificação aplicada ao contexto industrial. De acordo com Gil (2008) [14], pesquisas exploratórias e descritivas são frequentemente utilizadas quando se busca ampliar a compreensão sobre determinado fenômeno por meio da análise e sistematização do conhecimento existente na literatura.

A busca bibliográfica foi conduzida em bases de dados científicas consolidadas, com ênfase em publicações no período de 2015 a 2025. Foram selecionados artigos revisados por pares, revisões sistemáticas e estudos de natureza conceitual que examinassem a interface entre trabalho repetitivo, fatores psicossociais e estratégias de intervenção em ambientes industriais.

IV. DISCUSSÃO

A literatura analisada indica que a repetitividade característica da produção em massa está associada ao aumento da fadiga física e mental, com impactos diretos sobre a atenção, a motivação e a consistência do desempenho ao longo da jornada de trabalho. Estudos em ergonomia e psicologia do trabalho demonstram que a exposição prolongada a tarefas repetitivas e com baixa variabilidade compromete os mecanismos de autorregulação da atenção e do esforço, elevando a probabilidade de lapsos e erros operacionais [4]. Além disso, pesquisas clássicas sobre desenho do trabalho já apontavam que baixos níveis de autonomia e variedade tendem a reduzir a motivação intrínseca e a qualidade do desempenho [5].

Esses achados evidenciam que a produtividade não pode ser examinada exclusivamente sob a ótica da eficiência técnica, devendo incorporar a consideração dos limites físicos e cognitivos que condicionam o desempenho humano no sistema produtivo, sob pena de comprometer a sustentabilidade dos resultados organizacionais no longo prazo.

Nesse contexto, a gamificação desponta como estratégia potencial para atuar sobre dimensões psicossociais do trabalho repetitivo, funcionando como complemento às intervenções ergonômicas tradicionais. Ao estruturar metas claras, disponibilizar feedback contínuo e ampliar a percepção de controle e progresso, sistemas gamificados podem favorecer o engajamento e fortalecer o sentido atribuído às atividades desempenhadas, contribuindo para mitigar a monotonia inerente a tarefas altamente padronizadas.

Evidências provenientes da literatura sobre jogos educacionais na Engenharia de Produção indicam que abordagens baseadas em jogos e dinâmicas gamificadas tendem a tornar as atividades mais interativas e envolventes, favorecendo o engajamento e o desenvolvimento de competências relacionadas à resolução de problemas e à tomada de decisão. Revisões sistemáticas da literatura mostram que a utilização de jogos tem sido aplicada em diferentes áreas da Engenharia de Produção, como manufatura enxuta, logística e pesquisa operacional, demonstrando potencial para apoiar processos de aprendizagem e interação em ambientes organizacionais e educacionais [15].

Todavia, a literatura adverte que a gamificação não deve ser tratada como solução universal para os desafios do trabalho repetitivo. Experiências divulgadas em mídia especializada, como a adoção de mecanismos inspirados em jogos em uma indústria japonesa, ilustram simultaneamente o potencial e as limitações dessa abordagem. Nesse contexto, a ampliação de pesquisas empíricas no chão de fábrica representa, portanto, um campo promissor para avaliar, com maior rigor metodológico, os efeitos dessa estratégia sobre desempenho, engajamento e saúde ocupacional.

V. CONSIDERAÇÕES FINAIS

O presente artigo analisa o potencial da gamificação aplicada ao ambiente industrial como estratégia para mitigar os efeitos adversos do trabalho repetitivo, particularmente no que se refere à fadiga mental e à estabilidade do desempenho. A revisão da literatura evidenciou que a repetitividade excessiva, quando dissociada dos limites físicos e cognitivos dos

trabalhadores, tende a comprometer simultaneamente a saúde ocupacional e a eficiência do sistema produtivo, revelando a insuficiência de abordagens centradas exclusivamente na dimensão técnica da produtividade.

Conclui-se que a gamificação, quando concebida de forma integrada às práticas ergonômicas e alinhada aos princípios da Indústria 4.0, pode contribuir para a construção de ambientes de trabalho mais sustentáveis. Ao favorecer feedback contínuo, clareza de metas e maior percepção de controle sobre as atividades, essa abordagem apresenta potencial para fortalecer o engajamento, a atenção e a compreensão do processo produtivo por parte dos trabalhadores. Contudo, sua implementação exige planejamento criterioso, desenho participativo e avaliação contínua, a fim de evitar que se converta em instrumento de intensificação do trabalho ou de ampliação da carga mental.

Como limitação, destaca-se o caráter teórico do estudo, fundamentado exclusivamente em revisão bibliográfica. Recomenda-se, portanto, o desenvolvimento de pesquisas empíricas em contextos industriais brasileiros, com delineamentos quantitativos e qualitativos que permitam avaliar de maneira mais robusta os efeitos da gamificação sobre indicadores de saúde mental, engajamento e produtividade.

Em síntese, a gamificação não deve ser compreendida como solução tecnológica isolada, mas como componente de uma estratégia organizacional mais ampla, orientada à harmonização entre desempenho operacional e sustentabilidade humana nos sistemas produtivos contemporâneos.

VI. AGRADECIMENTOS

Os autores agradecem à Fundação Instituto Nacional de Telecomunicações (FINATEL) pelo apoio institucional e incentivo às atividades de pesquisa que tornaram possível o desenvolvimento deste trabalho.

REFERENCIAS

[1] N. Slack, A. Brandon-Jones, N. Burgess. Operations management. 10. ed. Harlow: Pearson, 2022.

[2] B. Hasanain. The role of ergonomic and human factors in sustainable manufacturing: a review. Machines, v. 12, n. 3, p. 159, 2024. DOI: https://doi.org/10.3390/machines12030159

[3] P. Freitas, M. Lopes, A. Assunção, A. Lopes. Health and work in Brazil: physical and psychosocial demands, v. 9, 2021.

[4] G. R. J. Hockey. The psychology of fatigue: work, effort and control. Cambridge: Cambridge University Press, 2013.

[5] J. R. Hackman, G. R. Oldham. Motivation through the design of work: test of a theory. Organizational Behavior and Human Performance, New York, v. 16, n. 2, p. 250–279, 1976.

[6] L. Cónego, R. Pinto, J. Pinto, G. Gonçalves. Leveraging Gamification in Industry 5.0: Tailored Solutions for Workplace' Employees, v. 232, p. 1769-1778, 2024.

[7] A. Bordin. Fábrica aposta em videogame para engajar trabalhadores em tarefas repetitivas. Portal Japão, 5 ago. 2025. Disponível em: https://portaljapao.com/fabrica-aposta-em-videogame-para-engajar-trabalhadores-em-tarefas-repetitivas. Acesso em: 11 mar. 2026.

[8] S. Deterding, D. Dixon, R. Khaled, L. Nacke. From game design elements to gamefulness: defining gamification. In: Proceedings of the 15th International Academic MindTrek Conference. Tampere: ACM, 2011. p. 9–15.

[9] R.M.C. Leite, L. R. G. Alves, L.S.P. Cardoso, H. M. M Neto. How Has Gamification in the Production Sector Been Developed in the Manufacturing and Construction Workplaces? Buildings 2023, 13, 2614. https://doi.org/10.3390/buildings13102614

[10] H. Hoppe. Uma teoria do socialismo e do capitalismo: economia, política e ética. São Paulo: Instituto Ludwig von Mises Brasil, 2013.

[11] E. Demerouti, A. B. Bakker, F. Nachreiner, W. B. Schaufeli. The job demands-resources model of burnout. Journal of Applied Psychology, v. 86, n. 3, p. 499–512, 2001.

[12] R. Saha, U. Jangbahadur, R. Israney, M. Sinha, S. Ahlawat, U. Akram, X. Su. The psychological and emotional factors of gamification and customer experience. A systematic literature review and future research agenda, v. 260, 2025.

[13] R. M. C. Leite, L. R. G. Alves. Estratégias de gamificação: um sistema web para disseminar as informações de produção para os trabalhadores da construção civil. Pesquisa, Sociedade e Desenvolvimento , [S. l.] , v. 10, n. 12, p. e241101220293, 2021.

[14] A. C. Gil. Métodos e técnicas de pesquisa social. 6. ed. São Paulo: Atlas, 2008.

[15] A. C. C. Trovão, M. I. Q. Netta, I. R. Uhlmann. Revisão sobre o uso de jogos educativos na Engenharia de Produção. XLV Encontro Nacional de Engenharia de Produção (ENEGEP), Natal, 2025.

AUTORES



Ana Beatriz Faria Gomes possui técnico em Administração pelo Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IFSULDEMINAS). Atualmente, é Aluna do curso de graduação em Engenharia de Produção no Instituto Nacional de Telecomunicações (INATEL).



Vinícius A. Montgomery de Miranda é graduado em Engenharia Elétrica pela Universidade Federal de Engenharia de Itajubá (UNIFEI, 1990), mestre em Engenharia de Produção pela Universidade Federal de Itajubá (UNIFEI, 1998). Doutor em Engenharia de Produção pela Universidade Federal de Itajubá (UNIFEI, 2026). Atualmente, é professor no INATEL – Instituto Nacional de Telecomunicações, atuando em Gestão Estratégica, Economia e Finanças. É professor do INATEL desde 2012.

Análise e Gerenciamento de Processos Industriais

Antônio Borges Olivera & Vinicius Antonio Montgomery de Miranda
Instituto Nacional de Telecomunicações – Inatel
antonio.borges@gep.inatel.br, vmontgomery@inatel.br

Abstract—*This study analyzes the relevance of installed capacity and production cycle times within the thermoplastic injection molding industry. Through an applied quantitative case study, production cycles were measured and integrated into a customized management dashboard. The results identified an average capacity utilization rate of 83.7%, revealing a 16.3% gap attributed to operational losses and setups. The implementation of a data-driven tool significantly improved decision-making speed for large-scale orders and reduced scrap by promoting process stability. The study concludes that transitioning from subjective estimates to systematic data analysis enhances planning accuracy, predictability, and organizational competitiveness.*

Index Terms—*Capacity planning, Data-driven decision making, Plastic injection molding, Production management.*

Resumo—Este estudo analisa a relevância da capacidade instalada e da dinâmica dos tempos de produção no setor de injeção termoplástica. Por meio de uma pesquisa aplicada com abordagem quantitativa e estudo de caso, os ciclos de produção foram mensurados e integrados a um *dashboard* gerencial customizado. Os resultados identificaram uma taxa média de utilização da capacidade de 83,7%, revelando um *gap* de 16,3% decorrente de perdas operacionais e *setups*. A implementação da ferramenta baseada em dados aumentou significativamente a agilidade na tomada de decisão para pedidos de larga escala e reduziu o índice de refugo (*scrap*) ao promover a estabilidade do processo. O estudo conclui que a transição de estimativas subjetivas para a análise sistemática de dados amplia a precisão do planejamento, a previsibilidade e a competitividade organizacional.

Palavras chave—Capacidade instalada, Gestão da produção, Decisão fundamentada em dados, Injeção de termoplásticos.

I - INTRODUÇÃO

A crescente intensificação da concorrência, aliada ao aumento do nível de exigência dos consumidores e à rápida incorporação de inovações tecnológicas, tem ampliado a pressão por eficiência nas organizações industriais, demandando maior precisão e fundamentação na tomada de decisão. Nesse cenário, a gestão da produção consolida-se como função estratégica, na medida em que a competitividade organizacional depende da capacidade de converter recursos em bens e serviços com qualidade consistente, custos controlados e agilidade de entrega. Como destacam Slack, Chambers e Johnston (2018) [1], o desempenho operacional está intrinsecamente associado ao modo como os processos produtivos são concebidos e administrados, impactando diretamente variáveis como custo, flexibilidade, confiabilidade e nível de serviço oferecido ao mercado.

A produtividade representa fator decisivo para a competitividade empresarial e para a geração de riqueza, tanto no âmbito organizacional quanto no plano macroeconômico. Ao traduzir a relação entre resultados obtidos e recursos empregados, ela evidencia a eficiência com que o sistema produtivo cria valor econômico. Conforme destacam Corrêa e Corrêa (2017) [2], ganhos de produtividade reduzem custos unitários, ampliam margens e fortalecem a posição competitiva das empresas, seja pela oferta de preços mais competitivos, seja pela capacidade de reinvestimento e inovação. Para Hoppe (2013) [3], a expansão da produtividade está associada à divisão do trabalho, à acumulação de capital e ao aprimoramento do conhecimento aplicado à produção, fatores que ampliam a capacidade de geração de bens e serviços e, consequentemente, de riqueza social. Nesse sentido, produtividade e competitividade configuram-se como dimensões interdependentes do processo de desenvolvimento econômico sustentável.

A capacidade instalada pode ser definida como o volume máximo de produção que um sistema é capaz de alcançar em determinado período, considerando suas restrições técnicas e organizacionais. De acordo com Martins e Laugeni (2015) [4], a adequada mensuração da capacidade produtiva é essencial para o planejamento agregado, o balanceamento de linhas e a tomada de decisões sobre investimentos e expansão. O nível de utilização da capacidade, por sua vez, permite avaliar se a empresa está operando de forma eficiente ou se há desperdícios decorrentes de subutilização ou gargalos que limitam o desempenho global do sistema.

Assim, a análise da capacidade instalada ultrapassa o âmbito operacional e assume relevância estratégica, pois está diretamente vinculada à geração de resultados econômicos. Uma capacidade mal dimensionada pode resultar em perda de vendas, aumento de custos fixos unitários ou deterioração da qualidade e da confiabilidade. Por outro lado, o alinhamento entre capacidade produtiva e demanda de mercado favorece a obtenção de lucro sustentável, ao combinar eficiência operacional com atendimento adequado ao cliente.

Diante desse contexto, o presente trabalho tem por objetivo analisar a relevância da definição da capacidade instalada na dinâmica dos tempos de produção, examinando de que forma sua mensuração acurada e sua gestão sistemática podem contribuir para a melhoria contínua dos processos, a mitigação de perdas operacionais e a ampliação da rentabilidade organizacional. Parte-se do pressuposto de que a adequada compreensão da capacidade produtiva constitui elemento estruturante para decisões relacionadas ao planejamento, programação e controle da produção.

Para atingir tal finalidade, desenvolve-se uma pesquisa aplicada em ambiente real de manufatura, em uma empresa do setor de transformação de plásticos. A investigação busca identificar oportunidades de aperfeiçoamento por meio da análise estruturada da capacidade produtiva disponível,

contemplando variáveis operacionais, índices de utilização e possíveis restrições do sistema produtivo. Espera-se, com isso, gerar subsídios técnicos que apoiem a tomada de decisão gerencial e promovam maior alinhamento entre demanda, recursos produtivos e desempenho operacional.

A. Delimitação da Pesquisa

A pesquisa caracteriza-se como aplicada, de abordagem quantitativa, com objetivo descritivo e analítico, desenvolvida por meio de um estudo de caso único em uma empresa do setor de transformação de plásticos que utiliza o processo de injeção termoplástica.

Segundo Yin (2015) [5], o estudo de caso é adequado quando se busca investigar fenômenos contemporâneos inseridos em seu contexto real, especialmente quando as fronteiras entre fenômeno e contexto não estão claramente definidas. Nesse sentido, a análise da capacidade instalada em ambiente fabril real justifica a escolha metodológica.

B. Organização do Trabalho

O trabalho está estruturado de forma a conduzir o leitor da fundamentação conceitual à aplicação prática dos resultados. Inicialmente, a Introdução contextualiza a relevância da gestão da capacidade produtiva no cenário competitivo atual e apresenta os objetivos da pesquisa. Em seguida, a Delimitação da Pesquisa explicita o enquadramento metodológico do estudo de caso. A Revisão de Literatura consolida os principais referenciais teóricos sobre capacidade instalada, taxa de utilização e produtividade, estabelecendo a base conceitual do estudo. Na sequência, a seção de Procedimentos Metodológicos descreve a coleta e o tratamento dos dados operacionais, bem como o desenvolvimento da planilha e do *dashboard*. Posteriormente, a seção de Análise dos Resultados apresenta a aplicação prática da ferramenta, os indicadores obtidos e os impactos gerenciais identificados. Por fim, as Considerações Finais sintetizam os principais achados, destacam as contribuições do trabalho e evidenciam sua relevância estratégica para a gestão da produção.

II - DESENVOLVIMENTO

A análise da capacidade produtiva se configura como etapa central no planejamento e no controle da produção, especialmente em ambientes industriais caracterizados por elevada intensidade tecnológica e forte pressão por eficiência. No contexto da indústria de transformação e, de modo particular, no setor de injeção de termoplásticos, a adequada mensuração da capacidade instalada e do nível de utilização dos equipamentos é determinante para assegurar competitividade, confiabilidade dos prazos e desempenho econômico sustentável. Conforme Tubino (2017) [6], a capacidade instalada corresponde ao volume máximo de produção que um sistema pode alcançar em determinado período, considerando condições normais de operação, disponibilidade de recursos produtivos e restrições tecnológicas. Trata-se, portanto, não apenas de uma estimativa de potencial produtivo, mas de um parâmetro essencial para a avaliação do desempenho efetivo da organização.

Sob a perspectiva estratégica, a capacidade deve ser compreendida como limite estrutural que orienta decisões relativas a investimentos, ampliação de instalações, aquisição de equipamentos, balanceamento de linhas e definição do mix de produtos. Slack, Chambers e Johnston (2018) [1] destacam que a gestão da capacidade envolve decisões de longo, médio e curto prazo, sendo crucial para alinhar a oferta produtiva às variações da demanda. Uma interpretação inadequada desse indicador pode gerar consequências econômicas relevantes: a ociosidade excessiva implica desperdício de recursos e elevação do custo fixo unitário, enquanto a operação sistemática próxima ao limite máximo tende a aumentar falhas, custos de manutenção e riscos de perda de qualidade.

Complementarmente, a taxa de utilização da capacidade constitui indicador operacional que expressa o grau de aproveitamento do potencial produtivo disponível. De acordo com Gaither e Frazier (2014) [7], a utilização é calculada pela relação entre a produção efetiva e a capacidade disponível, permitindo identificar gargalos, ineficiências e distorções no dimensionamento dos recursos. Em sistemas de injeção plástica, nos quais os ciclos produtivos são repetitivos, dependentes de parâmetros térmicos e temporais rigorosos, e sensíveis a variações de *setup*, a análise integrada da capacidade e de sua utilização assume relevância ainda maior. Pequenas oscilações no tempo de ciclo, na disponibilidade das máquinas ou na programação podem impactar significativamente o volume produzido e os custos operacionais.

Além disso, a literatura em gestão da produção enfatiza que a administração eficaz da capacidade está diretamente relacionada à produtividade e à rentabilidade. Corrêa e Corrêa (2017) [2] ressaltam que decisões adequadas sobre capacidade contribuem para reduzir desperdícios, estabilizar fluxos produtivos e melhorar o nível de serviço ao cliente, fatores que fortalecem a posição competitiva da empresa. Assim, no ambiente industrial contemporâneo, a análise da capacidade produtiva ultrapassa o âmbito meramente técnico, assumindo papel estratégico na busca por eficiência operacional, equilíbrio econômico e sustentabilidade organizacional.

III – PROCEDIMENTOS METODOLÓGICOS

O trabalho desenvolvido no âmbito da iniciação científica buscou mensurar e analisar a capacidade instalada de uma empresa do setor de transformação de plásticos, especializada na fabricação de componentes por meio do processo de injeção. Para alcançar esse propósito, realizou-se a coleta sistemática e padronizada dos tempos de ciclo referentes aos produtos manufaturados nas diferentes injetoras que compõem o parque industrial.

O tempo de ciclo, variável crítica para a determinação da produtividade e para o dimensionamento da capacidade, abrange as etapas de fechamento do molde, injeção do polímero fundido, aplicação de recalque (ou pressão de compactação), resfriamento, abertura do molde e extração da peça. A

mensuração precisa dessas fases é fundamental, uma vez que pequenas variações nos parâmetros operacionais podem impactar diretamente o rendimento produtivo, a estabilidade do processo e a qualidade final do produto. Conforme destacam Rosato e Rosato (2000) [8], variações mesmo que discretas nesses parâmetros podem impactar significativamente o rendimento do processo, afetando tanto o volume produzido quanto os custos operacionais, o que torna sua mensuração precisa indispensável em estudos de capacidade produtiva.

A partir do registro individualizado do ciclo de cada produto em cada equipamento, foi possível estimar o volume máximo de peças produzidas por hora, considerando o número de cavidades do molde, a eficiência operacional, a disponibilidade diária das máquinas e as paradas programadas para *setup* e manutenção. Tais variáveis são reconhecidas na literatura de gestão da produção como determinantes para a apuração da capacidade efetiva, distinguindo-se da capacidade teórica máxima [1]. Os dados coletados foram organizados em planilha estruturada com fórmulas automatizadas, permitindo o cálculo da capacidade instalada diária por máquina e a consolidação da capacidade global da empresa.

Além da mensuração da capacidade total, foi incorporada à ferramenta uma funcionalidade adicional destinada ao apoio à programação da produção. Por meio de uma aba específica, tornou-se possível inserir uma quantidade desejada de produção e obter automaticamente o tempo estimado necessário para sua fabricação, com base no tempo de ciclo e na produtividade real observada. Essa aplicação prática contribui para decisões relacionadas ao sequenciamento de ordens, estimativas de *lead time* e avaliação da viabilidade de atendimento a prazos comerciais, aspectos essenciais para o alinhamento entre operações e mercado [2].

A. Metodologia de Coleta e Tratamento de Dados

O desenvolvimento da ferramenta fundamentou-se em uma metodologia empírica baseada na coleta sistemática de dados em ambiente real de produção, ao longo de múltiplos dias operacionais. Para assegurar rigor técnico e confiabilidade estatística, o ciclo de injeção foi decomposto em etapas operacionais mensuráveis, notadamente: (i) tempo de preenchimento e recalque e (ii) tempo de extração (retirada da peça do molde).

As medições foram realizadas por meio de cronometragem direta, seguidas do cálculo da média aritmética dos ciclos observados, a qual foi adotada como parâmetro de entrada (*input*) para os cálculos automatizados de capacidade na planilha desenvolvida. Esse procedimento metodológico reduz a influência de observações atípicas e contribui para maior robustez nas estimativas de produtividade.

IV – ANÁLISE DOS RESULTADOS

Com os dados consolidados, foram elaborados gráficos e painéis gerenciais que possibilitaram a análise da taxa de utilização das injetoras, evidenciando máquinas operando próximas ao limite de capacidade, níveis de ociosidade e potenciais gargalos do sistema produtivo. A visualização comparativa entre produtos, turnos e equipamentos proporcionou uma compreensão sistêmica do desempenho da fábrica, favorecendo a identificação de desequilíbrios e oportunidades de melhoria.

A abordagem adotada permitiu à empresa compreender com maior precisão sua capacidade produtiva efetiva, reduzindo incertezas no planejamento e ampliando a previsibilidade operacional. Os resultados subsidiaram decisões relativas ao ajuste do mix de produção, priorização de ordens, incremento da eficiência e avaliação de investimentos futuros em ampliação ou modernização do parque fabril. Assim, a aplicação prática dos conceitos de capacidade instalada mostrou seu valor estratégico como instrumento de apoio à competitividade e à sustentabilidade econômica da organização.

A. Planilha e Dashboard

O trabalho resultou no desenvolvimento de uma planilha estruturada e de um *dashboard* (painel de controle) composto por gráficos e indicadores de desempenho voltados ao monitoramento das injetoras do parque fabril. A planilha foi concebida para consolidar e processar os dados operacionais, enquanto o *dashboard* foi estruturado para apresentar as informações de forma visual, sintética e orientada à tomada de decisão.

O painel contempla campos específicos para a delimitação do período de análise, mediante a inserção de data inicial e data final. Tal recurso possibilita a filtragem temporal das informações, por mês, turno ou qualquer intervalo customizado, viabilizando o acompanhamento sistemático do volume produzido no período selecionado, bem como a identificação das injetoras com maior disponibilidade e respectivos níveis de utilização.

Ao permitir a análise dinâmica dos dados operacionais, a ferramenta fortalece o controle gerencial, aprimora o processo de planejamento e programação da produção e amplia a assertividade das decisões operacionais. A Figura 1 apresenta um exemplo ilustrativo do *dashboard* desenvolvido.

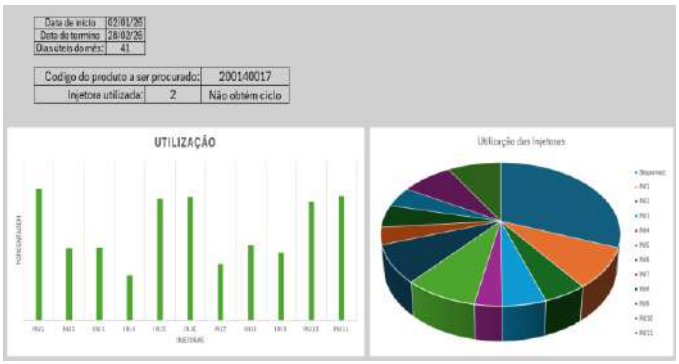


Fig. 1. *Dashboard* desenvolvido no âmbito do projeto.

B. Consulta de Ciclos e Capacidade

Na etapa subsequente, ao selecionar o código de determinado produto, o sistema executa automaticamente uma consulta à base de dados corporativa com o objetivo de verificar a existência de tempos de ciclo previamente registrados para o item em questão.

Com base nessa verificação, o sistema retorna o status “Obtém ciclo” ou “Não obtém ciclo”, sinalizando, respectivamente, a disponibilidade de parâmetros operacionais cadastrados e, consequentemente, a viabilidade técnica de produção na máquina selecionada. Esse procedimento assegura maior confiabilidade às análises de capacidade e reduz o risco de programação de itens sem dados técnicos validados.

As figuras 2 e 3 apresentam o armazenamento de dados utilizado para realizar as pesquisas na planilha.

Codigo do produto	Cliente	Produto
200020019	APTIV	SUPORTE (BOCAL INJETADO)
200020025	APTIV	CLIP MEDIO INJETADO
200060026	APTIV	BOCAL INJETADO TRASEIRO
200020033	APTIV	BASE DROPSLID
200020018	APTIV	RACK P/ CIRCUITOS
200020020	APTIV	DIVISOR DE CIRCUITOS
200020026	APTIV	CHIP CURTO
200060013	APTIV	CHIP P/ REGUA T
200060020	APTIV	HOLDER HOB
200060024	APTIV	CLIP ABRACADEIRA P/ TUBO
200020017	APTIV	BOCAL INJETADO PERCHEIRO
200060027	APTIV	BOCAL DIANTEIRO
200060023	APTIV	ALAVANCA PEBH 3498
200050005	APTIV	BORRACHA PROT. MINI APLICADOR
200050006	APTIV	BORRACHA PROT. MINI APLICADOR
200060008	APTIV	FIXADOR INJETADO PEBH 1322
200020065	APTIV	BOCAL DIANTEIRO INJETADO
200020068	APTIV	BOCAL DIANTEIRO INJETADO
200140017	BALTEAU	CAIXA PROTEÇÃO SECUNDARIA TIPO 2
200140021	BALTEAU	CX PROT. SECUNDARIA P/MT
200140001	BALTEAU	PERFIL VISOR DEMO L 157 MM
200140005	BALTEAU	ESPAÇADOR POLIAMIDA 150 MM
200140010	BALTEAU	PERFIL EXTERNO PARA NUCLEO
200140011	BALTEAU	INVOLUCRO NYLON BLA - 0,6
200140003	BALTEAU	CAIXA DE PROTEÇÃO SECUNDARIA TIPO 3
200140015	BALTEAU	TAMPA PLASTICA SECUNDARIA TCBT
200140002	BALTEAU	PERFIL INTERNO P/ NUCLEO
204760001	AUDAX	CAIXA PLASTICA

Fig. 2. Tabela de armazenamento de dados.

Injetora 1	Injetora 2	Injetora 3	Injetora 4	Injetora 5	Injetora 6	Injetora 7	Injetora 8	Injetora 9	Injetora 10	Injetora 11
38.0								37.0	42.0	
30.0				24.0			17.0			
30.0			37.0				25.0			
20.0										
25.0										
	23.0	25.0	30.0	25.0	20.0		23.0			
	20.0	22.5								
26.0	27.0	25.0		25.0			23.0			
	30.0	33.0	30.0				28.0			
30.0	30.0				30.0					
							38.0			
		35.0							42.0	
			30.0		30.0		25.0			
					30.0					
					30.0					
					18.0					
									42.0	
									42.0	
18.0				18.0						
33.5										
	35.0	25.0	40.0		38.0					
		21.0								
		25.0				25.0				
		30.0				40.0				
		25.0								
					25.0					
						25.0				
30.0		30.0				30.0		30.0		

Fig. 3 Continuação da tabela de dados.

Após a identificação da injetora compatível e do respectivo tempo de ciclo, o código do produto é automaticamente direcionado a outra seção da planilha destinada ao processamento dos dados de capacidade.

Nessa etapa, o sistema realiza o cálculo automatizado da capacidade instalada diária, considerando a carga horária efetiva de produção da empresa, bem como os parâmetros operacionais previamente cadastrados. Esse procedimento assegura padronização nos cálculos e maior confiabilidade às estimativas de capacidade produtiva, conforme apresentado na figura 4.

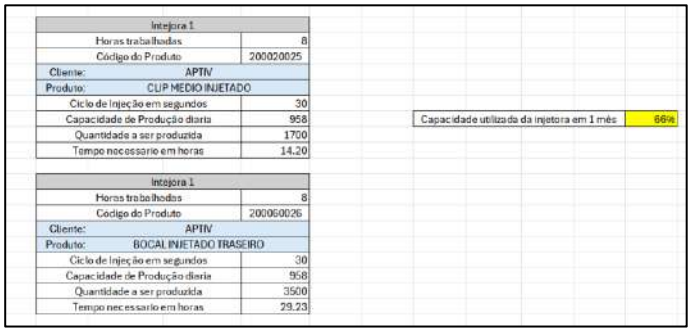


Fig. 4. Dimensionamento da Capacidade Produtiva e Estimativa de Tempo de Produção.

C. Análise de Demanda e Ocupação Mensal

Além de mensurar a capacidade instalada, a ferramenta possibilita estimar o tempo necessário (em horas) para o atendimento de uma demanda específica. A partir da quantidade solicitada e do tempo de ciclo validado, o sistema calcula automaticamente a carga-máquina requerida. Adicionalmente, é apresentado o percentual de ocupação da capacidade mensal da injetora para atender ao volume demandado, permitindo avaliar o impacto do pedido sobre o plano produtivo vigente. Esse recurso fornece suporte analítico ao planejamento agregado e à programação de médio e longo prazo, reduzindo incertezas quanto à disponibilidade de recursos.

D. Estabilidade do Processo e Variabilidade

A confiabilidade dos dados coletados está diretamente associada ao regime operacional da unidade produtiva, que atua em turno único. Essa característica contribuiu para a manutenção de baixo desvio padrão nos ciclos medidos, uma vez que minimizou fontes recorrentes de variabilidade presentes em operações multi-turno, tais como alternância de operadores, diferenças de ritmo produtivo e variações térmicas ambientais acentuadas.

Consequentemente, os parâmetros consolidados na ferramenta refletem um processo com elevada previsibilidade e estabilidade estatística, condição essencial para análises consistentes de capacidade produtiva.

E. Análise da Capacidade Real e Utilização

A aplicação da ferramenta evidenciou que as injetoras operam com um índice médio de utilização de 83,7% em relação ao tempo nominal disponível. Esse resultado revela a existência de um *gap* de 16,3% entre a capacidade teórica e a capacidade efetivamente realizada, decorrente de perdas inerentes ao ambiente fabril, como ajustes de máquina, pequenas paradas, trocas de molde e intervenções operacionais.

A incorporação desse índice aos cálculos da planilha permite que o planejamento da produção considere a capacidade real do sistema, promovendo agendamentos mais aderentes à realidade operacional. Dessa forma, torna-se possível otimizar a carga-máquina, reduzir riscos de sobrecarga e mitigar períodos de ociosidade não planejada.

F. Relação entre Planejamento Assertivo e Qualidade (Redução de Scrap)

O planejamento estruturado da carga-máquina influencia diretamente a estabilidade do processo produtivo. Em cenários caracterizados por pressão excessiva e reprogramações frequentes, é comum a ocorrência de instabilidades térmicas e operacionais nas injetoras, ampliando a variabilidade do processo e afastando-o de sua janela ótima de operação.

A utilização da ferramenta contribui para evitar gargalos inesperados e setups executados de forma precipitada, assegurando o respeito aos parâmetros técnicos de injeção previamente definidos. Como consequência, observa-se redução do índice de refugo (*scrap*), uma vez que a manutenção de ciclos estáveis e a ausência de sobrecarga operacional diminuem a incidência de defeitos dimensionais, rebarbas e falhas de preenchimento.

G. Análise de Ganhos: A Transição para a Gestão baseada em Dados

A implementação da ferramenta representou uma mudança estrutural na gestão da unidade produtiva, ao substituir práticas baseadas em estimativas subjetivas por decisões fundamentadas em dados sistematizados. Anteriormente, a ausência de registros consolidados sobre tempos de ciclo reais e disponibilidade das máquinas limitava a precisão no cálculo da capacidade instalada e comprometia a confiabilidade do planejamento.

Enquanto gestões que operam sem embasamento técnico tendem a estabelecer metas sobredimensionadas que geram desmotivação na equipe, a ferramenta desenvolvida neste estudo, ao considerar o índice de utilização real de 83,7% e dados sistemáticos, evita falhas de planejamento como as observadas por Silva, Souza Neto e Diniz (2025) [9]. No estudo referido, a meta real de 13.218 peças/mês era ignorada em favor de uma estimativa subjetiva de 18.000 peças, resultando em metas inexequíveis.

Com a estruturação das informações na planilha e sua integração ao *dashboard*, houve ganho expressivo de agilidade na resposta às demandas comerciais. A viabilidade de pedidos de grande escala por exemplo, a produção de 10.000 peças que anteriormente exigiria dias de análise ou seria respondida sem embasamento técnico, passou a ser avaliada em poucas horas, com base em parâmetros operacionais consolidados. Esse aumento de velocidade decisória confere maior robustez ao planejamento, amplia a confiabilidade junto aos clientes e fortalece a competitividade da empresa em um ambiente industrial orientado por desempenho e eficiência.

V – CONSIDERAÇÕES FINAIS

O presente trabalho evidenciou que a mensuração estruturada da capacidade instalada constitui elemento central para o aprimoramento do planejamento e controle da produção em ambientes industriais, especialmente no setor de transformação de plásticos por injeção. Ao sistematizar a coleta dos tempos de ciclo, tratar estatisticamente os dados e integrar essas informações a uma planilha automatizada e a um *dashboard* gerencial, foi possível transformar dados operacionais dispersos em informação estratégica para a tomada de decisão.

Os resultados evidenciaram que a capacidade nominal das injetoras difere da capacidade efetivamente realizada, sendo identificado um índice médio de utilização de 83,7%. Esse achado reforça a importância de considerar perdas operacionais inerentes ao processo produtivo nos cálculos de planejamento, evitando tanto a superestimação da capacidade quanto a geração de promessas de entrega inexequíveis. A incorporação do índice real de utilização aos modelos de cálculo permitiu maior aderência entre planejamento e execução, contribuindo para a redução de riscos operacionais.

Adicionalmente, verificou-se que o planejamento fundamentado em dados confiáveis impacta positivamente a estabilidade do processo e a qualidade do produto. Ao minimizar reprogramações emergenciais e sobrecargas nas injetoras, a ferramenta favorece a manutenção de parâmetros operacionais consistentes, reduzindo a variabilidade e, consequentemente, os índices de refugo (*scrap*). Dessa forma, a gestão da capacidade revela-se não apenas um instrumento quantitativo de dimensionamento produtivo, mas também um vetor de melhoria contínua e eficiência operacional.

Por fim, conclui-se que a implementação da ferramenta representou uma transição significativa para um modelo de gestão orientado por dados (*data-driven*), proporcionando maior previsibilidade, agilidade na análise de viabilidade de pedidos e suporte técnico às decisões estratégicas relacionadas à alocação de recursos e investimentos futuros. O estudo evidencia que a análise rigorosa da capacidade instalada, quando integrada a instrumentos de visualização e controle, constitui vantagem competitiva relevante em ambientes industriais caracterizados por elevada pressão por eficiência operacional, aumento de produtividade e redução de custos.

VI. AGRADECIMENTOS

Os autores agradecem à Fundação Instituto Nacional de Telecomunicações (FINATEL) pelo apoio institucional e incentivo às atividades de pesquisa que tornaram possível o desenvolvimento deste trabalho.

REFERENCIAS

[1] N. Slack, S. Chambers, R. Johnston. Administração da Produção. 5. ed. São Paulo: Atlas, 2018.

[2] H. L. Corrêa, I. G. Corrêa. Administração de produção e operações: manufatura e serviços: uma abordagem estratégica. 4. ed. São Paulo: Atlas, 2017.

[3] H. Hoppe. O que deve ser feito. São Paulo: Instituto Ludwig von Mises Brasil, 2013.

[4] P. G. Martins, F. P. Laugeni. Administração da produção. 3. ed. rev. e atual. São Paulo: Saraiva, 2015.

[5] R. K. Yin. Estudo de caso: planejamento e métodos. 5. ed. Porto Alegre: Bookman, 2015.

[6] D. R. Tubino. Planejamento e controle da produção: teoria e prática. 3. ed. São Paulo: Atlas, 2017.

[7] N. Gaither, G. Frazier. Administração da produção e operações. 8. ed. São Paulo: Cengage Learning, 2014.

[8] D. V. Rosato, D. V. Rosato. Injection molding handbook. 3. ed. Boston: Kluwer Academic Publishers, 2000.

[9] M. E. L. Silva, A. M. Souza Neto, H. H. L. Diniz. Análise de capacidade de produção em uma fábrica de confecção de roupas de bebê no setor de costura. In: XLV ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 2025, Natal. Anais... Natal: ENEGEP, 2025.

AUTORES



Antonio Borges Oliveira é graduando em Engenharia de Produção pelo Instituto Nacional de Telecomunicações (Inatel). Atualmente, atua como estagiário de Engenharia de Qualidade em uma empresa do setor automotivo especializada na fabricação de espelhos retrovisores. Desenvolve projeto de iniciação científica voltado para a análise e gestão de processos industriais, com foco no aumento da produtividade e na eficiência operacional. Seus principais interesses de pesquisa e atuação profissional concentram-se em análise estatística, métodos de cálculo numérico e na aplicação de ferramentas de programação para a análise de dados e suporte à tomada de decisão na gestão da produção.



Vinícius A. Montgomery de Miranda é graduado em Engenharia Elétrica pela Universidade Federal de Itajubá (UNIFEI, 1990), mestre em Engenharia de Produção pela Universidade Federal de Itajubá (UNIFEI, 1998) e Doutor em Engenharia de Produção pela Universidade Federal de Itajubá (UNIFEI, 2026). Atualmente, é professor no INATEL – Instituto Nacional de Telecomunicações, atuando em Gestão Estratégica, Economia e Finanças. É professor do INATEL desde 2012.

Qualidade de Energia Elétrica: Simulador de Distúrbios e Mitigação

Egidio Raimundo Neto, Geovane Mantuam Alvim, Júlia da Silva Maciel e Natan Barreto de Paula
Instituto Nacional de Telecomunicações – Inatel

***Abstract*—This work presents an interactive Power Quality simulator developed in Python using the NumPy, Matplotlib, and Streamlit libraries. The tool enables the generation and analysis of single-phase and three-phase electrical signals, allowing controlled harmonic insertion, simulation of typical power disturbances such as voltage sags, swells, and interruptions, and calculation of Total Harmonic Distortion using the Fast Fourier Transform. Nonlinear loads are also modeled, including single-phase full-wave rectifiers and induction motors approximated by an RL model, highlighting the effects of voltage harmonics on electrical current. The interactive graphical interface facilitates parameter configuration and both time-domain and frequency-domain analysis of signals, making the simulator suitable for educational and academic applications in power systems engineering.**

***Index Terms*—Computational Simulation, Harmonics, Power Quality, Power Systems, Total Harmonic Distortion.**

***Resumo*—Este trabalho apresenta um simulador interativo de Qualidade de Energia Elétrica desenvolvido em Python com uso das bibliotecas NumPy, Matplotlib e Streamlit. A ferramenta permite a geração de sinais monofásicos e trifásicos, a inserção de harmônicas, a simulação de distúrbios elétricos como afundamentos *(sag)*, elevações *(swell)* e interrupções, além do cálculo da Distorção Harmônica Total por meio da Transformada Rápida de Fourier. Também são modeladas cargas não lineares, como retificadores monofásicos de onda completa e motores de indução aproximados por um modelo RL, evidenciando os impactos das harmônicas na corrente elétrica. A interface gráfica interativa facilita a configuração dos parâmetros e a análise temporal e espectral dos sinais, tornando o simulador adequado para aplicações didáticas e acadêmicas em sistemas elétricos de potência.**

***Palavras chave*—Distorção Harmônica Total, Harmônicas, Qualidade de Energia Elétrica, Simulação Computacional, Sistemas Elétricos de Potência.**

I. Introdução

A elevada presença de cargas não lineares e conversores eletrônicos de potência nos sistemas elétricos contemporâneos consolidou-se como um dos principais desafios para a Qualidade de Energia Elétrica (QEE). Embora o ritmo de industrialização apresente relativa estabilidade em termos macroeconômicos na última década, observa-se uma contínua substituição tecnológica [1], na qual dispositivos semicondutores e sistemas chaveados passaram a compor a base de equipamentos como inversores de frequência e fontes de alimentação. Esses componentes introduzem distorções harmônicas e distúrbios de tensão que podem comprometer a eficiência, a confiabilidade do sistema e a vida útil dos equipamentos [2]. Nesse contexto, a análise e o controle desses fenômenos tornaram-se temas centrais na engenharia elétrica, sendo regulamentados por normas técnicas, como a IEEE 519 [3].

Ferramentas de simulação computacional têm papel fundamental na análise desses fenômenos, pois permitem a reprodução controlada de distúrbios elétricos e a avaliação de seus efeitos sobre o sistema [4, 5]. Além disso, simuladores interativos possibilitam a visualização dos impactos das harmônicas, dos afundamentos e elevações de tensão e das interrupções, contribuindo para uma melhor compreensão desses fenômenos.

Diante desse cenário, este trabalho tem como objetivo desenvolver um simulador interativo de Qualidade de Energia Elétrica para análise de distúrbios em sistemas elétricos. O simulador foi implementado em Python e integra análise no domínio do tempo e da frequência, permitindo a modelagem de harmônicas e a simulação de cargas não lineares. Diferentemente de simuladores comerciais de código fechado e alta complexidade, a principal contribuição deste artigo é de caráter misto (didático e experimental): fornece uma interface acessível que traduz modelos matemáticos de QEE em visualizações gráficas imediatas, além de propor uma validação direta entre os sinais simulados e medições físicas de laboratório. Assim, a ferramenta busca reduzir a distância entre a teoria acadêmica e a prática laboratorial na área de sistemas elétricos de potência.

II. Metodologia

A metodologia adotada neste trabalho foi estruturada em duas etapas principais: o desenvolvimento do simulador computacional e a validação experimental em laboratório. Na primeira etapa, foram implementados os modelos e rotinas de simulação responsáveis pela geração e análise dos sinais elétricos, permitindo reproduzir fenômenos associados à Qualidade de Energia Elétrica [5, 1]. Na segunda etapa, foram realizados ensaios experimentais com o objetivo de comparar os resultados obtidos por simulação com medições realizadas em sistemas reais. A validação foi conduzida por meio da análise comparativa entre as formas de onda no domínio do tempo e os espectros no domínio da frequência, bem como pela comparação dos valores de Distorção Harmônica Total (THD) obtidos experimentalmente e por simulação. Essa abordagem permitiu avaliar a consistência dos resultados e a capacidade do modelo em representar fenômenos reais de qualidade de energia elétrica.

A. Simulação Computacional

A primeira etapa consistiu no desenvolvimento de um simulador de Qualidade de Energia Elétrica em Python. A geração e manipulação dos sinais, bem como a implementação dos modelos matemáticos, foram realizadas com a biblioteca *NumPy*,

incluindo a síntese de harmônicas e a resolução de equações diferenciais associadas a cargas RL. A análise espectral foi conduzida por meio da Transformada Rápida de Fourier (FFT), permitindo a obtenção do espectro harmônico e o cálculo da Distorção Harmônica Total (THD). A visualização dos sinais no domínio do tempo e da frequência foi realizada com a biblioteca Matplotlib. Adicionalmente, foi desenvolvida uma interface interativa utilizando o *framework Streamlit*, possibilitando a configuração de parâmetros e a simulação de diferentes eventos de qualidade de energia. Essa abordagem assegura flexibilidade e clareza na análise dos fenômenos estudados [4].

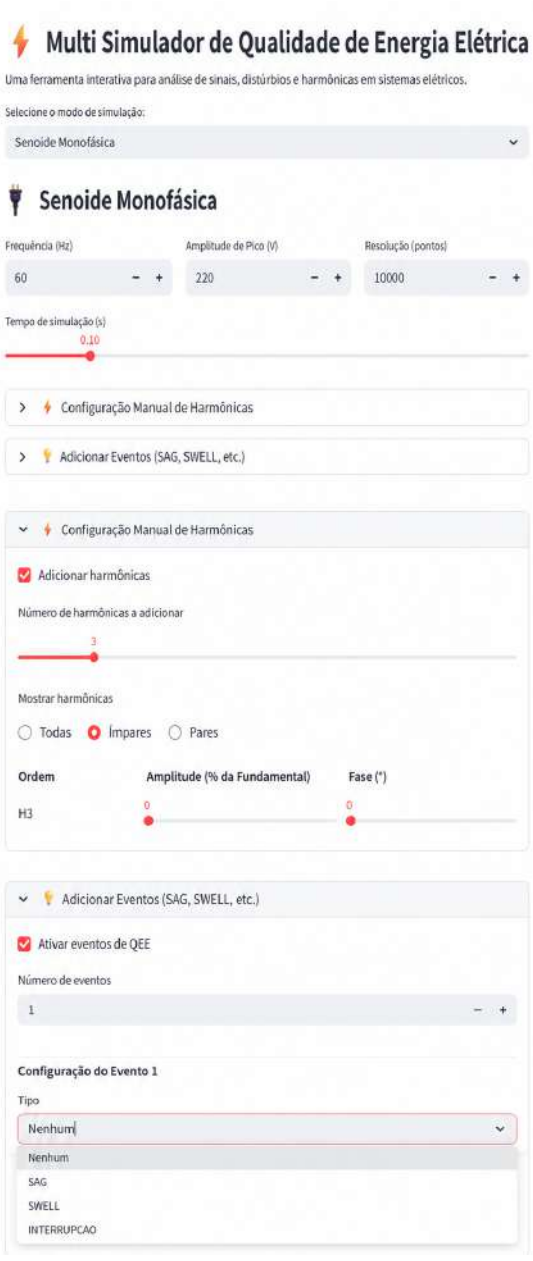


Fig. 1. Interface do simulador de Qualidade de Energia Elétrica.

A Figura 1 apresenta a interface principal do simulador. As demais formas de onda e resultados obtidos são discutidos nas seções seguintes, acompanhados de suas respectivas análises para diferentes condições de operação.

B. Validação Experimental em Laboratório

A segunda etapa da metodologia envolveu a realização de experimentos em laboratório com o objetivo de validar os resultados obtidos na simulação computacional. Foram conduzidos ensaios utilizando retificadores monofásicos, motores de indução trifásicos e diferentes configurações de ligação em estrela e triângulo, permitindo analisar o comportamento do sistema sob distintas condições de carregamento. A instrumentação utilizada incluiu osciloscópios digitais para a aquisição e análise das formas de onda de tensão no domínio do tempo, além de medidores de qualidade de energia empregados para a obtenção de parâmetros elétricos, como Distorção Harmônica Total (THD), espectro harmônico individual, fator de potência e desequilíbrio de tensão.

A Figura 2 apresenta o arranjo experimental utilizado, no qual a rede elétrica alimenta o sistema por meio de um inversor de frequência conectado ao motor de indução, enquanto o medidor Mult-K e o osciloscópio digital são responsáveis pela aquisição dos sinais e parâmetros elétricos ao longo do circuito.

Durante os ensaios, também foi realizado um teste prático utilizando um inversor de frequência associado a um medidor de grandezas elétricas, Mult-K, permitindo a observação das formas de onda e dos parâmetros elétricos do sistema em condições reais de operação, conforme ilustrado na Figura 3. As medições possibilitaram observar experimentalmente a influência das cargas não lineares na rede elétrica, evidenciando a presença de componentes harmônicas e alterações no perfil das correntes de fase e de neutro. Os resultados experimentais foram posteriormente comparados com os dados obtidos pelo simulador, possibilitando avaliar a consistência do modelo implementado e sua capacidade de representar fenômenos reais associados à qualidade de energia elétrica.

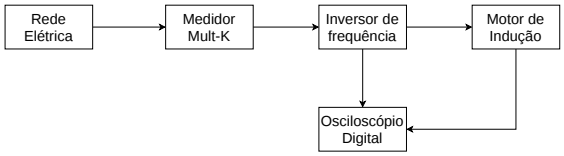


Fig. 2. Diagrama em blocos do cenário experimental, ilustrando a interligação entre a rede elétrica, o inversor de frequência, o motor de indução e os instrumentos de medição.

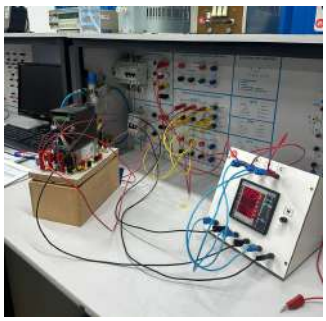


Fig. 3. Teste prático com inversor de frequência e medidor Mult-K.

III. Funcionamento

O simulador de Qualidade de Energia Elétrica é organizado em diferentes modos de operação que representam condições típicas de sistemas elétricos e permitem analisar diversos

fenômenos associados à qualidade da energia. Os resultados gerados pelo simulador, incluindo as formas de onda no domínio do tempo e suas representações espectrais, são apresentados nas figuras das subseções a seguir, permitindo a visualização direta das saídas para cada modo de operação. No modo de senoide monofásica, é gerado um sinal de tensão alternada ideal, no qual podem ser inseridas harmônicas de diferentes ordens e aplicados distúrbios como *sag*, *swell* e interrupções. O *sag* corresponde a uma redução temporária da amplitude da tensão, enquanto o *swell* representa uma elevação temporária dessa amplitude. Essa configuração permite observar a alteração da forma de onda e realizar análises tanto no domínio do tempo quanto no domínio da frequência [2]. O modo de senoide trifásica estende essa abordagem para sistemas com três fases defasadas de 120 graus elétricos, possibilitando a avaliação de efeitos harmônicos em sistemas equilibrados e desequilibrados, bem como a análise da interação entre as fases [6].

O modo de retificador monofásico de onda completa modela uma carga não linear a partir da retificação da tensão de entrada, resultando em sinais com elevado conteúdo harmônico e frequência fundamental duplicada, característica comum em equipamentos eletrônicos de potência [7]. Por fim, o modo de motor de indução, aproximado por um modelo elétrico RL, simula a resposta da corrente elétrica à tensão aplicada, permitindo avaliar como distorções presentes na tensão podem influenciar o comportamento da corrente em sistemas elétricos [8].

A. *Senoide Monofásica*

No modo de senoide monofásica, o código gera um sinal de tensão alternada ideal a partir de parâmetros definidos pelo usuário, como frequência, amplitude e fase. Esse sinal pode ser descrito por:

$$v(t) = A \sin(2\pi f t + \varphi) \tag{1}$$

onde A é a amplitude, f a frequência e φ a fase inicial. Sobre esse sinal, podem ser adicionadas harmônicas de ordens específicas, com amplitudes e defasagens configuráveis, além da aplicação de eventos de qualidade de energia, como afundamentos, elevações e interrupções de tensão, conforme descrito na literatura de qualidade de energia elétrica. O sinal resultante é analisado no domínio do tempo e da frequência, permitindo o cálculo da Distorção Harmônica Total.

A Figura 4 apresenta o sinal senoidal monofásico ideal gerado pelo simulador, utilizado como referência para a análise dos demais fenômenos. Já a Figura 5 ilustra a ocorrência de distúrbios de tensão, como o afundamento de tensão *(SAG)* e a elevação de tensão *(SWELL)*, evidenciando a alteração da forma de onda resultante.

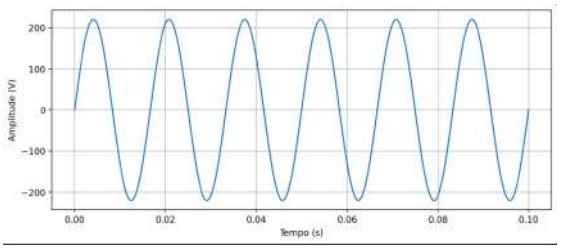


Fig. 4. Senoide monofásica sem alteração.

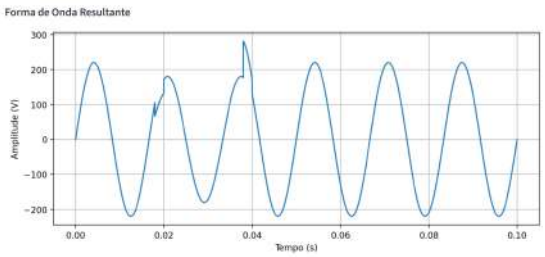


Fig. 5. Senoide monofásica com ocorrência de SAG e SWELL.

B. *Senoide Trifásica*

O modo de senoide trifásica estende essa abordagem para sistemas de três fases defasadas de 120 graus elétricos. Cada fase é gerada de forma independente, possibilitando a aplicação individual de harmônicas e distúrbios. Essa estrutura permite observar fenômenos típicos de sistemas trifásicos equilibrados e desequilibrados, bem como os efeitos de cancelamento ou amplificação de determinadas harmônicas [6].

A Figura 6 apresenta o sinal trifásico ideal gerado pelo simulador, no qual as três fases estão defasadas de 120° e mantêm comportamento senoidal equilibrado. Já a Figura 7 ilustra a aplicação de distúrbios de tensão, *SAG* e *SWELL*, evidenciando as alterações na forma de onda das fases.

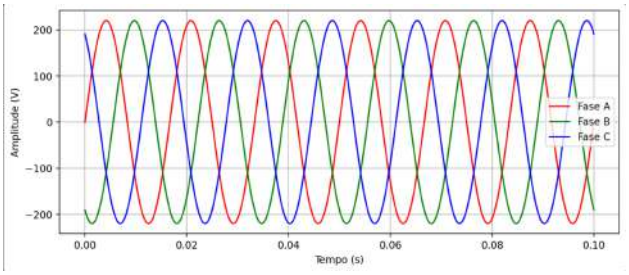


Fig. 6. Senoide trifásica sem alteração.

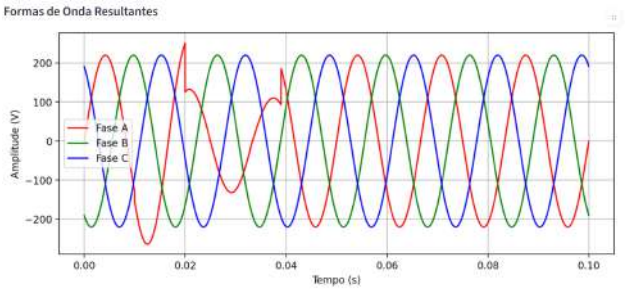


Fig. 7. Senoide trifásica com ocorrência de SAG e SWELL.

C. *Retificador Monofásico*

No modo de retificador monofásico de onda completa, o simulador modela uma carga não linear a partir da retificação da tensão senoidal de entrada. Esse processo resulta em uma forma de onda com frequência fundamental duplicada e elevado conteúdo harmônico, característica comum em equipamentos eletrônicos de potência [7]. O sinal retificado é analisado espectralmente, evidenciando o impacto da não linearidade sobre o índice de distorção harmônica.

A Figura 8 apresenta o resultado da simulação do retificador monofásico alimentado em 127 V, no qual é possível observar

a forma de onda retificada e a presença de componentes harmônicas decorrentes do processo de retificação.

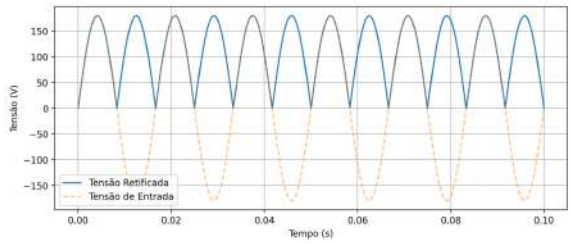


Fig. 8. Retificador monofásico alimentado em 127 V.

D. Motor de Indução

O modo de motor de indução, aproximado por um modelo elétrico RL [8], simula o comportamento da corrente elétrica a partir da aplicação de uma tensão senoidal ou distorcida. A equação diferencial que relaciona tensão, resistência e indutância é resolvida numericamente, permitindo avaliar como distorções harmônicas na tensão se refletem em correntes ainda mais distorcidas. Esse modo pode ser aplicado tanto em configurações monofásicas quanto trifásicas, ampliando a análise para sistemas industriais.

A Figura 9 apresenta a tensão de alimentação aplicada ao motor no processo de simulação. Já a Figura 10 ilustra a corrente resultante no motor, evidenciando como as características do circuito RL e as possíveis distorções presentes na tensão influenciam a forma de onda da corrente elétrica.

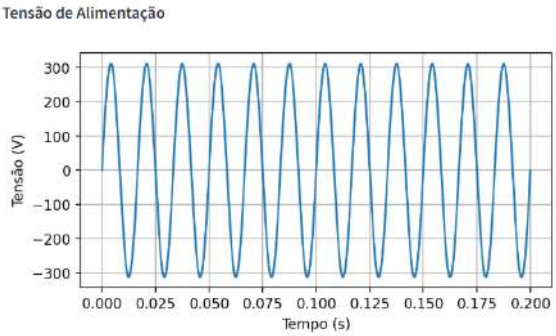


Fig. 9. Tensão de alimentação do motor.

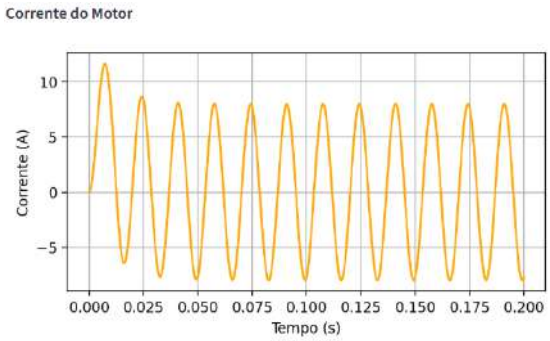


Fig. 10. Corrente do motor.

E. Harmônicas

Em sistemas monofásicos, as distorções harmônicas na tensão e na corrente propagam-se integralmente, pois não há mecanismos de cancelamento entre fases, resultando em elevação da Distorção Harmônica Total e impactos diretos no desempenho dos equipamentos [9, 10]. Em sistemas trifásicos, o comportamento das harmônicas depende do equilíbrio do sistema e da sequência harmônica considerada. Componentes de sequência positiva e negativa afetam a operação dos equipamentos, enquanto harmônicas de sequência zero, especialmente as de ordem tripla, podem se somar no condutor neutro em sistemas a quatro fios, provocando sobrecorrentes [6].

A análise desses fenômenos é essencial para a avaliação da qualidade de energia elétrica e para a verificação da conformidade com os limites estabelecidos por normas técnicas, como a IEEE 519 [3], além de ser amplamente abordada em estudos de classificação e caracterização de distúrbios de qualidade de energia [11].

Os efeitos dessas distorções podem ser observados nas formas de onda apresentadas nas Figuras abaixo. A Figura 11 apresenta a senoide monofásica com a presença de harmônicas, evidenciando a deformação da forma de onda em relação ao comportamento senoidal ideal. Na Figura 12 observa-se o comportamento das harmônicas em um sistema trifásico, permitindo analisar os efeitos da distorção em cada fase. Já as Figuras 13 e 14 ilustram, respectivamente, a tensão de alimentação e a corrente de um motor quando submetido a sinais com conteúdo harmônico, demonstrando como as distorções presentes na tensão podem resultar em correntes ainda mais distorcidas.

A análise quantitativa dos sinais simulados foi realizada por meio do cálculo da Distorção Harmônica Total (THD), um dos principais indicadores utilizados para avaliar o nível de distorção em sistemas elétricos [3, 12]. A Distorção Harmônica Total é calculada conforme:

$$THD = \frac{\sqrt{\sum_{k=2}^{N} V_k^2}}{V_1} \tag{2}$$

onde V_k representa a amplitude da k-ésima harmônica e V_1 a componente fundamental.

A Tabela I apresenta os valores de THD obtidos em diferentes cenários simulados, permitindo comparar o comportamento do sistema em condições ideais e na presença de distorções harmônicas.

TABELA I
VALORES DE THD OBTIDOS EM DIFERENTES CENÁRIOS SIMULADOS

Cenário analisado	THD (%)
Senoide ideal (sem harmônicas)	0
Senoide com harmônicas (3ª e 5ª ordem)	22, 36
Retificador monofásico (carga não linear)	22, 73
Motor de indução com tensão distorcida	7, 17

Os resultados apresentados na Tabela I evidenciam a influência das harmônicas e das cargas não lineares na qualidade da energia elétrica. Observa-se que, na condição ideal sem a presença de harmônicas, o sinal apresenta THD nulo (0,00%), caracterizando uma forma de onda senoidal pura. Por outro lado, quando componentes harmônicas são inseridas ou quando o retificador de onda completa é simulado, ocorre um aumento significativo da distorção harmônica (ultrapassando os 22%), evidenciando a degradação da qualidade da energia no sistema

analisado. Esse comportamento está em conformidade com estudos da área e com os limites recomendados por normas técnicas, como a IEEE 519 [3]. Além disso, a tabela destaca o comportamento do modelo de motor de indução (circuito RL) frente a uma tensão distorcida. Mesmo submetido a uma tensão de alimentação com 22,36% de THD, a corrente resultante do motor apresentou uma distorção de apenas 7,17%. Esse resultado demonstra a consistência física do simulador, evidenciando o efeito de filtragem natural da reatância indutiva para componentes de alta frequência, que atenuam a propagação harmônica na corrente do equipamento.

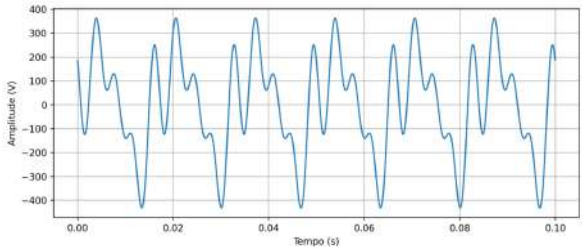


Fig. 11. Senoide Monofásica com Harmônicas.

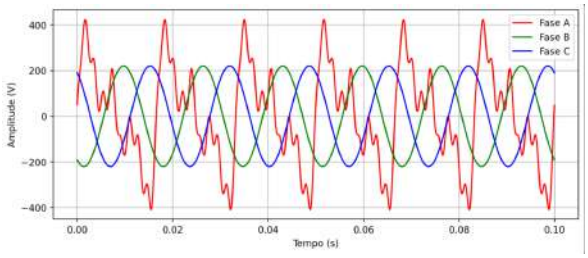


Fig. 12. Senoide Trifásica com Harmônicas.

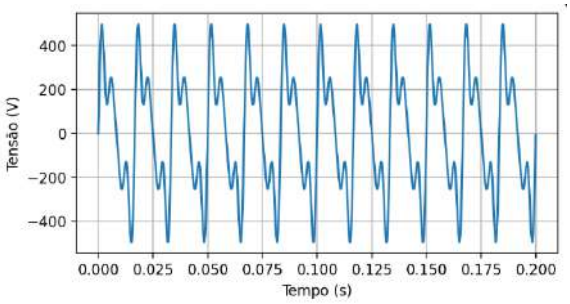


Fig. 13. Tensão de Alimentação do Motor com Harmônicas.

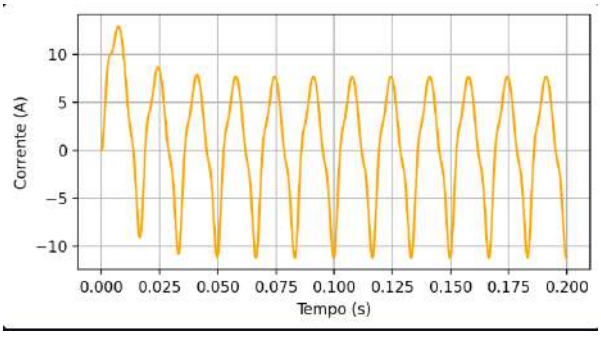


Fig. 14. Corrente do Motor com Harmônicas.

IV. Conclusão

Este trabalho apresentou as funcionalidades e aplicações de um simulador interativo de Qualidade de Energia Elétrica, capaz de reproduzir e analisar fenômenos típicos de sistemas elétricos, como distorções harmônicas, afundamentos e elevações de tensão.

Os resultados obtidos por meio das simulações permitiram analisar o impacto da inserção de distúrbios e componentes harmônicas na forma de onda dos sinais elétricos. A análise espectral realizada por meio da Transformada Rápida de Fourier possibilitou identificar as principais componentes de frequência presentes no sinal e calcular a Distorção Harmônica Total (THD). Em condições ideais, sem a presença de distúrbios ou harmônicas adicionais, os sinais apresentaram comportamento predominantemente senoidal, com valores de THD próximos de zero.

Quando componentes harmônicas foram inseridas, observou-se um aumento significativo do conteúdo harmônico e da Distorção Harmônica Total. Em alguns cenários simulados, os valores de THD ultrapassaram os limites recomendados para sistemas de potência, evidenciando a degradação da qualidade da energia elétrica. Esses resultados estão em conformidade com a literatura da área e com as recomendações da norma IEEE 519 [3].

Os resultados obtidos nas simulações apresentaram comportamento coerente com os experimentos realizados em laboratório, sendo essa validação conduzida por meio da comparação qualitativa entre as formas de onda e os níveis de Distorção Harmônica Total (THD) obtidos experimentalmente e por simulação. Durante os ensaios práticos com retificadores, motores elétricos e inversores de frequência, foram observadas distorções semelhantes nas formas de onda de tensão e corrente, bem como a presença de componentes harmônicas associadas a cargas não lineares. A comparação entre os resultados experimentais e computacionais reforça a validade do modelo implementado no simulador. Observou-se boa concordância entre os resultados simulados e experimentais, com comportamento semelhante das formas de onda e níveis de distorção harmônica.

Assim, o simulador desenvolvido demonstra ser uma ferramenta eficaz para o estudo e a visualização de fenômenos associados à qualidade de energia elétrica, podendo ser aplicado em atividades didáticas, estudos acadêmicos e análises preliminares de sistemas elétricos [12, 13].

Diferentemente de abordagens tradicionais encontradas na literatura, o simulador proposto integra, em uma única ferramenta interativa, modelagem matemática, visualização gráfica em tempo real e validação experimental, demonstrando capacidade de representar adequadamente fenômenos de distorção harmônica e validando sua aplicação como ferramenta de análise preliminar em sistemas elétricos reais.

Referências

[1] R. P. S. Leão e R. F. Sampaio. *Harmônicos em sistemas elétricos*. Rio de Janeiro, RJ: Elsevier, 2014, p. 354. ISBN: 978-85-352-7439-4.

[2] R. A. Lopez. *Qualidade na energia elétrica: efeitos dos distúrbios, diagnósticos e soluções*. São Paulo, SP: Artliber, 2013, p. 527. ISBN: 85-88098-02-4.

[3] *IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems*. New York: IEEE, 2014.

[4] Paulo Fernando Ribeiro e Ricardo Aldabó Lopes. *Qualidade de Energia Elétrica em Sistemas de Potência*. Rio de Janeiro: Ciência Moderna, 2017.

[5] Nagulapati Kiran. "Simulation of Power Quality Disturbances using MATLAB/SIMULINK". Em: *International Journal of Scientific Research in Engineering and Management (IJSREM)* 5.6 (2021), pp. 1–3. ISSN: 2582-3930.

[6] E. J. Robba. *Introdução a sistemas elétricos de potência: componentes simétricas*. São Paulo, SP: Edgard Blücher, 1973, p. 344.

[7] F. R. Martins e J. L. Barros. *Eletrônica de Potência: Conceitos e Aplicações*. Belo Horizonte: Editora UFMG, 2015.

[8] R. D. Gonçalves e A. H. Azevedo. *Máquinas Elétricas: Teoria e Prática*. São Paulo: LTC, 2016.

[9] Ivo Barbi. *Eletrônica de Potência*. Florianópolis: Editora da UFSC, 2015.

[10] O. Jeba Singh e Prince Winston. "A Survey on Classification of Power Quality Disturbances in a Power System". Em: *International Journal of Engineering Research and Applications (IJERA)* 4.8 (2014), pp. 80–84. ISSN: 2248-9622.

[11] Maísa L. F. Santos, Suzete E. N. Correia, Cleumar S. Moreira, Álvaro de M. Maciel e Edgard L. L. Fabricio. "Análise de Distúrbios da Qualidade da Energia Elétrica Empregando a Decomposição Variacional de Modos". Em: *Congresso Brasileiro de Automática (CBA)*. 2022.

[12] J. L. R. Pereira e E. J. Silva. *Qualidade de Energia Elétrica*. São Paulo: Editora Blucher, 2016.

[13] Agência Nacional de Energia Elétrica. *Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST: Módulo 8 – Qualidade da Energia Elétrica*. ANEEL. Brasília, 2023.

AUTORES



Geovane Mantuan Alvim é graduando do terceiro período do curso de Bacharelado em Engenharia de Controle e Automação pelo INATEL – Instituto Nacional de Telecomunicações, bolsista de Iniciação Científica e representante de categoria na equipe de robótica Inatel RobotBulls.



Júlia da Silva Maciel é graduanda do quinto período do curso de Bacharelado em Engenharia Elétrica pelo INATEL – Instituto Nacional de Telecomunicações,foi bolsista de Iniciação Científica e voluntária na International Society of Automation (ISA).



Natan Barreto de Paula é Técnico em Sistemas I na FINATEL – Fundação Instituto Nacional de Telecomunicações, atuando no Programa Inatel cas@viva, com foco em atividades de ensino, tecnologia e inovação. E graduando do décimo período do curso de Bacharelado em Engenharia Elétrica pelo INATEL – Instituto Nacional de Telecomunicações. E cofundador da startup Bullvest, iniciativa voltada a educação financeira gamificada e ao uso de inteligência artificial em simulações de investimento, e formado como Técnico em Eletromecânica pela ETPC – Escola Técnica Pandiá Calógeras.

ORIENTADOR



Egídio Raimundo Neto possui Graduação em Engenharia Elétrica (2011) e Mestrado em Telecomunicações (2014) pelo Instituto Nacional de Telecomunicações. Atua como Professor e Coordenador do curso de Engenharia Elétrica do INATEL. Suas áreas de interesse incluem Sistemas de Controle Clássico, Discreto e Moderno, Redes Industriais, Robótica, Drones Autônomos, Eletrônica de Potência e Smart Grid.

Desenvolvimento e Validação de um Sistema Automatizado para Calibração e Caracterização Metrológica de Sensores de Umidade de Solo de Baixo Custo

Mateus Sabio Martins & Carlos Alberto De Francisco
Universidade Federal de São Carlos (UFSCar)
mateussabiomartins@estudante.ufscar.br, cfrancisco@ufscar.br

***Abstract*—**Precision agriculture relies critically on reliable soil moisture data to optimize water resource usage and maximize productivity. However, low-cost sensors used in IoT systems suffer from low accuracy and high variability, especially under fluctuating temperature conditions. In this work, an Automated Thermal Control Calibration Platform was proposed and validated to calibrate and characterize multiple soil moisture sensors simultaneously, with a specific focus on thermal compensation. The system was built using a calibration chamber with closed-loop temperature control, a high-resolution data acquisition system (DAQ), and a rigorous protocol for manual incremental moisture addition (gold standard). The gravimetric method was used as the reference, allowing the generation of precise calibration curves at different thermal plateaus to quantify the influence of temperature on the electrical conductivity of the porous medium. Preliminary results demonstrate a strong positive correlation between temperature and apparent humidity, highlighting the need for compensation algorithms for professional applications.

***Index Terms*—**IoT, Precision Agriculture, Sensor Calibration, Soil Moisture, Thermal Compensation.

***Resumo*—**A agricultura de precisão depende criticamente de dados confiáveis sobre a umidade do solo para otimizar o uso de recursos hídricos e maximizar a produtividade. No entanto, os sensores de baixo custo utilizados em sistemas IoT sofrem com baixa precisão e alta variabilidade, especialmente em condições de temperatura flutuante. Neste trabalho, propõe-se e valida-se uma Plataforma de Calibração com Controle Térmico Automatizado desenvolvida para calibrar e caracterizar múltiplos sensores de umidade simultaneamente, com foco específico na compensação térmica. O sistema foi construído utilizando uma câmara de calibração com controle de temperatura em malha fechada, um sistema de aquisição de dados (DAQ) de alta resolução e um protocolo rigoroso para adição manual incremental de umidade (padrão-ouro). Utilizou-se o método gravimétrico como referência, permitindo gerar curvas de calibração precisas em diferentes patamares térmicos para quantificar a influência da temperatura na condutividade elétrica do meio poroso. Os resultados preliminares demonstram uma forte correlação positiva entre a temperatura e a humidade aparente, evidenciando a necessidade de algoritmos de compensação para aplicações profissionais.

***Palavras-chave*—**Agricultura de Precisão, Calibração de Sensores, Compensação Térmica, IoT, Umidade do Solo.

I. INTRODUÇÃO

A gestão eficiente da água na agricultura é um dos maiores desafios globais deste século. Segundo a Organização das Nações Unidas para Alimentação e Agricultura (FAO), a agricultura é responsável por 70% do consumo de água doce no mundo. Nesse cenário, o monitoramento preciso da umidade do solo é a chave para garantir a sustentabilidade da irrigação, permitindo diagnósticos precisos e evitando tanto o desperdício hídrico quanto o estresse das plantas [1]. No contexto da agricultura familiar e de cultivos protegidos (estufas), observa-se que o uso de substratos comerciais orgânicos é predominante devido às suas propriedades físicas favoráveis à retenção de água e aeração das raízes.

A democratização da tecnologia IoT permitiu que pequenos produtores tivessem acesso a ferramentas antes restritas a grandes latifúndios. No entanto, a "ilusão da precisão" criada por sensores não calibrados pode levar a decisões agronômicas errôneas, como a irrigação deficitária (quebra de safra) ou excessiva (lixiviação de nutrientes e desperdício de energia e água [1]). Portanto, validar a robustez desses dispositivos não é apenas uma questão técnica, mas econômica e ambiental [1].

Entretanto, identificou-se uma lacuna tecnológica crítica que é a confiabilidade de sensores de baixo custo quando aplicados em substratos orgânicos adubados. A maioria dos sensores acessíveis no mercado baseia-se na medição da condutividade elétrica ou da permissividade dielétrica. É crucial distinguir as tecnologias: enquanto sensores capacitivos medem a constante dielétrica do meio, o sensor HD-38 (objeto deste estudo) é do tipo resistivo, medindo diretamente a resistência elétrica entre seus eletrodos. A presença elevada de íons dissolvidos em substratos adubados, somada à variação de temperatura, altera drasticamente a condutividade galvânica do meio. O aumento da temperatura reduz a viscosidade da solução do solo e aumenta a mobilidade iônica, o que diminui a resistência elétrica aparente medida pelo sensor resistivo, mimetizando um aumento de umidade. Embora trabalhos seminais como o de Kizito et al. [2] foquem em sensores capacitivos, o princípio físico da mobilidade iônica afeta de forma ainda mais direta os sensores resistivos [3], [6].

A dependência térmica da condutividade elétrica (σ) em meios porosos pode ser descrita, em primeira aproximação, por uma relação linear com a temperatura (T):

$$\sigma\_T = \sigma\_25 [1 + \alpha(T - 25)] \quad (1)$$

Onde σ_{25} é a condutividade elétrica a 25°C e α é o coeficiente de temperatura, que tipicamente varia entre 0.019 e 0.022 $°C^{-1}$ para soluções salinas padrão [2]. No entanto, em solos complexos, a interação entre a fase sólida e a líquida introduz não-linearidades que desviam deste modelo ideal [4]. Estudos recentes demonstram que sensores resistivos de baixo custo, como o modelo HD-38, podem apresentar erros de medição de umidade volumétrica (VWC) superiores a 5% apenas devido a flutuações diurnas de temperatura, sem qualquer alteração real no teor de água [5], [6].

A dependência térmica da permissividade dielétrica da água é um fenômeno físico bem conhecido, onde a constante dielétrica diminui com o aumento da temperatura. No entanto, em solos salinos ou substratos ricos em nutrientes, o efeito da condutividade iônica (que aumenta com a temperatura) frequentemente domina, criando um comportamento complexo e difícil de prever sem dados empíricos. Sensores comerciais de alto custo (TDR/FDR) possuem termistores integrados e algoritmos proprietários para compensar esses erros, mas essas soluções são economicamente inviáveis para pequenos produtores ou para redes de sensores de alta densidade pois esses sensores de alto custo custam 100 vezes mais que os de sensores de baixo custo.

Além disso, a variabilidade intrínseca entre unidades do mesmo modelo de sensor de baixo custo exige uma calibração individual ou por lote. A maioria dos trabalhos na literatura foca na calibração em solos minerais (areia, argila), ignorando a complexidade dos substratos organominerais utilizados em estufas, que possuem alta capacidade de troca catiônica (CTC) e retêm umidade de forma diferente [7].

A principal contribuição e originalidade deste trabalho consistem no desenvolvimento e validação de uma Plataforma de Calibração com Controle Térmico Automatizado, projetada para isolar matematicamente os efeitos da umidade e da temperatura num substrato comercial organomineral. O objetivo direto é provar a necessidade mandatória de acoplar sensores de temperatura aos de umidade, permitindo a futura criação de algoritmos de compensação que elevem o desempenho de equipamentos de baixo custo a padrões de precisão científica.

II. MATERIAIS E MÉTODOS

A metodologia foi desenhada para estabelecer uma curva de calibração robusta, onde se confronta a resposta elétrica do sensor com o método gravimétrico (padrão-ouro) em toda a faixa de disponibilidade hídrica do solo, sob condições térmicas controladas.

A. Caracterização do Substrato

Utilizou-se um substrato comercial organomineral (Terra Vegetal Adubada Orgânica Mogifertil) [11], amplamente empregado na horticultura brasileira. Escolheu-se este material pois sua composição rica em turfa negra, casca de pinus, serragem e carvão vegetal cria uma matriz de baixa densidade, alta porosidade e alta complexidade dielétrica. Este tipo de meio representa o "pior caso" para testes de estabilidade de sensores, pois a decomposição da matéria orgânica libera íons continuamente, e a estrutura porosa pode criar bolsões de ar que afetam o contato solo-sensor. A densidade aparente do substrato seco foi determinada em 0,45 g/cm³, significativamente menor que a de solos minerais típicos (1,2 – 1,4 g/cm³).

B. Hardware e Sistema de Aquisição (DAQ)

A arquitetura do sistema foi projetada (Fig. 1) dividida em dois subsistemas independentes baseados no microcontrolador ESP32, garantindo a modularidade e evitando interferências elétricas entre a alta potência (aquecimento) e a instrumentação sensível (sensores).

1. Subsistema de Controle da Câmara (ESP32 Independente 1): Este módulo é responsável exclusivamente pela termodinâmica da câmara de calibração.

- Câmara Térmica: Construiu-se a câmara utilizando uma caixa de isopor de 17 litros com revestimento térmico de Etaflon, garantindo o isolamento adiabático.

- Atuadores: Instalou-se uma lâmpada halógena de 100W (elemento resistivo de aquecimento) e uma ventoinha (convecção forçada) para homogeneização da temperatura do ar. Ambos são acionados por um módulo MOSFET de alta potência (D4184, 40V/50A) via PWM [12], [13].

- Sensor de Referência: Um sensor digital DS18B20 posicionado no centro geométrico da caixa (no ar) fornece a realimentação para o controle de temperatura.

- Controle PID (IMC): Implementou-se um algoritmo de controle Proporcional-Integral-Derivativo (PID) no firmware do ESP32 para manter a temperatura estável. A sintonia dos ganhos (K_p, K_i, K_d) foi realizada utilizando o método de Controle por Modelo Interno (IMC) [14], que permite um ajuste robusto baseado na constante de tempo e no ganho estático do sistema térmico identificados em malha aberta. A saída do controle $u(t)$ é calculada como:

$$u(t) = K\_p\, e(t) + K\_i \int e(\tau)\, d\tau + K\_d\, de(t)/dt \qquad (2)$$

Onde $e(t)$ é o erro entre o setpoint e a temperatura medida. Isso permite rampas de aquecimento suaves e manutenção precisa da temperatura (± 0,5°C).

A escolha do método IMC para sintonia do PID justifica-se pela sua robustez frente a variações nos parâmetros da planta. Diferente de métodos heurísticos como Ziegler-Nichols, que frequentemente resultam em respostas oscilatórias (inaceitáveis para calibração biológica), o IMC permite projetar o controlador visando uma resposta de malha fechada específica, garantindo que a temperatura do solo atinja o equilíbrio sem ultrapassagens (overshoot) que poderiam induzir histerese térmica nos sensores.

2. Subsistema de Aquisição de Dados (ESP32 Independente 2): Este módulo foca na leitura precisa dos sensores inseridos no solo.

- Conversão A/D (ADS1115): Para superar a não-linearidade do ADC interno do ESP32, utilizou-se um módulo ADS1115 externo (16 bits, I2C). O ganho foi configurado para ±4,096V, obtendo uma resolução de 0,125mV por bit. Optou-se por esta faixa, mesmo com o ESP32 operando em 3,3V, para garantir a linearidade da leitura longe dos limites de saturação e simplificar o circuito, aceitando a leve redução na faixa dinâmica efetiva.

- Sensores de Umidade (HD-38): O elemento sensor sob teste é o modelo resistivo HD-38 (Fig. 2), que atua como um resistor variável dependente da umidade.

- Sensores de Temperatura do Solo (DS18B20): Sensores adicionais foram enterrados junto com as sondas de umidade para monitorar a temperatura real da amostra, que difere da temperatura do ar devido à inércia térmica.

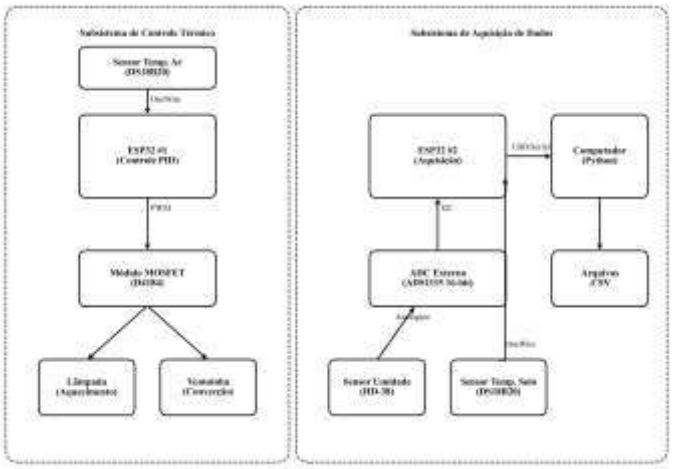


Fig. 1. Diagrama de blocos da arquitetura de hardware e fluxo de dados.



Fig. 2. Sensor de umidade de solo resistivo HD-38 com modulo incluso utilizado nos ensaios.

C. Protocolo de Calibração

Para garantir a integridade e reprodutibilidade dos dados, adotou-se uma metodologia híbrida que combina a gravimetria com o controle ativo de temperatura. Vale ressaltar que, nesta etapa da pesquisa, a adição de água foi realizada manualmente (padrão-ouro), embora a plataforma permita a automação futura deste processo [9], [10].

1. Preparação das Amostras: Foram preparados 2 corpos de prova (potes) independentes com substrato, cobrindo diferentes níveis de umidade para validar a resposta do sensor. Para a análise apresentada neste trabalho, utilizou-se o Pote 1 (10% de UG) e o Pote 2 (21% de UG) para demonstrar que a variação térmica afeta significativamente a leitura do sensor em diferentes patamares de umidade. A umidade de cada pote foi ajustada artificialmente por balanço de massa. Após a adição de água, os potes foram vedados hermeticamente com filme plástico para impedir a evaporação e garantir que a massa de água (e, portanto, a umidade gravimétrica) permanecesse constante durante todo o ciclo de aquecimento.

2. Instrumentação: Posicionou-se um pote por vez dentro da câmara térmica para evitar interferências térmicas mútuas e garantir a homogeneidade do aquecimento. Os sensores de umidade e temperatura foram inseridos no centro geométrico de cada amostra, garantindo contato íntimo com o substrato.



Fig. 3. Configuração experimental com um único pote dentro da câmara térmica.

3. Ciclo de Teste Automatizado: O firmware opera como uma Máquina de Estados Finitos, conforme detalhado no fluxograma da Fig. 4. O sistema ajusta a temperatura da câmara para setpoints definidos (ex: 30°C, 55°C) e monitora a temperatura do solo. O sistema só inicia a gravação dos dados de umidade quando a variação de temperatura do solo é menor que 0,1°C por minuto, indicando equilíbrio térmico.

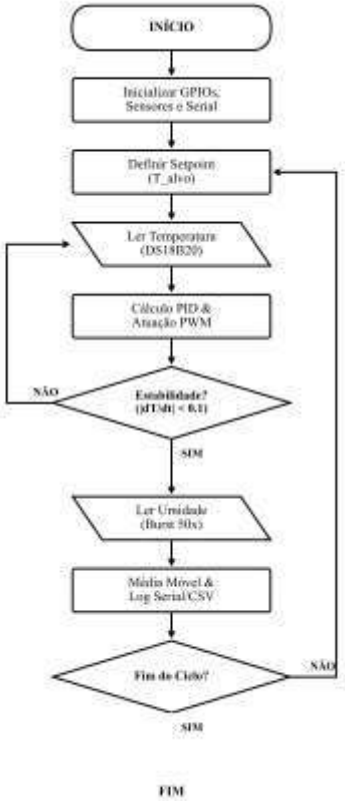


Fig. 4. Fluxograma do firmware do sistema de controle e aquisição.

4. Aquisição e Processamento: Assim que o equilíbrio térmico é atingido, o ESP32 realiza uma amostragem em burst (50 leituras em 1 segundo) e calcula a média para filtrar ruídos elétricos de alta frequência. Os dados consolidados são transmitidos via porta serial para um computador. Desenvolveu-se um script em Python (serial_logger.py) que utiliza a biblioteca pandas para estruturar e salvar esses dados em arquivos CSV, facilitando a análise posterior.

III. RESULTADOS PRELIMINARES E DISCUSSÃO

A. Caracterização da Câmara Térmica

Inicialmente, validou-se a resposta dinâmica do sistema de aquecimento realizando um ensaio em malha aberta (degrau de potência). O gráfico de caracterização (Fig. 5) demonstrou que o sistema, partindo da temperatura ambiente (~24°C), atingiu 50°C em aproximadamente 320 segundos (5,3 minutos). A curva de aquecimento apresentou um comportamento exponencial característico de sistemas térmicos de primeira ordem, sem overshoot significativo.

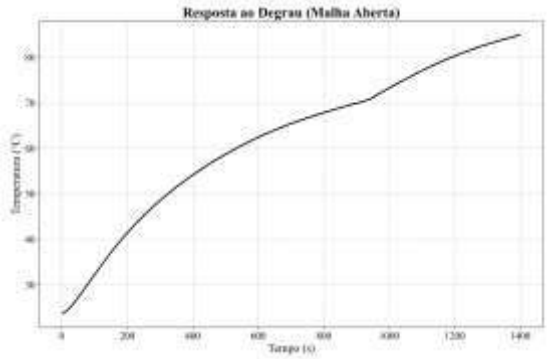


Fig. 5. Identificação da Planta Térmica em Malha Aberta (Resposta ao Degrau).

Para o controle em malha fechada, utilizou-se o método IMC para sintonizar o controlador PID. O desempenho do controlador pode ser visualizado na Figura 6, onde o sistema segue o setpoint com precisão e estabilidade, mesmo após a breve abertura da câmara (distúrbio).

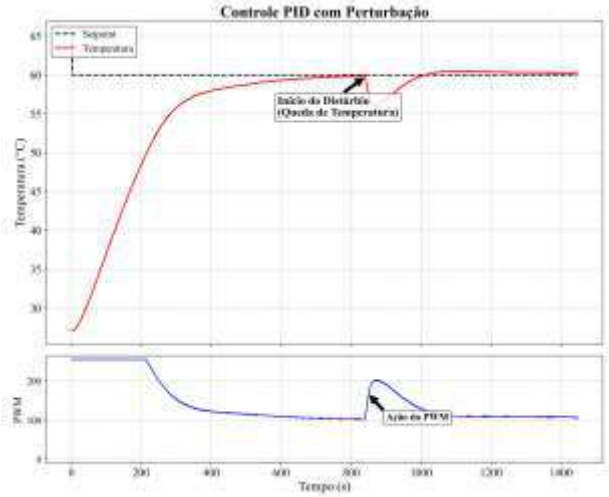


Fig. 6. Resposta em Malha Fechada do Controlador PID sob Perturbação Externa.

B. Análise Experimental dos Potes

Realizaram-se os ensaios de calibração com níveis controlados de umidade. A Tabela I apresenta os dados obtidos durante a preparação das amostras, onde utilizou-se uma balança de precisão para determinar a Umidade Gravimétrica (UG) real com exatidão. A variação entre a UG calculada e a real foi inferior a 1,1%, validando o protocolo de preparação.

TABELA I
DADOS DE PREPARAÇÃO DAS AMOSTRAS DE SOLO

Amostra	Terra Seca (g)	Água Calc. (ml)	UG Alvo	Água Real (ml)	UG Real	Peso Total (g)
Pote 1	475	47,5	10%	48	10,11%	538
Pote 2	507	101,4	20%	107	21,10%	649

Para compreender a dinâmica térmica do sistema solo-sensor, é fundamental observar primeiramente o comportamento temporal das variáveis durante a janela de aquecimento controlada de 30°C a 58°C. As Figuras 7 e 8 apresentam a evolução temporal da temperatura e da leitura bruta de umidade para os Potes 1 e 2, respectivamente.

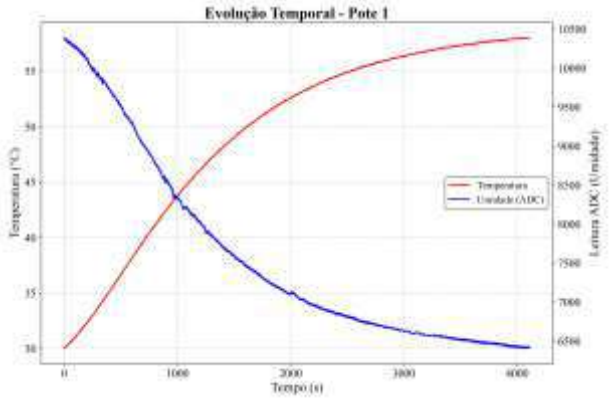


Fig. 7. Evolução temporal da temperatura e umidade no Pote 1 (Janela 30°C - 58°C).

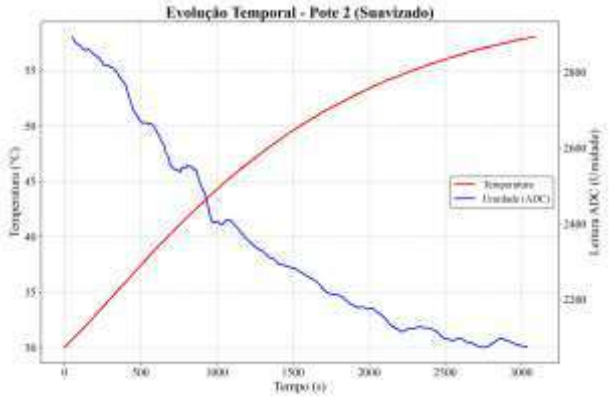


Fig. 8. Evolução temporal da temperatura e umidade no Pote 2 (Janela 30°C - 58°C).

Observa-se nas Figuras 7 e 8 que a resposta do sensor de umidade (curva azul) acompanha inversamente a elevação da temperatura (curva vermelha) de forma quase instantânea. No Pote 1 (10% UG), a leitura bruta cai de aproximadamente 10.500 para 6.500 unidades ADC, uma variação de quase 40% apenas devido ao efeito térmico. No Pote 2 (21% UG), a variação é menor em magnitude absoluta, mas igualmente significativa, caindo de ~3.000 para ~2.100. A suavização

aplicada aos dados do Pote 2 (média móvel) permitiu visualizar a tendência clara de queda, eliminando ruídos de alta frequência inerentes à medição eletrônica.

É importante esclarecer a relação entre a grandeza física e a leitura digital. O circuito de condicionamento do sensor HD-38 opera como um divisor de tensão onde o solo atua como uma resistência variável. O conversor ADS1115 lê a tensão sobre o sensor. Neste arranjo, o aumento da temperatura eleva a mobilidade iônica, aumentando a condutividade (σ) e reduzindo a resistência elétrica (R_solo) do substrato. Como a leitura do ADC é proporcional à resistência neste circuito, a leitura bruta diminui com o aumento da condutividade, mimetizando um falso aumento de umidade.

A partir desses dados temporais, correlacionou-se a leitura do sensor diretamente com a temperatura para quantificar a dependência térmica. As Figuras 9 e 10 apresentam a análise de regressão linear para os Potes 1 e 2.

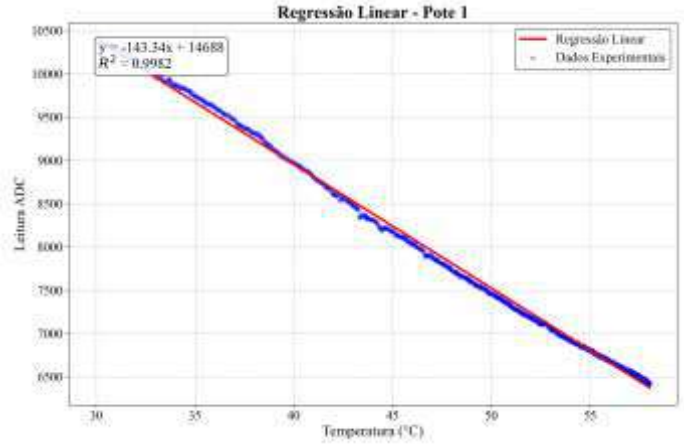


Fig. 9. Regressão linear da leitura de umidade vs temperatura no Pote 1 (R^2 = 0.9982).

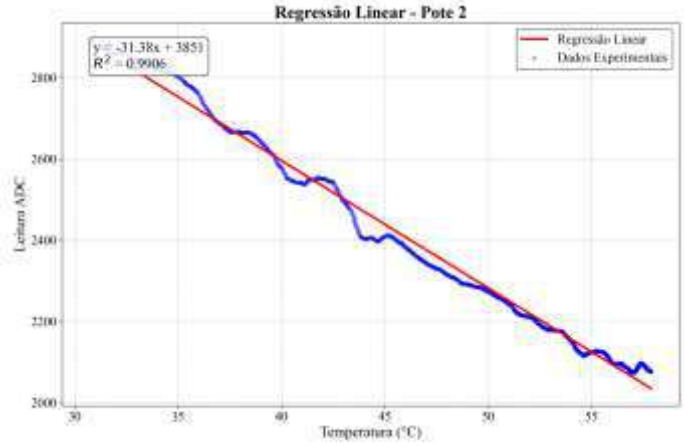


Fig. 10. Regressão linear da leitura de umidade vs temperatura no Pote 2 (R^2 = 0.9910).

No Pote 2 (21% de Umidade), obteve-se um R^2 = 0,9910 com a equação y = -31,30x + 3.847. Embora a linearidade se mantenha alta, nota-se uma mudança drástica na inclinação da curva (coeficiente angular de -31,30 comparado a -143,34 do Pote 1). Essa variação na sensibilidade térmica confirma as observações de Wang et al. [10] e Kizito et al. [2], que alertam para a dependência do coeficiente de temperatura em relação ao conteúdo volumétrico de água (θ). A presença de mais água altera a concentração iônica e a permissividade dielétrica efetiva, modificando a resposta térmica do conjunto sensor-solo.

A alta linearidade comprovada pelos valores de $R^2 > 0{,}98$ em ambos os casos valida a hipótese de que a resposta térmica do sensor é previsível. Entretanto, dado que o coeficiente angular varia significativamente com a umidade, propõe-se uma abordagem de calibração baseada em uma superfície de resposta gerada a partir de 10 potes com umidades distintas (cobrindo a faixa de 0% a 40% de Umidade Gravimétrica). Através da coleta de dados em uma rampa térmica de 25 °C a 70 °C para cada nível de umidade, a umidade real (UG) é determinada por uma função de interpolação bidimensional:

$$UG = f(ADC_{lido}, T)$$

Onde o espaço de estados é definido pela matriz de calibração que cruza os valores brutos do ADC e a temperatura lida. A ausência de histerese significativa observada durante a rampa de aquecimento e o comportamento monotônico reforçam a viabilidade de utilizar algoritmos de interpolação (como a bilinear ou por *splines*) para mapear esses sensores de baixo custo, permitindo que o sistema compense as flutuações térmicas em tempo real e atinja precisão científica.

Esses resultados corroboram a hipótese inicial e estão alinhados com estudos recentes da Embrapa e literatura internacional [5], [9], que apontam a temperatura como a principal fonte de erro em sensores dielétricos e resistivos de baixo custo.

C. Limitações do Estudo

Sendo uma investigação em fase preliminar, o presente estudo possui limitações que devem ser consideradas. A base experimental focou-se na demonstração da plataforma através da análise detalhada de dois níveis de umidade gravimétrica (10% e 21%). Embora os resultados comprovem a elevada correlação térmica, a generalização do método requer a expansão da base de dados para abranger uma gama mais ampla de umidades, diferentes modelos de sensores e distintos tipos de substratos agrícolas. Além disso, a atual adição manual de água carece de automação para garantir uma repetibilidade integral do processo nos trabalhos futuros.

IV. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho, apresentou-se o desenvolvimento e a validação de uma plataforma automatizada de baixo custo para a caracterização metrológica de sensores de umidade do solo. A arquitetura modular provou ser eficaz para garantir a estabilidade das medições. Os resultados quantitativos comprovaram que a variação de temperatura introduz um erro sistemático crítico: observou-se uma queda na leitura digital

de até 40% (de 10.500 para 6.500 unidades ADC no Pote 1) induzida puramente pelo aquecimento, mimetizando falsos aumentos de umidade.

As regressões lineares demonstraram coeficientes de determinação bastante elevados ($R^2 = 0,9982$ e $R^2 = 0,9910$), validando a viabilidade de modelar matematicamente este erro. Confirma-se, assim, que as leituras destes sensores não podem ser utilizadas na agricultura de precisão sem compensação térmica algorítmica. Como trabalhos futuros, pretende-se alargar a base experimental para gerar superfícies de resposta tridimensionais completas, integrando totalmente a automação da adição de água para garantir a generalização do método a diferentes substratos.

REFERÊNCIAS

[1] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, and K.-K. R. Choo, "Internet of Things for the Future of Smart Agriculture: A Comprehensive Survey of Emerging Technologies," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 4, pp. 718-752, 2021.

[2] F. Kizito et al., "Frequency, electrical conductivity and temperature analysis of a low-cost capacitance soil moisture sensor," Journal of Hydrology, vol. 352, no. 3-4, pp. 367–378, 2008.

[3] Z. A. Bello, S. Shittu, and I. O. A. Odeh, "Investigation of temperature effects and performance of capacitance soil moisture sensors," Measurement, vol. 138, pp. 634–643, 2019.

[4] D. Hillel, Environmental Soil Physics. San Diego: Academic Press, 1998.

[5] S. Adla et al., "Impact of calibrating a low-cost capacitance-based soil moisture sensor using multiple least squares and machine learning models," Journal of Environmental Management, vol. 352, p. 120263, 2024.

[6] J. M. Datta et al., "Laboratory Calibration and Performance Evaluation of Low-Cost Capacitive and Very Low-Cost Resistive Soil Moisture Sensors," Sensors, vol. 20, no. 2, p. 363, 2020.

[7] A. Szypłowska et al., "Soil texture, salinity and frequency effects on capacitance soil moisture sensors: A laboratory assessment," Journal of Hydrology, vol. 579, p. 124155, 2019.

[8] Vaz, C.M.P., Jones, S., Meding, M. and Tuller, M. (2013) Evaluation of Standard Calibration Functions for Eight Electromagnetic Soil Moisture Sensors. Vadose Zone Journal, 12.

[9] L. H. Bassoi et al., "Calibração de sensores de umidade do solo," Embrapa Semiárido, Comunicado Técnico 145, 2012.

[10] Y. Wang et al., "Methods to correct temperature-induced changes of soil moisture sensors to improve accuracy," Vadose Zone Journal, vol. 23, no. 1, e20305, 2024.

[11] MOGIFERTIL**.** Terra Vegetal Adubada Orgânica: substrato organomineral. Mogi das Cruzes, SP, 2026. Disponível em: [https://www.mogifertil.com.br/produto/terra-vegetal/]. Acesso em: 4 mar. 2026.

[12] Alpha & Omega Semiconductor, "AOD4184A: N-Channel Enhancement Mode Field Effect Transistor," Datasheet, Rev. 5.0, 2011 .Acesso em: 4 mar. 2026.

[13] RobotDyn, **"MOSFET Transistor Module: DC Control 5-40V 50A,"** Technical Specification, 2022.

[14] Rivera, D. E., M. Morari, and S. Skogestad, “Internal model control: PID controller design,” Ind. Eng. Chem. Process Des. Dev., vol. 25, no. 1, pp. 252–265, 1986.

AUTORES

Mateus Sabio Martins é estudante de Engenharia Elétrica na Universidade Federal de São Carlos (UFSCar). Atua em projetos de pesquisa na área de instrumentação agrícola e IoT.

Carlos Alberto De Francisco é Professor do Departamento de Engenharia Elétrica na Universidade Federal de São Carlos (UFSCar) e orientador do projeto. Possui experiência em sistemas embarcados e automação.

PROJETO DE UM PRÉ AMPLIFICADOR DE ÁUDIO DE ALTO DESEMPENHO SEM REALIMENTAÇÃO NEGATIVA GLOBAL

Lucas Alan da Conceição Batista, Carlos Alberto De Francisco
Universidade Federal de São Carlos - UFSCar
lucasbatista@estudante.ufscar.br, cfrancisco@ufscar.br

***Abstract*—The design of audio amplifiers is traditionally based on the use of Global Negative Feedback (G-NFB). However, this technique has recently raised discussions regarding its possible impact on perceived sound quality. In this context, this paper presents the development of a three-stage audio preamplifier powered by a symmetric 30 V supply, designed to amplify an input signal of 20 mVrms to an output level of 2 Vrms. The adopted methodology is based on parametric variation of the biasing conditions of each stage, allowing the identification of operating points that minimize distortion while preserving bandwidth and dynamic response. The obtained results demonstrate satisfactory performance, achieving a Total Harmonic Distortion (THD) of 0.0032%, a Slew Rate of 3.48 V/μs, 39.88dB gain and an upper cutoff frequency of 268.5 kHz.**

***Index Terms*—audio amplifiers, global negative feedback, parametric variation, preamplifier.**

***Resumo*—O projeto de amplificadores de áudio é tradicionalmente baseado no uso de Realimentação Negativa Global (G-NFB). Entretanto, essa técnica tem sido alvo de debates quanto à possível degradação da qualidade sonora. Nesse contexto, este artigo apresenta o desenvolvimento de um pré-amplificador de áudio de três estágios, alimentado por tensão simétrica de 30 V, projetado para amplificar um sinal de entrada de 20 mVrms para uma saída de 2 Vrms. A metodologia adotada baseia-se na variação paramétrica dos pontos de polarização de cada estágio, permitindo a determinação de condições de operação que minimizem a distorção e preservem o desempenho em frequência e velocidade de resposta. Os resultados obtidos indicam desempenho satisfatório, com Distorção Harmônica Total (THD) de 0,0032%, *Slew Rate* de 3,48 V/μs, ganho de 39,88dB e frequência de corte superior de 268,5 kHz.**

***Palavras chave*—amplificadores de áudio,pré-amplificador, realimentação negativa global, variação paramétrica.**

I. Introdução

Desde os primeiros projetos de amplificadores de áudio, tem-se adotado intensivamente a Realimentação Negativa Global (G-NFB), do inglês Global Negative Feedback. Nessa técnica, parte do sinal de saída é invertida e reinjetada na entrada para subtrair o erro causado pelas não linearidades dos estágios de amplificação. A utilização desse método traz diversas melhorias no desempenho do amplificador, como aumento da impedância de entrada, redução da impedância de saída, ampliação da resposta em frequência e controle da distorção harmônica total (THD), do inglês Total Harmonic Distortion [1].

Entretanto, discussões sobre o uso dessa técnica têm crescido recentemente na comunidade audiófila. Críticos argumentam que esse recurso pode provocar compressão excessiva do sinal, restringir a sensação de espaço sonoro e reduzir a musicalidade do sistema [2, 3].

A fundamentação teórica dessas críticas está principalmente na geração de harmônicos de ordens mais elevadas. Devido às não linearidades naturais dos dispositivos eletrônicos, surgem inicialmente harmônicos de segunda ordem em configurações *single-ended* e de terceira ordem em topologias *push-pull* ou de par diferencial. Quando essas componentes de ordem mais baixa são reinjetadas ao circuito pela própria realimentação, podem gerar harmônicos de quarta, quinta, sexta e até sétima ordem na saída — frequências ainda mais audíveis e, segundo muitos ouvintes, menos agradáveis [4, 3].

Diante dessa problemática, fabricantes de equipamentos de alta fidelidade passaram a desenvolver amplificadores sem o uso de realimentação negativa global [5]. Contudo, muitas dessas topologias são protegidas por direitos comerciais e não estão disponíveis publicamente. Além disso, o processo de projeto de amplificadores tradicionais frequentemente baseia-se na realimentação negativa global. Nesse caso, o circuito completo é inicialmente projetado e o controle dos principais parâmetros de desempenho ocorre posteriormente por meio do fator de realimentação. Ou seja, o ajuste do sistema acontece apenas ao final do projeto, sem grande preocupação com as não linearidades introduzidas em cada estágio, já que estas serão compensadas posteriormente pelo próprio fator de realimentação.

Nesse contexto, Visintin et al. [6] propõem uma metodologia estruturada para o desenvolvimento de amplificadores de áudio sem o uso de realimentação negativa global. Ela consiste em projetar os estágios do amplificador sequencialmente, usando realimentação negativa local e escolhendo cuidadosamente pontos de operação e ganhos individuais, de modo que cada estágio seja otimizado considerando os anteriores, minimizando a propagação de distorções e garantindo o desempenho global do sistema.

Assim sendo, este trabalho apresenta o desenvolvimento de um pré-amplificador de áudio de alto desempenho sem o uso da G-NFB, com o objetivo de aplicar e validar a metodologia proposta por [6].

II. Materiais e Métodos

A. Escolha da arquitetura final

Para evitar a propagação da distorção, deve-se escolher uma arquitetura com o menor número possível de estágios.

B. Escolha dos componentes comerciais

O projeto de um pré-amplificador difere em alguns aspectos do projeto de um amplificador de potência. Portanto, em cada estágio a seleção dos componentes deve ser feita cuidadosamente. Os principais fatores de escolha são: baixa densidade espectral de ruído, menores capacitâncias parasitas — favorecendo uma maior taxa de variação de tensão, do inglês, *Slew Rate* (SR)— e alta tensão de trabalho, permitindo que, mesmo para grandes excursões de sinal, o circuito opere na região linear.

C. Otimização paramétrica dos estágios para determinação de Slew-rate, THD e resposta em frequência

Após a definição da arquitetura e dos dispositivos utilizados, cada estágio do circuito é otimizado por meio de varreduras paramétricas das correntes de polarização e dos ganhos locais. As rotinas de simulação foram automatizadas por meio de scripts em Python, responsáveis por executar as simulações no LTspice e extrair os dados de interesse [7, 8]. Inicialmente, o primeiro estágio é analisado isoladamente para a determinação do ponto ótimo de operação. Em seguida, o segundo estágio é conectado ao primeiro e seus parâmetros são ajustados considerando a influência do estágio anterior. O mesmo procedimento é aplicado aos estágios subsequentes.

III. Critérios de Desempenho

O pré-amplificador atua sobre sinais de baixa amplitude — como os provenientes de microfones, instrumentos musicais ou fontes de linha — elevando-os a níveis adequados para posterior amplificação. Assim, o circuito projetado tem como objetivo elevar um sinal típico de $20\,\mathrm{mV_{rms}}$ para $2\,\mathrm{V_{rms}}$. A Tabela I apresenta os principais requisitos de desempenho considerados para o projeto.

A. Slew Rate

Para a determinação do SR, existem dois critérios que podem ser utilizados. O primeiro tem como finalidade estimar o valor mínimo necessário para evitar distorções transitórias [9], podendo ser calculado pela Equação 1:

$$SR = 2\pi V_p f_{\max} \tag{1}$$

em que V_p representa a tensão de pico do sinal de saída e $f_{\max}$ a frequência máxima considerada.

Outros autores, entretanto, sugerem um critério mais rigoroso — apresentado na Equação 2 — com a justificativa de proporcionar melhor qualidade sonora [10]:

$$SR = 0,5\,V_p \tag{2}$$

B. THD

No estado da arte é possível encontrar pré-amplificadores com THD da ordem de 0.003% [11]. Entretanto, esse nível de desempenho está geralmente associado a equipamentos de custo bastante elevado. Por outro lado, dispositivos comerciais de alta qualidade apresentam valores de THD próximos de 0.1% [12, 13].

C. Resposta em Frequência

O circuito deve ser projetado para atender à faixa audível do sistema auditivo humano, compreendida aproximadamente entre $20\,\mathrm{Hz}$ e $20\,\mathrm{kHz}$. Entretanto, como os modelos *SPICE* não consideram alguns efeitos práticos que podem afetar a resposta em frequência — como capacitâncias parasitas e limitações do arranjo físico — é recomendável que, durante o projeto, os limites inferior e superior de frequência sejam especificados com uma margem adicional de segurança. Com base nos critérios discutidos, foram determinados os requisitos de projeto, que são apresentados na Tabela 1.

TABELA I
PARÂMETROS DE DESEMPENHO ESPERADOS PARA O PRÉ-AMPLIFICADOR

Parâmetro	Valor Esperado
Ganho de tensão (A_v)	40 dB
Impedância de entrada (Z_{in})	$\geq 100\ k\Omega$
Impedância de saída (Z_{out})	$\leq 100\ \Omega$
Resposta em frequência	10 Hz – 40 kHz
Distorção Harmônica Total (THD)	$\leq 0,1\%$
Slew Rate (SR)	$\geq 1,41\ V/\mu s$

IV. Resultados

A. Escolha da arquitetura final

A escolha da arquitetura foi idealizada com o circuito possuindo três estágios - que serão descritos nesta seção - Um menor número de estágios possibilita melhor controle da distorção, reduzindo a necessidade da realimentação global [6].

a.1) 1° Estágio : Par diferencial

O primeiro estágio, apresentado na Figura 1, utiliza um par diferencial com transistores *JFET*, responsável por fornecer alto ganho inicial, boa rejeição de modo comum e maior linearidade ao sinal de entrada.

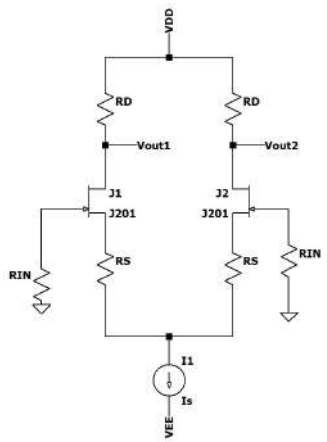


Fig. 1. Primeiro estágio: Par diferencial com *JFET*

A queda de tensão no resistor R_D deve permitir boa excursão de sinal, manter o dispositivo dentro de seus limites de operação e possibilitar ganhos elevados sem levar o estágio à saturação, além de garantir bom desempenho para os estágios subsequentes. Dessa forma, foi adotada uma queda de tensão de 15V.

a.2) Buffer

Para evitar o carregamento do primeiro estágio pelo segundo, utiliza-se um circuito de casamento, ou *"buffer"*, conforme mostrado na Figura 2. Devido à sua alta impedância de entrada, esse estágio permite que o sinal proveniente do primeiro estágio seja transferido ao estágio seguinte praticamente sem atenuação.

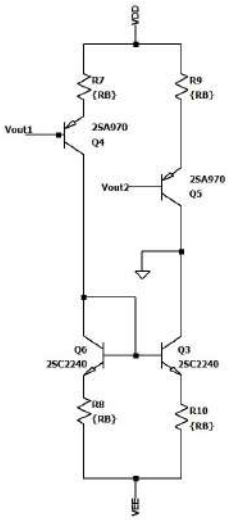


Fig. 2. Bipolar com coletor comum e espelho de corrente

a.3) 2° Estágio: Célula de ganho

O segundo estágio é responsável por complementar o ganho inicial fornecido pelo primeiro estágio.

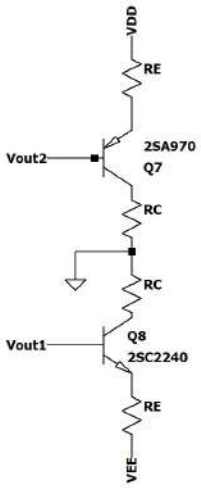


Fig. 3. Célula de ganho com TBJ operando em contra fase

Na Figura 3, observa-se que os transistores Q7 e Q8 estão em configuração emissor comum e operam em contrafase, acompanhando os dois ramos do par diferencial.

a.4) 3° Estágio: Push-Pull.

A Figura 4 apresenta o estágio de saída na configuração *push-pull*. Os transistores Q1 e Q2 são responsáveis por minimizar a distorção de *crossover*, sendo polarizados pelo multiplicador de VBE (Q7) em conjunto com os resistores R1_VBE e R2_VBE.

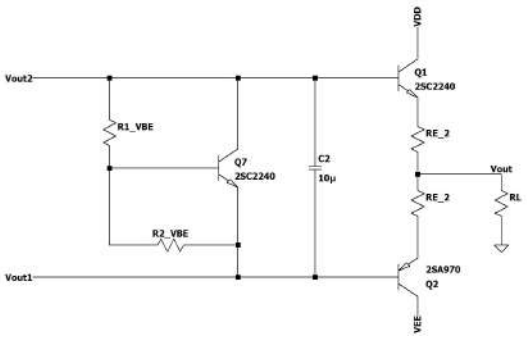


Fig. 4. Configuração *Push- Pull*.

B. Escolha dos componentes comerciais

A seleção dos componentes para o primeiro estágio seguiu os critérios descritos em IV-A1. Assim, o *JFET* J201 foi escolhido por suas características de baixo ruído e baixas capacitâncias parasitas. Para os estágios seguintes, foram utilizados os transistores 2SC2240 e 2SA970, um par complementar amplamente empregado em aplicações de áudio devido ao alto ganho de corrente e boa capacidade de operação em tensões mais elevadas.

C. Otimização paramétrica dos estágios para determinação de Slew-rate, THD e resposta em frequência

O ganho total do circuito é determinado pela equação

$$A_{v_t} = A_{v_1} A_{v_2} A_{v_3} \tag{3}$$

onde A_{v_t} é o ganho total e A_{v_1}, A_{v_2} e A_{v_3} são os ganhos do primeiro, segundo e terceiro estágios, respectivamente. Como A_{v_3} é aproximadamente igual a 1, o ganho global desejado de 100 está distribuído entre os dois primeiros estágios.

c.1) 1° Estágio

Para o primeiro estágio, cujo circuito é apresentado na Figura 1, os parâmetros variados foram a corrente de polarização I_s e os resistores R_D e R_S, que determinam o ganho do par diferencial.

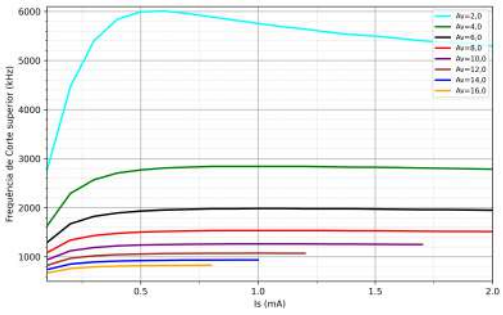


Fig. 5. Frequência de corte superior em função da corrente de polarização I_s, para diversos ganhos

A Figura 5 mostra que, quanto maior o ganho, menor é a resposta em frequência. Esse comportamento ocorre principalmente devido ao efeito Miller, no qual a capacitância de entrada — interna ao próprio *JFET* — é multiplicada pelo ganho do estágio.

Outro aspecto observado é que, para ganhos mais elevados, a extensão percorrida no eixo x diminui. Isso indica que, nesses casos, a variação da corrente I_s é menor. Tal comportamento

ocorre porque, à medida que o ganho aumenta, o valor do resistor R_S diminui, aproximando-se de um limite físico de redução.

Além disso, a diminuição de R_S reduz a realimentação local, fazendo com que o ganho real do circuito se afaste do valor projetado. Para garantir que o ganho calculado estivesse próximo do ganho obtido em simulação, foi estabelecida no algoritmo uma precisão de 97%. Dessa forma, o ganho foi inicialmente calculado e posteriormente medido no LTspice; caso o erro fosse inferior a 3%, o ponto era considerado válido e incluído nas análises.

Observa-se, entretanto, que nenhuma das curvas apresenta frequência de corte inferior ao mínimo especificado de $40\,kHz$. Assim, todos os pontos analisados atendem a esse critério de desempenho.

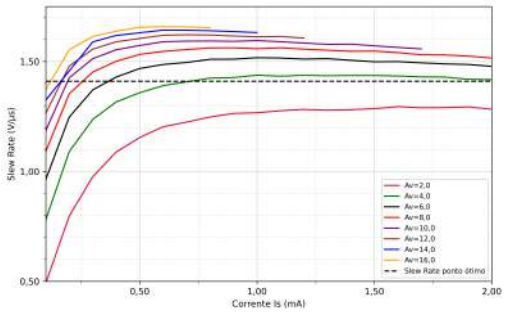


Fig. 6. SR vs I_s, para diversos ganhos

Na Figura 6 observa-se a relação entre I_s, A_v e o SR. O critério adotado foi o apresentado na Equação 2, considerando que, ao conectar os demais estágios ao primeiro, a capacitância associada tende a aumentar, isso pode reduzir o SR. Assim, o ponto de corte escolhido foi de $1{,}41\,V/\mu s$ — conforme apresentado na Tabela I — garantindo maior margem de segurança. Dessa forma, um ganho mínimo de 6 é necessário para atender a esse critério.

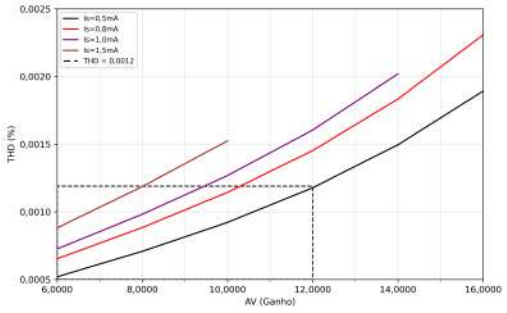


Fig. 7. THD vs Av, para diversas correntes de polarização I_s

A Figura 7 apresenta a relação entre I_s e a THD para diferentes valores de ganho. A análise mostra que menores valores de I_s resultam em níveis mais baixos de distorção. Considerando que a adição de estágios tende a introduzir novos comportamentos não lineares ao circuito, foi adotado um valor de distorção de $0{,}0012\%$, significativamente inferior ao limite total estabelecido de $0{,}1\%$.

Sendo assim, a faixa de ganho para atender os requisitos descritos é de 6 a 12 e a corrente I_S foi definida como $0{,}5\,mA$. Esses pontos serão considerados para a análise dos próximos estágios.

c.2) BUFFER

O acréscimo do *buffer* melhora o circuito conforme descrito na Seção IV-A2; entretanto, implica na adição de novas fontes de distorção ao sistema. Assim, para a determinação dos resistores do buffer, o circuito deve ser otimizado por meio da variação da corrente que circula por esses resistores.

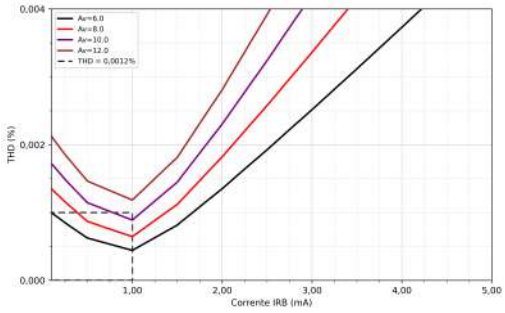


Fig. 8. THD vs IRB, para os pontos selecionados

Conforme observado na Figura 8, a corrente de $1{,}0\,mA$ apresenta os menores níveis de distorção. Esse resultado permite definir o valor dos resistores R_B, apresentados na Figura 2.

c.3) 2° ESTÁGIO

Para o segundo estágio, a variação da corrente de coletor é obtida por meio da alteração dos resistores R_E e R_C (Figura 3). O ganho desse estágio é calculado de forma que o produto entre A_{v_1} e A_{v_2} resulte em 100 ($40\,dB$).

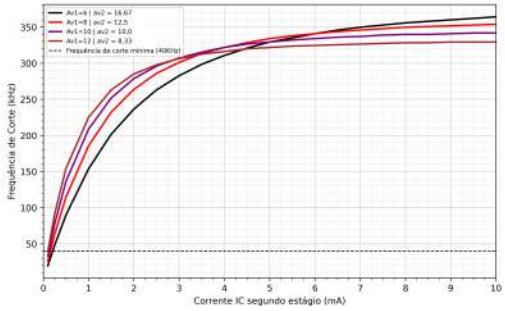


Fig. 9. Frequência de corte superior vs IC, para os pontos selecionados

A Figura 9 mostra que a resposta em frequência permanece acima da margem especificada para qualquer corrente de coletor superior a $0{,}5\,mA$.

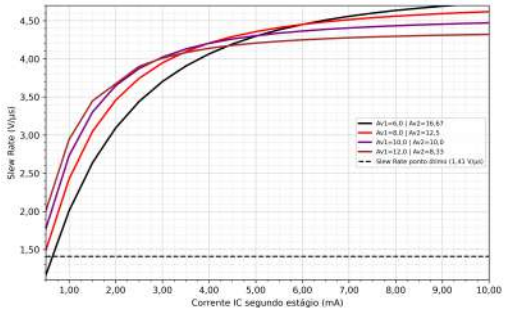


Fig. 10. SR vs IC, para os pontos selecionados

Embora fosse esperado que o SR pudesse diminuir com o acréscimo de novos componentes em relação ao primeiro

estágio, o comportamento observado foi o oposto. Isso se explica pelo aumento da amplitude do sinal ao longo da cadeia de amplificação. Embora o acréscimo de estágios introduza capacitâncias adicionais ao circuito, os estágios seguintes também são capazes de fornecer correntes maiores para carregar e descarregar essas capacitâncias. Dessa forma, a taxa de variação máxima do sinal acaba sendo determinada principalmente pela corrente disponível nos estágios posteriores, resultando em um aumento do SR observado na saída do circuito. Sendo assim, é visto na Figura 10 que o SR mínimo é atendido para todas as curvas com a IC maior que $1{,}0\,mA$.

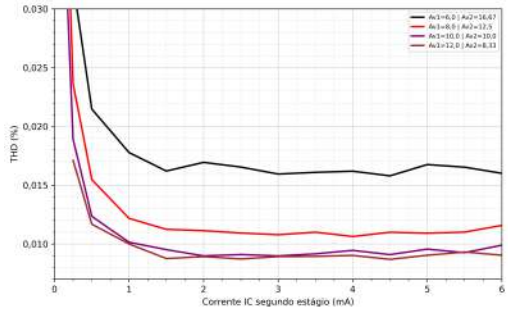


Fig. 11. THD vs IC, para os pontos escolhidos

O comportamento observado na Figura 11 é a diminuição da THD conforme o aumento da corrente, até atingir um regime estável. A explicação está relacionada principalmente ao ponto de operação dos transistores: para correntes de polarização menores, o dispositivo opera mais próximo da região ôhmica; à medida que a corrente aumenta, a operação tende à região mais linear.

O aumento da THD — em média cerca de 10 vezes — ao comparar as Figuras 8 e 11 é explicado pelo aumento do ganho em proporção semelhante, evidenciando o impacto da adição do segundo estágio.

Mediante as análises realizadas e priorizando a menor distorção possível, os ganhos A_{v_1} e A_{v_2} de 12 e 8,33, respectivamente, com $I_C = 4{,}5\,mA$, apresentam a menor distorção e resposta satisfatória para todos os critérios analisados.

c.4) 3° Estágio

Com as correntes de polarização e os ganhos dos estágios anteriores devidamente analisados e escolhidos, a otimização do último estágio é feita escolhendo a corrente quiescente de RE_2 através da variação de R1_VBE, mostrados na Figura 4.

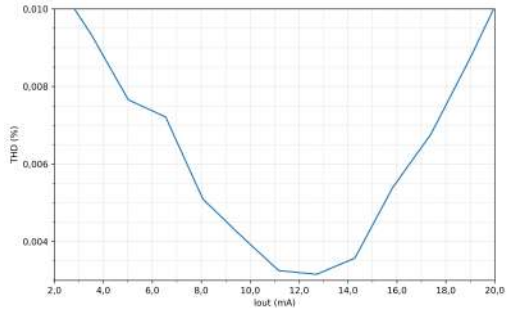


Fig. 12. THD vs Corrente Quiescente do terceiro estágio

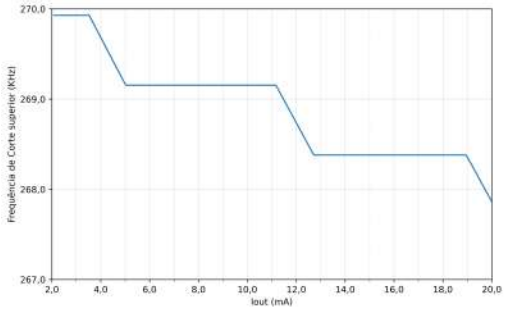


Fig. 13. Frequência de corte superior vs Corrente Quiescente do terceiro estágio

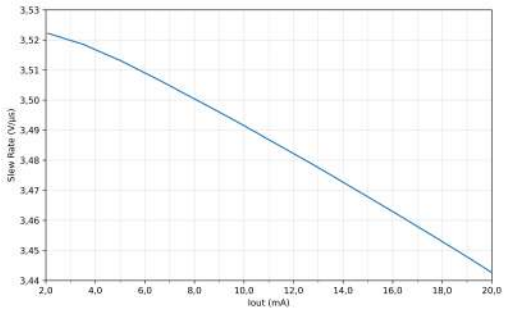


Fig. 14. SR vs Corrente Quiescente do terceiro estágio

A Figura 12 apresenta a variação da distorção harmônica total em função da corrente de polarização do estágio *push-pull*. Observa-se um comportamento aproximadamente parabólico, característico de estágios classe AB.

Para correntes de polarização reduzidas, a distorção é dominada pelo efeito de *crossover*, resultante da região morta existente durante a transição entre os dispositivos de saída. À medida que a corrente de polarização aumenta, essa região é progressivamente reduzida, levando a uma diminuição da distorção.

Entretanto, para correntes mais elevadas, a distorção volta a crescer devido às não linearidades intrínsecas dos transistores bipolares, associadas à relação exponencial entre corrente de coletor e tensão base-emissor. Dessa forma, observa-se a existência de um ponto ótimo de polarização em aproximadamente 13 mA, no qual a distorção é minimizada.

A Figura 14 mostra que, para o ponto escolhido, o SR permanece dentro do valor esperado. O comportamento levemente decrescente ocorre porque, ao aumentar a corrente quiescente nos resistores do último estágio, além da diminuição de V_{CE}, há também uma pequena modulação no ganho do estágio, tornando-o ligeiramente menor. Entretanto, essa variação é reduzida e, portanto, o impacto sobre a análise do SR é mínimo.

A Figura 13 apresenta a redução da resposta em frequência com a variação de I_{OUT}. Observa-se que, para o mesmo ganho apresentado na Figura 9, a frequência de corte situa-se em torno de 300 kHz. A interpretação desse comportamento é semelhante à apresentada para o SR.

Essa redução na resposta em frequência constitui, por outro lado, uma possível explicação para a diminuição da THD observada entre o segundo e o terceiro estágio, uma vez que harmônicas de ordem mais elevada tendem a ser mais fortemente atenuadas. O circuito completo com os valores comerciais é apresentado na Figura 15.

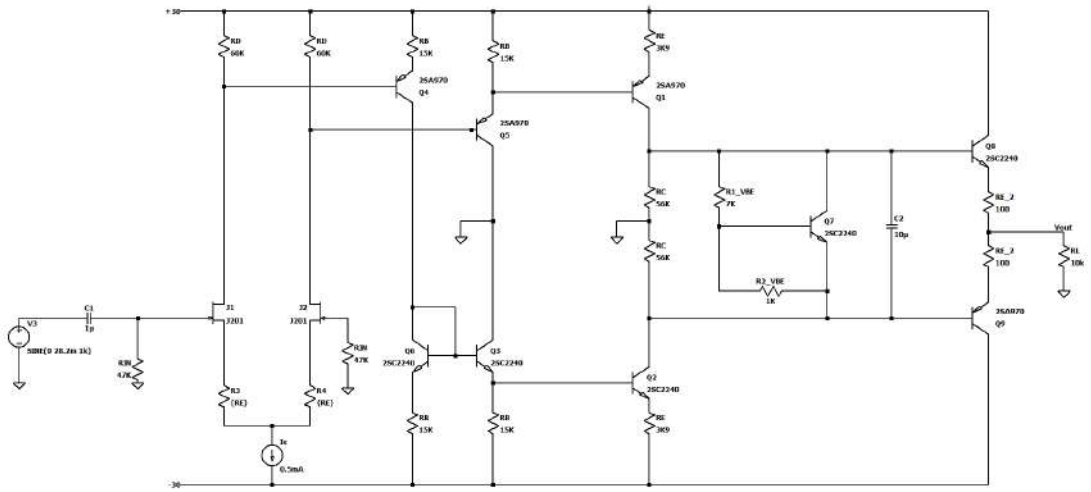


Fig. 15. Circuito completo com valores comerciais

V. Conclusão

Os resultados obtidos demonstraram que a abordagem empregada é capaz de atingir desempenho comparável ao de pré-amplificadores comerciais de alta qualidade. O circuito projetado apresentou distorção harmônica total de aproximadamente 0,0032%, slew rate de 3,48 V/μs, ganho de 39,88 dB e frequência de corte superior de 268,5 kHz, atendendo aos critérios de desempenho estabelecidos no projeto. O uso de um JFET no primeiro estágio apresentou resultados extremamente satisfatórios para o projeto. Portanto, o estudo de uma topologia completa utilizando esse componente pode ser aprofundado. Além disso, observou-se que o maior ganho no primeiro estágio resultou em bons resultados, de modo que esse aspecto também pode ser explorado de forma mais detalhada nessa mesma topologia. Estudos sobre o comportamento térmico e de fase constituem etapas futuras do trabalho, assim como a montagem física do circuito para validação experimental dos critérios em bancada.

VI. Agradecimentos

Agradeço imensamente ao meu orientador, Prof. Dr. Carlos Alberto de Francisco, pela oportunidade de aprendizado, bem como por sua constante disponibilidade e compreensão ao longo do desenvolvimento deste trabalho.

Referências

[1] Douglas Self. *Audio Power Amplifier Design Handbook*. 5ª ed. Oxford: Newnes, 2009.

[2] AnalogPlanet. *Negative Feedback and High-End Audio: An Endless Debate*. URL: https://www.analogplanet.com/ (acesso em 10/06/2025).

[3] Nelson Pass. *Audio, Distortion and Feedback*. URL: http://www.passdiy.com/technical/audio-distortion-and-feedback (acesso em 10/06/2025).

[4] J. Moir. "High-quality audio amplifier design". Em: *Wireless World* 87.1534 (1981), pp. 32–39. ISSN: 0043-6039. DOI: https://doi.org/10.1049/ew.1981.0323.

[5] Dartzeel. *Sobre a filosofia de design da Dartzeel*. URL: https://www.dartzeel.com/ (acesso em 11/06/2025).

[6] Rafael Morelli Visintin, André Carlos Paim, Wesley Machado Molina e André Felipe Paim. "A design methodology for audio amplifiers with no global negative feedback: Proposal and validation". Em: *Semina: Ciências Exatas e Tecnológicas* 43.2 (2022), pp. 105–120. ISSN: 1679-0375. DOI: https://doi.org/10.5433/1679-0375.2022v43n2p105. URL: https://ojs.uel.br/revistas/uel/index.php/semexatas/article/view/46530.

[7] Áudio e Acústica Grupo de Pesquisa em Eletrônica. *Como fazer simulações automatizadas pelo Python no LTSPice*. YouTube. 2024. URL: https://www.youtube.com/watch?v=zt3BgBCHfPE&t=938s (acesso em 15/03/2026).

[8] Lucas Alan da Conceição Batista. *OPTIMIZATION OF A MULTISTAGE AUDIO PREAMPLIFIER WITHOUT G-NFB USING PYTHON AND LTSPICE*. GitHub. 2024. URL: https://github.com/LucasAlan03/-OPTIMIZATION-OF-A-MULTISTAGE-AUDIO-PREAMPLIFIER-WITHOUT-G-NFB-USING-PYTHON-AND-LTSPICE (acesso em 15/03/2026).

[9] I. Sinclair e et al. *Audio Engineering: know it all*. Oxford: Elsevier, 2008.

[10] B. Duncan. *High performance audio power amplifiers for music performance and reproduction*. Boston: Newnes, 1997.

[11] John Atkinson. *Ayre Acoustics KX-8 Preamplifier Measurements*. Stereophile. 2024. URL: https://www.stereophile.com/content/ayre-acoustics-kx-8-preamplifier-measurements (acesso em 27/06/2025).

[12] FOSI. *Pré-amplificador Phono Toca Disco Fosi Box X1*. 2025. URL: https://www.primeinhp.com.br/pre-amplificador-phono-toca-disco-fosi-box-x1-capsula-mm (acesso em 27/06/2025).

[13] LOGOS. *Pré-amplificador cerâmico*. 2025. URL: https://www.amplificadoreslogos.com.br/pre-amplificadores/pre-ceramico (acesso em 27/06/2025).

Desenvolvimento de um tubo de impedância para caracterização acústica de materiais

Gabriel Farias de Oliveira Junior, Carlos Alberto De Francisco
Universidade Federal de São Carlos - UFSCar
gabrieljunior@estudante.ufscar.br, cfrancisco@ufscar.br

***Abstract*—In the field of acoustics, parameters such as absorption and reflection coefficients are fundamental for selecting porous, fibrous, or micro-perforated materials. In this context, the impedance tube emerges to meet the need for measuring such values. Using derivations from current standards, it was possible to obtain graphs demonstrating the effects of each measurement method and arrangement for high and low frequencies. Upon analyzing the results, it was noted that instead of doubling a sample's thickness to increase absorption, adding an air gap between it and the reference increases the coefficient and reduces material costs. Therefore, despite being a low-cost device, the results obtained were satisfactory and consistent with the theory, proving to be an excellent alternative to professional or more expensive measurement equipment.**

***Index Terms*—Air gap, Impedante tube, Porous materials, Sound absorption.**

***Resumo*—No campo da acústica, parâmetros como os coeficientes de absorção e reflexão são fundamentais para a escolha de materiais porosos, fibrosos ou microperfurados. Nesse contexto, o tubo de impedância surge para suprir a necessidade de medir tais valores. Utilizando as deduções das normas vigentes, foi possível obter gráficos que demonstram os efeitos de cada método de medição e arranjo para altas e baixas frequências. Ao analisar os resultados, notou-se que, em vez de dobrar a espessura de uma amostra para elevar a absorção, adicionar um espaço de ar entre ela e o referencial aumenta o coeficiente e reduz custos com material. Portanto, apesar de ser um dispositivo de baixo custo, foram obtidos resultados satisfatórios e condizentes com a teoria, mostrando-se uma excelente alternativa a equipamentos de medição profissionais ou mais caros.**

***Palavras chave*—Absorção acústica, Espaço de ar, Materiais porosos, Tubo de impedância.**

I. Introdução

No estudo da acústica em ambientes, é de suma importância conhecer os parâmetros acústicos dos materiais utilizados no tratamento acústico, como materiais porosos, fibrosos, microperfurados ou a combinação entre eles, com o objetivo de se alcançar o controle acústico desejado.

Dentre os principais parâmetros que caracterizam esses materiais, destacam-se o coeficiente de absorção, o coeficiente de reflexão e a impedância acústica. Esses parâmetros são fundamentais para a escolha do material mais adequado em diferentes aplicações, uma vez que estão diretamente relacionados com a forma como a onda sonora interage com o meio.

Uma das formas mais utilizadas para a determinação desses parâmetros é o uso do tubo de impedância, conforme descrito nas normas ISO 10534-2 [1] e ASTM E-1050 [2]. Essas normas estabelecem critérios construtivos e operacionais que influenciam diretamente na faixa de frequência analisada e na precisão das medições.

Dentre os métodos existentes, o método dos dois microfones se destaca por sua relativa simplicidade de implementação e boa precisão experimental. No entanto, ele apresenta algumas limitações, como a restrição da faixa de frequência, a necessidade de incidência normal da onda sonora e a dependência da qualidade da amostra analisada.

Apesar dessas limitações, estudos mostram que é possível obter resultados satisfatórios utilizando equipamentos de baixo custo, como microfones e sistemas de aquisição simples [3].

Dessa forma, o presente trabalho tem como objetivo analisar e implementar a medição de parâmetros acústicos utilizando o método dos dois microfones em tubo de impedância, avaliando sua viabilidade experimental, limitações e precisão, especialmente considerando o uso de componentes de baixo custo.

II. Fundamentação Teórica

Os materiais utilizados no tratamento acústico podem apresentar diferentes mecanismos de atuação, dependendo de sua estrutura física.

Para os materiais porosos ou fibrosos, sua funcionalidade ocorre devido às irregularidades estruturais, que dificultam a propagação da onda sonora em seu interior, promovendo a dissipação de energia ao longo do caminho [4]. Um exemplo de material fibroso pode ser visto na Figura 1.



Fig. 1. Exemplo de material poroso (Autoral).

Já os materiais microperfurados têm seu funcionamento baseado no dimensionamento do diâmetro dos furos e na fração da área perfurada da placa [5]. Um exemplo teórico pode ser observado na Figura 2.

Além disso, é possível combinar placas microperfuradas com materiais porosos, formando sistemas híbridos. Essa combinação pode proteger o material poroso contra poeiras que podem comprometer seu funcionamento, além de alterar suas características de absorção, dependendo de parâmetros como a distância entre as camadas, o diâmetro dos furos e a área de contato entre os materiais [6, 7].

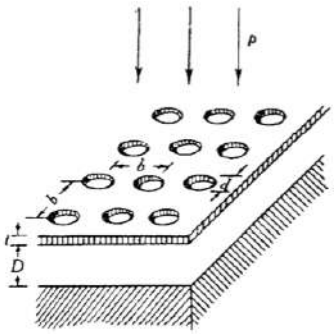


Fig. 2. Exemplo de placa microperfurada (Adaptado de [5]).

Para caracterizar esses materiais, alguns parâmetros acústicos são fundamentais. O coeficiente de absorção representa a eficiência do material em absorver a onda sonora, variando entre 0, onde não há absorção, e 1, onde ocorre absorção máxima. Esse coeficiente depende diretamente da frequência da onda incidente, podendo variar significativamente conforme o material e sua espessura.

A impedância acústica, por sua vez, é análoga à impedância elétrica, sendo uma medida da oposição do material à propagação da onda sonora, relacionando a pressão acústica com a velocidade de partícula [8].

Uma das principais formas de medir esses parâmetros é por meio do tubo de impedância, cujo funcionamento é descrito em normas como a ISO 10534-2 [1] e a ASTM E-1050 [2]. Essas normas estabelecem critérios que influenciam diretamente na faixa de frequência analisada, como as dimensões do tubo e o posicionamento dos sensores.

Dentre os métodos utilizados, destaca-se o método dos dois microfones, no qual os sensores devem estar posicionados rente ao tubo, permitindo a medição das variações de pressão acústica ao longo do seu comprimento. A partir dessas medições, é possível determinar o coeficiente de reflexão e, consequentemente, os demais parâmetros acústicos.

Apesar de ser amplamente utilizado, esse método apresenta algumas limitações, como a necessidade de incidência normal da onda sonora, restrições na faixa de frequência e dependência das condições da amostra analisada [8].

III. Metodologia

A. Materiais

Para a fabricação do tubo, foi utilizado um cano de PVC de 10 cm de diâmetro devido ao baixo custo e seu padrão já vir com o interior liso para refletir as ondas.

Os microfones são do tipo condensador uni-direcional de eletreto, a placa de aquisição trabalha com uma taxa de 48kHz apresentando uma entrada de linha, para ocorrer a divisão entre canal esquerdo e direito, essencial na hora dos cálculos com os sinais.

Para gerar o sinal de excitação, é utilizado o computador que envia o sinal via placa de aquisição para um amplificador que, posteriormente, manda o sinal final para o alto-falante acoplado à uma das extremidades do tubo.

Na realização dos cálculos, é utilizado um computador, em que são enviados e recebidos todos os sinais para calcular os valores dos coeficientes, como de reflexão ou absorção. Para esses cálculos, foi desenvolvido em código utilizando a linguagem Python e suas bibliotecas para fazer toda a análise dos dados.

Por fim, é de suma importância também, analisar as condições do local, como temperatura, umidade e pressão. Assim, foi utilizado um microcontrolador STM32 ligado a um sensor BME280 para esse propósito, a fim de se chegar a um valor de velocidade de som mais fiel possível ao do ambiente e minimizar o máximo possível os erros.

B. Método

O método para se medir no tubo de impedância foi baseado na dedução a partir das ondas refletidas e incidentes [9], onde, é necessário primeiro, calcular a Função Resposta em Frequência (FRF), sendo ela, a variável $H(f)$, que representa a razão entre os espectros das pressões sonoras adquiridas pelos microfones 1, P_1, e 2, P_2, como é mostrado na Equação 1

$$H_{12} = \frac{P_2}{P_1} \tag{1}$$

Para capturar os sinais dos microfones, é enviado pelo auto falante um ruído branco gaussiano, o qual abrange todas as frequências com potências iguais em todas elas, além de ser simples de ser realizado via código, tanto no computador quanto em microcontroladores.

Os sinais adquiridos pelos microfones em conjunto com a placa de aquisição são enviados para o computador para a realização dos cálculos matemáticos para se chegar aos valores dos coeficientes.

O princípio dessa medição, é baseado no espectro pressão de onda incidente P_I e no espectro da pressão de onda refletida P_R no qual ambas estão no plano de referência, assim, temos as Equações 2 e 3 que demonstram as equações da onda incidente e refletida respectivamente.

$$P_I = P_I^{'} e^{jk_0 x} \tag{2}$$

$$P_R = P_R^{'} e^{-jk_0 x} \tag{3}$$

Com a variável k_0 sendo a constante de propagação da onda, x a distância do microfone em relação ao plano de referência, e $P_I^{'}$ e $P_R^{'}$ os espectros das ondas incidente e refletida no plano de referência.

Esse plano de referência pode ser a terminação própria do tubo ou o começo da amostra a ser medida pelo tubo de impedância.

Assim, definida a equação de referência das ondas, podemos calcular P_1 e P_2, desse modo, são obtidas as Equações 4 e 5, que demonstram a soma das pressões incidentes e refletidas nos microfones 1 e 2 respectivamente.

$$P_1 = P_I^{'} e^{jk_0 x_1} + P_R^{'} e^{-jk_0 x_1} \tag{4}$$

$$P_2 = P_I^{'} e^{jk_0 x_2} + P_R^{'} e^{-jk_0 x_2} \tag{5}$$

Com x_1, sendo a distância entre o microfone mais distante em relação ao plano de referência e x_2 a distância do microfone mais próximo do plano de referência.

Sabendo que $P_R^{'} = R \cdot P_I^{'}$, com R sendo o coeficiente de reflexão, podemos substituir as Equações 4 e 5 na Equação 1 e, reescrevendo, é possível chegar na Equação 6

$$H_{12} = \frac{P_2}{P_1} = \frac{e^{jk_0x_2} + R \cdot e^{-jk_0x_2}}{e^{jk_0x_1} + R \cdot e^{-jk_0x_1}} \tag{6}$$

Com o objetivo de se conseguir o coeficiente de reflexão, pode-se reorganizar novamente a Equação 6 e chegar à Equação 7, que demonstra o cálculo para obter o coeficiente de reflexão, que pode ser implementado no meio computacional, onde a variável x_{12} representa a distância entre os dois microfones.

$$R = \frac{H_{12} - e^{-jk_0x_{12}}}{e^{jk_0x_{12}} - H_{12}} e^{2jk_0x_1} \tag{7}$$

Com o valor do coeficiente de reflexão, é possível expressar o coeficiente de absorção do material a ser estudado, assim, temos a Equação 8 para obtê-lo.

$$\alpha = 1 - |r|^2 \tag{8}$$

Por fim, devido aos microfones não serem igualmente calibrados, é necessário utilizar um meio de correção no momento das medições para corrigir esse desvio de fase e amplitude que pode ocorrer no cálculo da Transformada de Fourier [2]. Para isso, é realizada a primeira medição (H_1) com o tubo normal montado, após isso, é realizada uma segunda medição (H_2) com os microfones trocados de posição, com isso é possível calcular o coeficiente de correção Hc como mostra a Equação 9

$$H_C = \sqrt{H_1 \cdot H_2} \tag{9}$$

Feito o cálculo para o fator de correção, é usado o tubo de impedância mais uma vez nesse ciclo para fazer a terceira e última medição $(\bar{H})$ sendo então aplicada a correção como demonstra a Equação 10 para calcular o H_{12}

$$H_{12} = \frac{\bar{H}}{H_C} \tag{10}$$

Agora, com esse valor corrigido é possível utilizar as Equações 8 e 7 para calcular os coeficientes desejados, seja via microcontrolador ou linguagens de programação como o Python.

IV. Resultados

Após o estudo da faixa de frequência que se deseja medir, foi possível dimensionar o tubo de impedância. Ele foi feito de PVC, com um diâmetro de 10 cm e comprimento de aproximadamente 73,45 cm, feito furos para colocar os microfones a uma distância x_1 de 12,2 cm e x_2 de 5,2 cm, o que resulta numa distância x_{12} entre eles de 7 cm. A construção do tubo pode ser vista na Figura 3

O terceiro microfone, embora não seja utilizado na medição, permanece instalado para vedar o orifício não utilizado, minimizando interferências externas no tubo melhorando as medições feitas.

Para dar mais segurança no posicionamento dos microfones, foram fabricados suportes para posicioná-los, que podem ser retirados e rotacionados para fazer o cálculo do fator de correção necessário. Além de suporte no próprio tubo para impedir pequenas vibrações que possam ocorrer e desestabilizar o sistema. Existem dois suportes diferentes, que podem ser alterados dependendo de qual medição se deseja fazer, sendo o de cima para medições em altas frequências e o de baixo para medições em baixas frequências como mostra a Figura 4

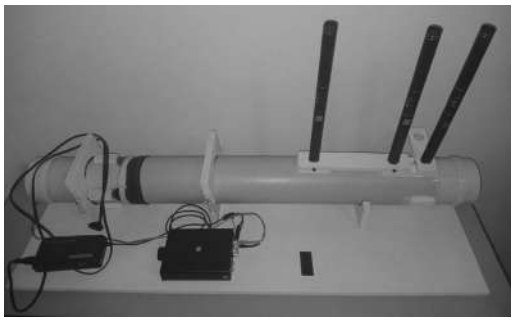


Fig. 3. Tubo de Impedância (Autoral).



Fig. 4. Suportes para ajustar os microfones (Autoral).

A amostra é posicionada no tubo utilizando dois suportes diferentes dependendo do tipo de medição. Na Figura 5, à esquerda é mostrado um suporte para uma amostra com espessura de 45 mm, enquanto a da direita é para suportar uma amostra de 90 mm, ou ainda, fazer medidas com um espaço de ar de 45 mm entre a terminação do suporte e a amostra, o que desenvolve outro tipo de medição e altera a curva de absorção. Ambos os suportes possuem como terminações uma chapa de aço devidamente lixada para melhorar a reflexão da onda dentro do tubo.



Fig. 5. Suportes para alocar as amostras (Autoral).

Além disso, para minimizar os erros que possam ocorrer durante as medições, foram alterados os microfones para deixá-los mais rentes ao tubo, retirando a capa protetora da ponta, deslocando a cápsula mais para fora e colocando uma espuma para suporte, além de melhorar a vedação do som, fator muito importante para as medições utilizando o tubo.

Devido às limitações teórica do tubo, foram utilizadas as Equações 11 e 12 [1] para descobrir as frequências máxima e mínima que o tubo é capaz de medir de maneira satisfatória, assim, pode-se estabelecer esse limite nos gráficos, eliminando as faixas de frequências que não importam nesse teste.

$$f_{max} = 0,58 \cdot \frac{c_0}{d} = 0,58 \cdot \frac{343}{0,1} = 1.989,4Hz \quad (11)$$

$$f_{min} = 0,05 \cdot \frac{c_0}{s} = 0,05 \cdot \frac{343}{0,07} = 245Hz \quad (12)$$

Com c_0 sendo a velocidade do som, d o diâmetro interno do tubo em metros e s a distância entre os microfones. Assim, todos os gráficos foram gerados respeitando esse limite e colocando uma margem de segurança para assegurar maior confiabilidade nos resultados, dessa forma, o limite inferior foi considerado 300 Hz e o superior de 1800 Hz.

Para melhorar a precisão das medidas no meio computacional, foram realizadas várias medições dos valores de distância entre os microfones e diâmetro do tubo, assim, o valor final encontrado para eles, devido às imperfeições de fabricação, foram de 12,2 cm e 6,935 cm respectivamente.

Antes de efetuar as medições, foi realizado um teste com o tubo completamente vazio para verificar a resposta do sistema, assim, foi gerado o gráfico da Figura 6.

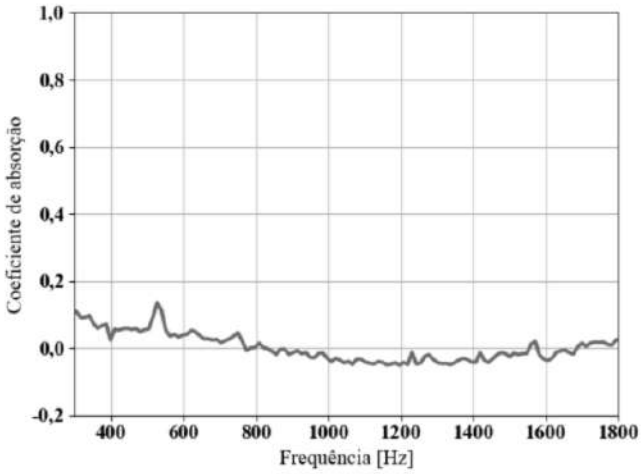


Fig. 6. Gráfico do sistema sem amostra em altas frequências (Autoral).

Para um tubo de impedância profissional, é desejável que o valor do coeficiente de absorção seja 0 em toda a faixa de frequência quando excitado sem amostra. Contudo, devido a esse ser um de baixo custo, a curva ficou abaixo de 0,1, o que é esperado nesse caso devido ao tubo ser de baixo custo e não estar completamente vedado, assim, o som pode escapar pelos furos, o que diminui um pouco a precisão.

Após todas as etapas do tubo feitas, foram realizadas as medições utilizando o arranjo para altas frequências e espuma porosa de 45 mm mostrada na Figura 1 como amostra, que gerou o gráfico da Figura 7.

Pelo gráfico, essa subida em linha reta é característica de materiais porosos, onde a curva aumenta até determinada frequência e após isso ocorre algumas ondulações até se

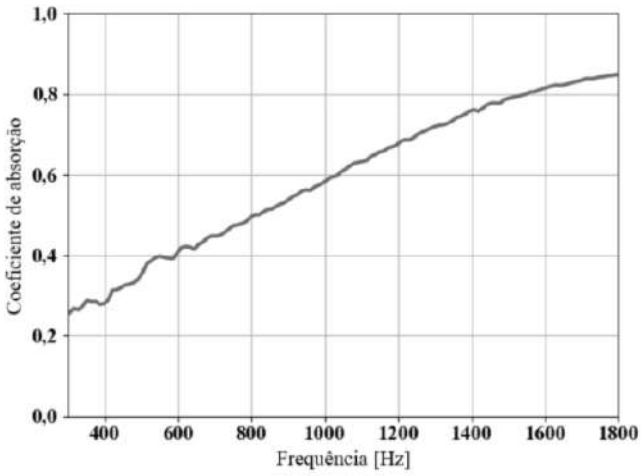


Fig. 7. Gráfico da medição com amostra de 45 mm em altas frequências(Autoral).

estabilizar em 1. Além disso, é possível notar que esse material não é muito eficiente nas baixas frequências, apresentando baixo coeficiente de absorção, contudo, conforme a frequência aumenta até os 1800 Hz, já se torna um bom material para uso prático.

Em seguida, foi realizada a medição para uma espessura da amostra com 90 mm, onde se espera que o coeficiente de absorção aumente significativamente em baixas frequências, o resultado pode ser visto no gráfico da Figura 8.

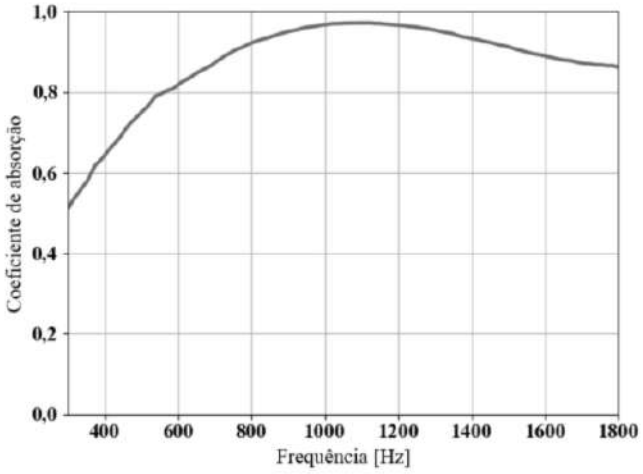


Fig. 8. Gráfico da medição com amostra de 90 mm em altas frequências (Autoral).

Nele, já é possível observar que o valor em baixas frequências é significativamente mais alto, chegando a mais que dobrar em algumas frequências, além disso, ele alcança o pico da primeira ondulação muito mais rápido que o anterior, assim, pode-se ver o feito de aumentar a espessura de uma amostra, onde ela agora já é uma boa opção para baixas frequências.

A seguir, foi feita uma medição deixando um espaço de ar de 45 mm entre a terminação do suporte e a amostra, também de 45 mm, para ver seu efeito no coeficiente de absorção, esse teste pode ser visto na Figura 9.

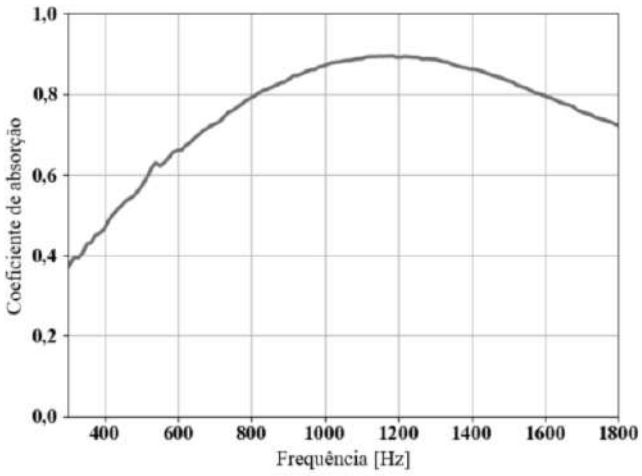


Fig. 9. Gráfico da medição com amostra de 45 mm com espaço de ar de 45 mm em altas frequências (Autoral).

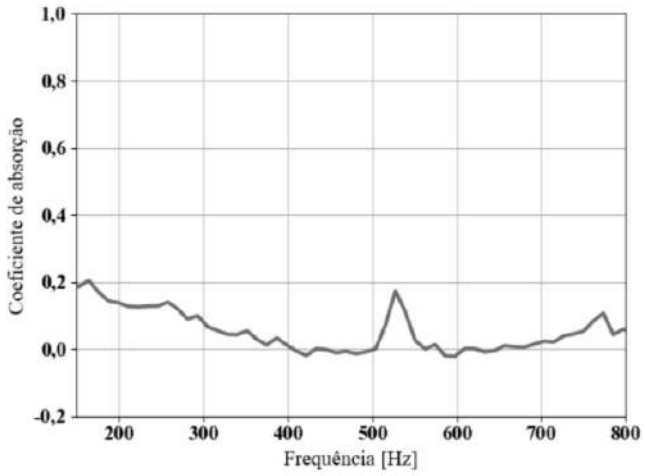


Fig. 10. Gráfico da medição sem amostra em baixas frequências (Autoral).

A presença de um espaço de ar atrás do material poroso resulta em uma melhora na absorção em baixas frequências, além de permitir que o coeficiente de absorção atinja valores elevados de forma mais rápida quando comparado à amostra de 45 mm sem espaçamento. Essa configuração possibilita um melhor desempenho acústico sem a necessidade de aumentar a quantidade de material, evitando assim acréscimos significativos de peso ou de custo. Dessa forma, o uso de uma camada de ar se apresenta como uma alternativa interessante para otimizar a absorção sonora com menor consumo de material.

Entretanto, observa-se que essa configuração também apresenta um comportamento mais pronunciado de pico em uma determinada faixa de frequência, que neste caso, ocorre aproximadamente em 1200 Hz. Após essa região de máxima absorção, há um decaimento mais acentuado do coeficiente, seguido posteriormente por uma leve tendência de estabilização.

Analogamente às medidas anteriores, foi trocado o arranjo dos microfones, utilizando o arranjo para medir baixas frequências para verificar o efeito do mesmo, onde foram utilizadas as Equações 13 e 14 [1], com espaçamento entre os microfones s de 17 centímetros.

$$f_{max2} = 0,45 \cdot \frac{c_0}{s} = 0,45 \cdot \frac{343}{0,17} = 907,94Hz \qquad (13)$$

$$f_{min2} = 0,05 \cdot \frac{c_0}{s} = 0,05 \cdot \frac{343}{0,17} = 100,88Hz \qquad (14)$$

Novamente, para deixar uma margem de segurança, os gráficos foram limitados entre 150 Hz e 800 Hz.

Antes das medições, foi realizado novamente o teste sem amostra para verificar a precisão do tubo nesta configuração, sendo obtido o gráfico da Figura 10.

Pela análise da curva, foi possível afirmar que esse arranjo ainda tem problemas de performance, que pode ser observado com valores mais altos de coeficiente de absorção em comparação ao arranjo anterior, principalmente abaixo dos 300 Hz e um pico em aproximadamente 525 Hz, o que pode significar possíveis erros nesses valores, contudo, há boa validação nas outras faixas, com valores muito próximos de 0.

Assim, utilizando novamente a amostra de 45 mm, foi possível obter o gráfico da Figura 11.

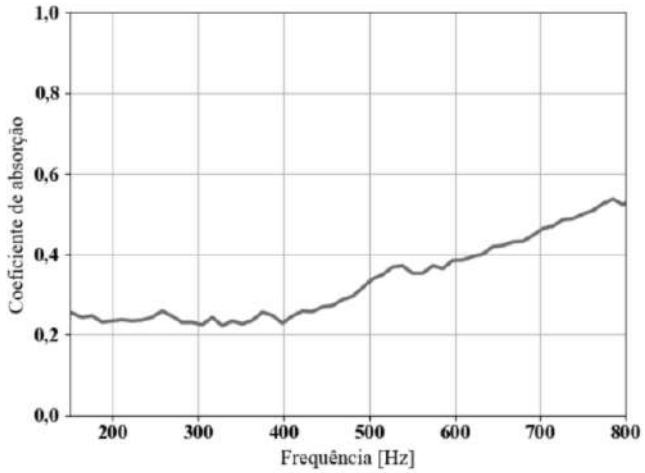


Fig. 11. Gráfico da medição com amostra de 45 mm em baixas frequências (Autoral).

Comparando-o com o gráfico da Figura 7 é possível notar semelhanças nas faixas que se interpolam, entre 350 e 800 Hz, o que evidencia que estão seguindo a mesma curva, apesar de ter pequenas variações pontuais devido, provavelmente, a problemas estruturais.

De forma análoga, foi obtido o gráfico da Figura 12, onde foi utilizada a amostra de 90 mm.

Novamente, apesar da presença de uma anomalia em aproximadamente 250Hz, foi possível obter uma curva que se assemelha à medida em altas frequências, com uma subida bem acentuada devido ao aumento de espessura, o que, apesar de não apresentar uma resposta um pouco pior no teste sem amostra, ainda sim, apresenta boa confiabilidade.

Por fim, foi feito o gráfico da Figura 13, onde é alocada a amostra de 45mm adicionando um espaço de ar também de 45mm entre a amostra e o final.

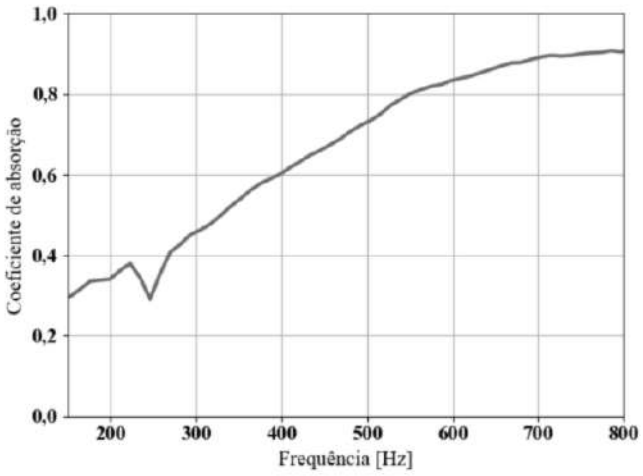


Fig. 12. Gráfico da medição com amostra de 90 mm em baixas frequências (Autoral).

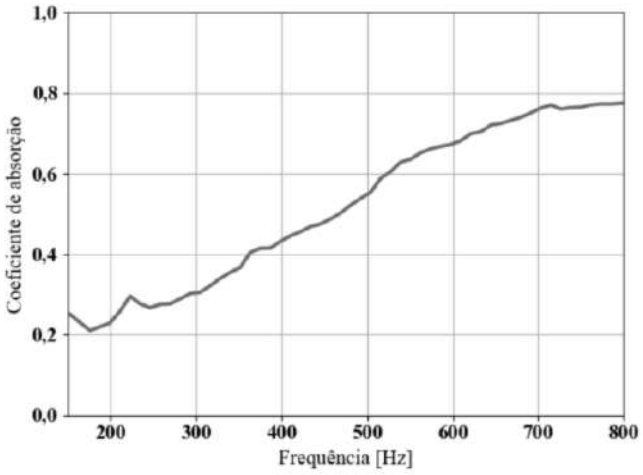


Fig. 13. Gráfico da medição com amostra de 45 mm com espaço de ar de 45 mm em baixas frequências (Autoral).

Nesse gráfico, é possível ver o mesmo efeito observado anteriormente mesmo para baixas frequências, onde o coeficiente de absorção é elevado ao custo de somente mais espaço, mostrando novamente o efeito e benefícios da implementação dessa configuração.

V. Conclusão

Com a construção do tubo de impedância de baixo custo, foi possível realizar medições do coeficiente de absorção sonora na faixa de 300 Hz a 1800 Hz com o arranjo de altas frequências e na faixa de 150 Hz a 800 Hz no arranjo de baixas frequências, obtendo resultados coerentes com o comportamento esperado para materiais porosos. Os testes realizados com o tubo vazio indicou que o sistema apresenta valores próximos de zero para o coeficiente de absorção, dentro da margem esperada para um sistema desse tipo, mostrando que o equipamento é capaz de fornecer medições satisfatórias, especialmente o arranjo de altas frequências.

Nas medições com as amostras de espuma, foi possível observar a influência da espessura do material no desempenho acústico. A amostra de 45 mm apresentou aumento gradual da absorção com a frequência, enquanto a amostra de 90 mm mostrou melhora significativa em baixas frequências, atingindo valores elevados de absorção de forma mais rápida. Já a configuração com 45 mm de material e 45 mm de espaço de ar também apresentou melhora na absorção em baixas frequências quando comparada à amostra de 45 mm sem espaçamento, mostrando que o uso de um air gap pode ser uma alternativa interessante para melhorar o desempenho acústico sem aumentar a quantidade de material.

De forma geral, os resultados obtidos indicam que o tubo desenvolvido é capaz de reproduzir tendências conhecidas no comportamento de materiais absorvedores, mostrando que sistemas de baixo custo podem ser utilizados para análises experimentais e estudos comparativos de absorção sonora.

Referências

[1] ISO. *ISO 10534-2: Acoustics. Determination of sound absorption coefficient and impedance in impedance tubes. Part 2: Transfer-function method. Norma ISO*. Disponível em: https://www.iso.org/standard/4472.html. 1998.

[2] *Standard Test Method for Impedance and Absorption of Acoustical Materials Using a Tube, Two Microphones, and a Digital Frequency Analysis System*. West Conshohocken, PA: ASTM International, 2024. URL: https://www.astm.org/e1050-24.html.

[3] Thaynan Oliveira, Paulo Mareze, Matheus Pereira, Sergio Aguirre, William Fonseca, Eric Brandao e Rogerio Pirk. "Sound absorption coefficient measurement based on the Transfer Function Method specified by the standard ISO 10534-2 using low-cost alternatives". Em: *22nd International Congress of Acoustics - ICA 2016*. Buenos Aires, Argentina, set. de 2016.

[4] Ovidiu VASILE. "The Analysis of Factors That Influence the Sound Absorption Coefficient of Porous Materials". Em: *Romanian Journal of Acoustics and Vibration* 11.2 (2014), pp. 105–108.

[5] DAH-YOU Maa. "Theory and Design of Microperforated Panel Sound-Absorbing Constructions". Em: *Scientia Sinica* 18.1 (1975), pp. 55–71. ISSN: 1000-0002.

[6] Davide BORELLI, Corrado SCHENONE e Ilaria PITTALUGA. "Analysis of sound absorption behaviour of polyester fibre material faced with perforated panels". Em: *The Journal of the Acoustical Society of America* 133 (2013), p. 3309. DOI: 10.1121/1.4805497.

[7] Iljae LEE, Ahmet SELAMET e Norman T. HUFF. "Acoustic impedance of perforations in contact with fibrous material". Em: *The Journal of the Acoustical Society of America* 119.5 (2006), pp. 2785–2797.

[8] Eric BRANDÃO. *Acústica de salas: projeto e modelagem*. 1ª ed. E-book. Disponível em: https://plataforma.bvirtual.com.br. Acesso em: 14 fev. 2026. São Paulo: Blucher, 2016.

[9] Rodolfo Thomazelli e Stelamaris Rolla Bertoli. "Tubo de impedância para baixas frequências: construção e validação". Em: *Acústica e Vibrações* 34.51 (dez. de 2019), pp. 59–71. DOI: 10.55753/aev.v34e51.81. URL: https://revista.acustica.org.br/acustica/article/view/aev51_tubo.

Otimização de interface de sistemas IoT aplicados em ambiente hospitalar

Rafael Gonçalves Magalhães, Pedro Henrique de Paiva Carvalho,
Luma Rissatti Borges do Prado
Instituto Nacional de Telecomunicações - Inatel
rafael.gm@gec.inatel.br, pedro.paivacarvalho@gec.inatel.br, luma.prado@inatel.br

***Abstract*—This work presents the functional expansion of an IoT-based monitoring system designed for hospital environments. The study focuses on improvements implemented in the web application layer, including automated PDF report generation for historical temperature data and a graphical spatial visualization of monitored equipment based on a digital floor plan. Additionally, visual interface adjustments were performed to enhance the organization and presentation of information in the monitoring module. The implemented features were qualitatively validated in a development environment, demonstrating proper functional behavior while preserving the original system architecture. The results indicate the potential for incremental software enhancements to increase the operational value of IoT-based monitoring systems in healthcare contexts.**

***Index Terms*—Graphical visualization, Hospital monitoring, Internet of Things, Web application**

***Resumo*—Este trabalho apresenta a ampliação funcional de um sistema de monitoramento baseado em IoT aplicado a ambientes hospitalares. O estudo concentra-se nas melhorias implementadas na camada de aplicação web, incluindo a geração automatizada de relatórios em PDF contendo dados históricos de temperatura e a implementação de uma visualização gráfica espacial dos equipamentos monitorados com base em uma planta digital. Além disso, foram realizados ajustes visuais na interface do núcleo de monitoramento, buscando aprimorar a organização e a apresentação das informações. As funcionalidades implementadas foram validadas qualitativamente em ambiente de desenvolvimento, demonstrando funcionamento adequado e preservando a arquitetura original do sistema. Os resultados evidenciam o potencial de melhorias incrementais na camada de software para ampliar o valor operacional de sistemas IoT aplicados ao monitoramento hospitalar.**

***Palavras chave*—Aplicação web, Internet das Coisas, Monitoramento hospitalar, Visualização gráfica**

I. Introdução

A incorporação de tecnologias baseadas em Internet das Coisas (IoT) no contexto hospitalar tem se mostrado uma estratégia fundamental para aprimorar a gestão de recursos, otimizar processos clínicos e elevar a segurança dos pacientes [1]. A capacidade de coletar, processar e analisar dados em tempo real permite o monitoramento contínuo de equipamentos e ambientes, além de possibilitar a automação de tarefas críticas e a redução de falhas operacionais. Estudos recentes destacam que a integração entre dispositivos IoT, sistemas em nuvem e aplicações *web* representa um dos pilares da chamada Saúde 4.0, promovendo maior rastreabilidade, interoperabilidade e suporte à tomada de decisão clínica [2].

No âmbito dessa tendência, o projeto IoT Santa Casa foi concebido em parceria do eHealth Innovation Center com o Hospital Santa Casa de Belo Horizonte, com o objetivo de desenvolver e validar uma arquitetura IoT voltada ao monitoramento e gerenciamento inteligente de recursos hospitalares. A proposta inicial contemplou três núcleos interconectados: Monitora, Localiza e Chamada. O núcleo Monitora destina-se ao acompanhamento contínuo da temperatura de refrigeradores utilizados para o armazenamento de medicamentos e insumos biológicos, garantindo a integridade dos materiais por meio de sensores de precisão e alertas automatizados. O núcleo Localiza implementa um sistema de rastreamento de equipamentos médico-hospitalares com base em comunicação *Bluetooth Low Energy* (BLE), permitindo a identificação e localização em tempo real de ativos dentro do ambiente clínico. Por fim, o núcleo Chamada busca aprimorar a comunicação entre pacientes no leito e a equipe médica, por meio de dispositivos conectados e interfaces *web* responsivas, otimizando o tempo de resposta e a coordenação das atividades assistenciais [3].

Apesar dos resultados positivos obtidos na fase inicial do projeto, observa-se na literatura que muitos sistemas IoT aplicados à saúde enfrentam desafios relacionados à escalabilidade, desempenho da aplicação *web* e capacidade analítica das plataformas de visualização [4, 5]. Em especial, a camada de *software* e visualização tende a se tornar um gargalo à medida que o volume de dados cresce e o número de dispositivos conectados aumenta, comprometendo a responsividade e a confiabilidade do sistema. Essa lacuna evidencia a necessidade de aprimoramentos arquiteturais que garantam maior robustez, modularidade e eficiência computacional nas aplicações hospitalares baseadas em IoT.

Dando continuidade às etapas anteriores da pesquisa, esta nova fase concentrou-se na validação e otimização das soluções previamente desenvolvidas, com ênfase na melhoria da robustez, escalabilidade e confiabilidade do sistema, especialmente na camada *web*. No âmbito do *software* embarcado, foram realizadas refatorações estruturais com foco em redução de latência, melhor gerenciamento de memória e maior estabilidade na comunicação com a nuvem, práticas amplamente recomendadas para sistemas IoT distribuídos [6].

Dessa forma, o presente estudo propõe a evolução do projeto IoT Santa Casa em direção a uma solução integrada, modular e escalável, capaz de atender a múltiplas demandas da gestão hospitalar contemporânea, ao mesmo tempo em que fortalece a aplicação de tecnologias IoT como instrumento estratégico de inovação no setor da saúde.

II. Metodologia

A arquitetura original do projeto IoT Santa Casa, representada na Figura 1, foi estruturada de forma modular, integrando dispositivos IoT baseados em microcontroladores ESP32 e ESP8266, comunicação via rede Wi-Fi, armazenamento em nuvem por meio do *Firebase Firestore* e uma aplicação *web* desenvolvida com o *framework Django* (Python) [3].

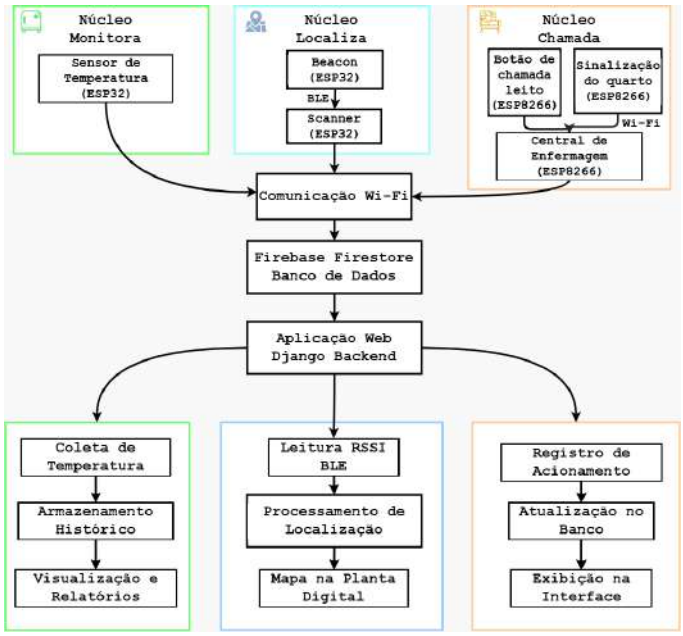


Fig. 1. Arquitetura do sistema IoT, mostrando o fluxo de dados entre os dispositivos de coleta, rede de comunicação, banco de dados e aplicação *web*.

A organização do sistema segue o modelo cliente-servidor, no qual os dispositivos embarcados realizam a coleta e a transmissão de dados para o banco em nuvem, enquanto a aplicação *web* atua como camada de visualização e gerenciamento das informações. O sistema é composto por três núcleos de funcionamento — Monitora, Localiza e Chamada —, cada um voltado a uma instância específica de requisições hospitalares, e integrados por meio de uma única aplicação *web* centralizada.

As modificações realizadas neste trabalho concentraram-se exclusivamente na ampliação funcional dessa aplicação *web*, mantendo-se inalterada a arquitetura física previamente validada [3]. A adoção de boas práticas de desenvolvimento *web* e de arquitetura orientada a serviços é amplamente reconhecida como fator determinante para a sustentabilidade de sistemas hospitalares digitais [7].

A. Núcleo Monitora

O Núcleo Monitora é responsável pelo acompanhamento contínuo da temperatura de refrigeradores clínicos utilizados para o armazenamento de medicamentos e insumos biológicos. Os dispositivos embarcados, baseados no microcontrolador ESP32 e no sensor *DS18B20*, coletam os dados de temperatura periodicamente e os transmitem ao *Firebase Firestore*, onde são armazenados tanto os valores em tempo real quanto o histórico de variações [3].

Na camada da aplicação *web*, a principal funcionalidade implementada neste trabalho consiste na geração automatizada de relatórios em formato PDF contendo o histórico completo de temperaturas monitoradas, recurso relevante para auditorias hospitalares e conformidade regulatória [8].

Adicionalmente, foram incorporadas melhorias na organização do *backend* e no aprimoramento visual da interface de visualização, visando reduzir o tempo de resposta da aplicação e aumentar sua escalabilidade.

A validação desta funcionalidade foi realizada por meio de testes com o protótipo em ambiente de desenvolvimento. Foram submetidas requisições de geração de relatório para diferentes intervalos temporais, e a resposta do sistema foi verificada diretamente no *Firebase Firestore*, confirmando que os registros de temperatura recuperados correspondiam corretamente ao período solicitado. O documento PDF gerado foi inspecionado para verificar a integridade das informações apresentadas, incluindo a identificação do equipamento, o intervalo analisado e os registros tabulares de temperatura.

B. Núcleo Localiza

O Núcleo Localiza realiza o controle de localização de equipamentos médico-hospitalares por meio de dispositivos com comunicação *Bluetooth Low Energy* (BLE). O sistema é composto por dois tipos de dispositivos físicos: o *beacon*, fixado ao equipamento rastreado, e o *scanner*, instalado em pontos estratégicos do ambiente hospitalar. A estimativa de proximidade é obtida a partir da intensidade do sinal recebido (*Received Signal Strength Indicator* – RSSI), e as leituras são enviadas ao *Firebase Firestore* para processamento [3].

Na camada da aplicação *web*, as informações são recuperadas do banco de dados e processadas no *backend*, onde é identificado o *scanner* com maior intensidade de sinal para cada *beacon*, determinando a área de maior proximidade do equipamento. A partir dessa análise, são construídas estruturas de dados que relacionam cada dispositivo ao *scanner* correspondente e aos metadados previamente cadastrados, como nome e tipo do equipamento.

A funcionalidade a implementada neste trabalho foi a visualização gráfica que permite a identificação espacial de equipamentos fora de sua localização designada, contribuindo para maior eficiência operacional e rastreabilidade de ativos.[9].

Para validar o funcionamento do módulo, foram realizados testes com o protótipo em ambiente de desenvolvimento, utilizando locais fictícios previamente cadastrados na planta digital. As leituras de RSSI dos dispositivos BLE foram registradas e enviadas ao *Firebase Firestore*, e a resposta do sistema foi monitorada diretamente no banco de dados. Verificou-se que alterações nos valores registrados refletiam corretamente na interface gráfica, evidenciando a consistência entre os dados armazenados e a representação visual dos equipamentos sobre a planta digital.

C. Núcleo Chamada

O Núcleo Chamada é responsável por um sistema de chamada de leitos, permitindo que pacientes contatem a equipe médica por meio de dispositivos físicos instalados nos quartos. Cada acionamento registra um evento no *Firebase Firestore*, que é então refletido na interface *web*, onde a equipe médica

pode visualizar quais leitos possuem chamadas ativas e o tempo decorrido desde o acionamento [3].

Dessa forma, foram realizadas modificações no contador do banco de dados, com o objetivo de tornar o registro das chamadas mais consistente e garantir a atualização adequada do estado de cada chamada no banco de dados.

A validação funcional foi conduzida por meio de testes com o protótipo em ambiente de desenvolvimento. O botão de acionamento foi utilizado diretamente na interface, e a resposta do sistema foi verificada em tempo real no *Firebase Firestore*. Constatou-se que, após cada acionamento, o estado correspondente era atualizado imediatamente no banco de dados, confirmando a comunicação adequada entre a aplicação *web* e o ambiente em nuvem.

III. Resultados e Discussão

A. Ajustes na Interface do Núcleo Monitora

Na interface do Núcleo Monitora foi implementado um mecanismo automatizado de geração de relatórios em formato PDF contendo o histórico completo das temperaturas monitoradas. A funcionalidade foi desenvolvida no *backend* da aplicação utilizando o *framework Django* e consiste em uma *view* responsável por receber requisições HTTP contendo parâmetros como o identificador do equipamento e o intervalo temporal de interesse.

Após a validação dos parâmetros recebidos, o sistema realiza a consulta aos dados históricos armazenados no banco de dados em nuvem *Firebase Firestore*, recuperando os registros de temperatura correspondentes ao período selecionado. Os dados obtidos são então processados e organizados em estruturas tabulares por meio da biblioteca *Pandas*, permitindo a padronização das informações antes da geração do relatório.

A criação do documento PDF é realizada utilizando a biblioteca FPDF, responsável pela construção dinâmica do arquivo, incluindo cabeçalhos, tabelas e elementos de formatação. O relatório gerado apresenta informações como identificação do equipamento monitorado, período analisado, faixa aceitável de temperatura e uma tabela detalhada contendo data, hora, valor de temperatura registrado e o respectivo status operacional (normal ou condição de alerta). Registros que excedem os limites configurados são destacados visualmente, facilitando a identificação de possíveis anomalias térmicas.

O documento é gerado diretamente em memória e disponibilizado ao usuário para download imediato pela interface *web*. Isso permite a exportação estruturada dos dados para fins de auditoria, rastreabilidade e documentação das condições de armazenamento.

Além da geração de relatórios, foram realizados ajustes na interface do núcleo Monitora, com melhorias visuais voltadas à organização dos elementos exibidos e à apresentação das informações de forma mais clara e esteticamente consistente. Essas alterações tiveram como objetivo aprimorar a experiência de utilização da plataforma, sem modificar sua lógica de funcionamento.

Embora não tenha sido conduzido estudo formal de usabilidade, a interface resultante apresentou melhor organização visual em ambiente de desenvolvimento. A organização visual obtida pode ser observada na Figura 2, evidenciando a melhoria na disposição dos elementos e na apresentação das informações.



Fig. 2. Interface do núcleo Monitora após ajustes visuais implementados.

B. Melhorias no Núcleo Localiza

No núcleo Localiza, a funcionalidade implementada teve como principal aplicação a visualização espacial de equipamentos médicos dentro do ambiente monitorado. O módulo permite associar cada dispositivo cadastrado a uma posição específica na planta digital, possibilitando a identificação visual de sua localização no ambiente hospitalar.

Essa abordagem contribui para a organização e rastreabilidade dos equipamentos médicos, facilitando a identificação de sua disposição física e auxiliando na gestão operacional dos ativos monitorados.

Os dados processados, correspondentes às leituras de temperatura coletadas pelo equipamento monitorado, organizadas e associadas a registros de data e hora para utilização na aplicação, são então enviados à interface da aplicação *web*, onde são disponibilizados em formato JSON e utilizados na visualização gráfica do sistema. A planta digital do ambiente hospitalar é representada por um elemento gráfico em SVG (*Scalable Vector Graphics*), no qual as diferentes áreas são previamente definidas e associadas a coordenadas espaciais na interface.

Com base nessas coordenadas, *scripts* em *JavaScript* realizam a renderização dinâmica dos equipamentos sobre a planta digital. Cada dispositivo é representado por marcadores gráficos posicionados de acordo com a área correspondente ao *scanner* responsável pela detecção. Além disso, a interface realiza atualizações periódicas por meio de requisições assíncronas ao sistema, com intervalo configurável entre 5 e 300 segundos, permitindo a atualização automática das informações exibidas conforme novos dados são registrados no banco.

O sistema também permite o cadastro de locais dentro da planta digital, possibilitando a definição de áreas ou pontos específicos do ambiente hospitalar. Dessa forma, novos equipamentos podem ser associados aos locais previamente configurados, facilitando a organização espacial e a identificação dos dispositivos monitorados.

Durante os testes realizados em ambiente de desenvolvimento, foram utilizados locais fictícios previamente definidos na planta digital para simular a disposição dos equipamentos no ambiente hospitalar. Observou-se que alterações nos da-

dos registrados refletiam adequadamente na interface gráfica, evidenciando a consistência entre o banco de dados e a representação visual do sistema, representada na Figura 3.



Fig. 3. Visualização gráfica espacial implementada no núcleo Localiza, evidenciando a disposição dos equipamentos cadastrados sobre a planta digital do ambiente monitorado.

As implementações realizadas ampliaram as capacidades funcionais do sistema ao incorporar mecanismos de visualização espacial dos equipamentos sem modificar a arquitetura original da plataforma.

Embora não tenham sido realizados testes quantitativos de desempenho ou estudos formais de usabilidade, as funcionalidades apresentaram comportamento consistente durante os testes em ambiente de desenvolvimento, demonstrando funcionamento adequado dentro do escopo proposto.

C. Validação Funcional do Núcleo Chamada

Durante a utilização do sistema foram identificadas limitações relacionadas ao funcionamento do núcleo Chamada, principalmente no registro e na contabilização do tempo associado às solicitações realizadas pelos usuários. Observou-se que o controle temporal não ocorria de forma totalmente confiável, podendo gerar divergências entre o momento de acionamento da chamada e o registro efetivo das informações no banco de dados. Essas inconsistências dificultavam o acompanhamento do tempo de resposta e comprometiam a confiabilidade dos dados utilizados para monitoramento do sistema.

Diante dessas limitações, foram realizadas correções e melhorias no contador do banco de dados, com o objetivo de tornar o registro das chamadas mais consistente e garantir a atualização adequada de seu estado no banco de dados. No núcleo Chamada, a verificação funcional foi realizada por meio da interação direta com o botão de acionamento. Ao realizar o clique na interface, constatou-se a atualização imediata do estado correspondente no banco de dados, confirmando a comunicação adequada entre a aplicação *web* e o banco de dados do sistema.

As funcionalidades apresentaram comportamento consistente durante os testes realizados em ambiente de desenvolvimento, demonstrando funcionamento adequado dentro do escopo proposto.

IV. Conclusão

A evolução apresentada neste trabalho contribui para a ampliação funcional do sistema IoT Santa Casa, incorporando recursos que expandem sua aplicabilidade prática e promovem melhorias na apresentação das informações ao usuário, sem alterar a arquitetura previamente consolidada. A implementação da geração automatizada de relatórios em PDF fortalece o potencial de utilização do sistema como ferramenta de registro e acompanhamento documental, enquanto a visualização gráfica baseada na planta estrutural adiciona uma camada interpretativa que facilita a compreensão espacial dos dados, ao mesmo tempo em que aprimora a interface visual da plataforma.

Embora as implementações tenham sido validadas em ambiente de desenvolvimento, os resultados obtidos indicam que a plataforma possui potencial para evoluções futuras voltadas à integração com demais sistemas de monitoramento hospitalar, à realização de testes em ambiente hospitalar real, bem como ao aprimoramento da interface e à expansão das funcionalidades analíticas da plataforma.

Dessa forma, o presente trabalho reforça a viabilidade de expansão progressiva de sistemas IoT aplicados ao contexto hospitalar, demonstrando que melhorias incrementais na camada de *software* podem ampliar significativamente o valor operacional da solução, mantendo a estabilidade da estrutura arquitetural existente.

V. Agradecimentos

Gostaríamos de expressar nossa gratidão à Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG, ao Instituto Nacional de Telecomunicações - INATEL, e ao eHealth Innovation Center pelo apoio técnico e recursos viabilizados para o desenvolvimento deste projeto.

Referências

[1] A. Atadoga, T. T. Omaghomi, O. A. Elufioye, I. P. Odilibe, A. I. Daraojimba e O. R. Owolabi. “Internet of Things (IoT) in healthcare: A systematic review of use cases and benefits”. Em: *International Journal of Science and Research Archive* 11.1 (2024), pp. 1511–1517. DOI: 10.30574/ijsra.2024.11.1.0243.

[2] H. Ahmadi et al. “Smart Hospitals Based on Internet of Things: Architecture, Applications and Challenges”. Em: *Journal of Healthcare Engineering* 2023 (2023), pp. 1–15.

[3] Felipe Silveira Barbosa, Filipe B. Vilela, Luma Rissatti Borges do Prado e Marcos Henrique Barbosa Pereira. “Validação e otimização de sistemas IoT de monitoramento e alertas aplicados em ambiente hospitalar”. Em: *Anais do 37º Congresso de Iniciação Científica do Inatel (INCITEL 2025)*. Santa Rita do Sapucaí, MG, Brasil: Instituto Nacional de Telecomunicações (INATEL), 2025, pp. 65–70.

[4] Sabrina Sicari, Alessandra Rizzardi, Luigi Grieco e Alberto Coen-Porisini. “Security, privacy and trust in Internet of Things: The road ahead”. Em: *Computer Networks* 76 (2015), pp. 146–164.

[5] Abdulrahman Almulhim. “Challenges in IoT-Based Healthcare Systems”. Em: *Sensors* 22.3 (2022).

[6] Eleonora Borgia. “The Internet of Things vision: Key features, applications and open issues”. Em: *Computer Communications* 54 (2014), pp. 1–31.

[7] Len Bass, Paul Clements e Rick Kazman. *Software Architecture in Practice*. Addison-Wesley, 2012.
[8] World Health Organization. *Temperature Sensitivity of Vaccines*. WHO/IVB/11.05. 2011.
[9] K. Wang et al. “IoT-Based Asset Tracking and Monitoring in Healthcare Environments”. Em: *IEEE Access* 6 (2018), pp. 12345–12354.

AUTORES



Pedro Henrique de Paiva Carvalho é aluno de graduação do curso de Engenharia da Computação do Inatel (Instituto Nacional de Telecomunicações). Bolsista de iniciação científica pelo grupo eHealth Inatel, atuando no desenvolvimento e aprimoramento de sistemas IoT aplicados ao contexto hospitalar. Participa de projetos voltados à integração entre dispositivos embarcados e aplicações web, contribuindo para a implementação e validação funcional de soluções tecnológicas.



Rafael Gonçalves Magalhães é graduando em Engenharia de Computação no Instituto Nacional de Telecomunicações (Inatel), atualmente no terceiro ano do curso. Atua há mais de um ano em projeto de Iniciação Científica, com foco em pesquisa e desenvolvimento de sistemas embarcados e soluções em Internet das Coisas (IoT), com ênfase em aplicações para ambientes hospitalares.



Luma Rissatti Borges do Prado graduada em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (Inatel), é mestranda em Engenharia Biomédica pela Escola Politécnica da Universidade de São Paulo (USP) e possui MBA em Gestão de Projetos e Metodologias Ágeis. Atua como Analista Sênior de Desenvolvimento de Produtos na Philips e é professora no curso de Engenharia Biomédica do Inatel, onde também é pesquisadora e orientadora no eHealth Innovation Center.

Sistema de inteligência artificial e redes neurais para suporte a tomada de decisão em tratamentos fisioterapêuticos

João Carlos Rodrigues Franco & Filipe Bueno Vilela
Instituto Nacional de Telecomunicações – Inatel
joao_carlos@ges.inatel.br, filipe.bueno@inatel.br

Abstract— *Physiotherapy diagnosis is a very common clinical practice. Many diagnoses present similar symptoms but require different therapeutic approaches, which may lead to diagnostic errors and consequently affect the patient's clinical outcome. To assist in this process, this project aims to develop an artificial intelligence–based system capable of supporting the generation of patient treatment plans, helping to reduce diagnostic errors and improve decision-making in physiotherapy.*

Index Terms— *Artificial intelligence; Physiotherapy; AI in physiotherapy; Machine learning.*

Resumo— O diagnóstico fisioterapêutico é uma prática clínica muito comum. Muitos diagnósticos apresentam sintomas semelhantes, porém requerem abordagens terapêuticas distintas, o que pode levar a erros diagnósticos e, consequentemente, influenciar negativamente o desfecho clínico do paciente. Para auxiliar nesse processo, este projeto tem como objetivo desenvolver um sistema baseado em inteligência artificial capaz de apoiar a geração de planos de tratamento para pacientes, contribuindo para a redução de erros diagnósticos e para a melhoria da tomada de decisão em fisioterapia.

Palavras-chave— Inteligência artificial; Fisioterapia; IA em fisioterapia; Aprendizado de máquina.

I. Introdução

O diagnóstico fisioterapêutico constitui uma etapa fundamental no processo de avaliação clínica e planejamento terapêutico. Na prática profissional, pequenas informações ou detalhes clínicos que passam despercebidos durante a avaliação inicial podem alterar significativamente o direcionamento do tratamento e a prescrição dos exercícios terapêuticos. Tais variações podem impactar diretamente o tempo necessário para a recuperação do paciente e, consequentemente, os resultados clínicos obtidos ao longo do processo de reabilitação. Em muitos casos, diferenças sutis nos sinais e sintomas apresentados pelos pacientes podem conduzir a estratégias terapêuticas distintas, ainda que as condições clínicas aparentem ser semelhantes [1].

Nesse contexto, torna-se evidente a necessidade de ferramentas que auxiliem o fisioterapeuta na análise sistemática das informações clínicas, reduzindo a possibilidade de erros ou inconsistências no processo diagnóstico. A aplicação de técnicas de inteligência artificial tem se mostrado uma abordagem promissora para apoiar processos decisórios em diferentes áreas da saúde. O estudo e a aplicação de métodos de inteligência artificial podem contribuir para a análise estruturada de dados clínicos, auxiliando tanto na identificação de padrões diagnósticos quanto na definição de estratégias terapêuticas mais adequadas [1, 2].

Nos últimos anos, a aplicação de técnicas de inteligência artificial na área da saúde tem crescido de forma significativa, especialmente em sistemas de apoio à decisão clínica. Métodos baseados em aprendizado de máquina e aprendizado profundo têm sido utilizados para analisar grandes volumes de dados clínicos, identificar padrões complexos e auxiliar profissionais de saúde na tomada de decisões mais precisas. No contexto da reabilitação e da fisioterapia, essas tecnologias podem ser aplicadas na análise de dados provenientes de avaliações funcionais, registros clínicos e monitoramento de movimentos corporais. Dessa forma, sistemas inteligentes podem contribuir para a identificação de padrões associados a determinadas disfunções musculoesqueléticas, além de auxiliar na recomendação de intervenções terapêuticas mais adequadas para cada perfil de paciente [3] .

Além disso, avanços recentes em visão computacional, processamento de linguagem natural e análise de movimento têm ampliado as possibilidades de utilização da inteligência artificial em processos de avaliação e acompanhamento fisioterapêutico. Ferramentas baseadas em análise de vídeo, sensores vestíveis e algoritmos de reconhecimento de padrões permitem monitorar a execução de exercícios terapêuticos, avaliar a qualidade do movimento e fornecer feedback em tempo real para pacientes e profissionais. Estudos recentes demonstram que a integração dessas tecnologias em sistemas clínicos pode melhorar a precisão das avaliações funcionais, ampliar a capacidade de monitoramento remoto e contribuir para estratégias de reabilitação mais personalizadas. Nesse cenário, o desenvolvimento de sistemas inteligentes voltados ao suporte do diagnóstico fisioterapêutico representa uma área promissora de pesquisa, com potencial para aprimorar a qualidade da prática clínica e otimizar os resultados do tratamento [4].

I.1. IA para medição de escala de esforço em fisioterapia

Tradicionalmente, a avaliação do esforço físico ou da dor percebida pelo paciente em contextos fisioterapêuticos é realizada por meio de instrumentos subjetivos, como a Borg Rating of Perceived Exertion Scale (RPE) e escalas visuais analógicas de dor. Embora amplamente utilizadas na prática clínica, essas abordagens dependem da autorreferência do paciente e da interpretação do profissional, o que pode introduzir variabilidade nos resultados e limitar a padronização das avaliações durante o processo de reabilitação. Em função dessas limitações, pesquisas recentes têm investigado o uso de técnicas de inteligência artificial para tornar a mensuração do esforço físico mais objetiva, utilizando dados provenientes de sensores fisiológicos, análise de movimento e processamento de imagens [5].

Nesse contexto, algoritmos de aprendizado de máquina e visão computacional têm sido empregados para estimar automaticamente o nível de esforço ou fadiga durante a realização de exercícios terapêuticos. Alguns sistemas utilizam análise de expressões faciais para identificar padrões associados a dor, desconforto ou esforço físico, enquanto outros combinam informações provenientes de sensores vestíveis, como frequência cardíaca, eletromiografia ou acelerômetros, para estimar o nível de esforço do paciente em tempo real. Esses modelos são capazes de reconhecer padrões complexos nos dados coletados e classificá-los em diferentes níveis de intensidade ou fadiga, fornecendo ao fisioterapeuta uma avaliação mais objetiva do desempenho do paciente. Dessa forma, a integração de inteligência artificial em sistemas de monitoramento fisioterapêutico pode contribuir para avaliações mais consistentes, além de permitir um acompanhamento mais preciso da evolução do paciente durante o processo de reabilitação [5, 6].

I.2. Software de análise de movimento baseado em inteligência artificial para reabilitação esportiva e fisioterapia

O OpenPose [7] é uma ferramenta computacional baseada em técnicas de inteligência artificial voltada para a análise automatizada de movimentos humanos. O software utiliza algoritmos de visão computacional capazes de identificar e rastrear pontos-chave do corpo humano em imagens ou vídeos, permitindo a reconstrução da postura e da dinâmica dos movimentos realizados pelo indivíduo.

No contexto da fisioterapia e da reabilitação esportiva, essa tecnologia possibilita a análise detalhada da execução de exercícios físicos, contribuindo para a identificação de padrões de movimento, correção postural e monitoramento do progresso do paciente ao longo do tratamento. Além disso, a utilização de sistemas baseados em inteligência artificial para análise de movimento tem se mostrado especialmente relevante para aplicações em reabilitação remota. A necessidade de soluções desse tipo tornou-se ainda mais evidente durante a pandemia de COVID-19, período em que o atendimento presencial foi significativamente reduzido, evidenciando a importância de ferramentas digitais capazes de auxiliar no acompanhamento terapêutico à distância.

I.3. Medme

O Medme é um sistema voltado para a prevenção, diagnóstico e acompanhamento de tratamentos fisioterapêuticos que incorpora recursos de inteligência artificial em sua estrutura. A plataforma foi desenvolvida com o objetivo de auxiliar profissionais da área da saúde no monitoramento da execução de exercícios terapêuticos realizados pelos pacientes [4].

Por meio de algoritmos de análise de movimento, o sistema é capaz de identificar possíveis inadequações na execução dos exercícios prescritos, realizando correções automáticas e fornecendo orientações para melhorar a qualidade da movimentação. Além disso, a plataforma gera feedback automático para o profissional responsável pelo acompanhamento do paciente, permitindo um monitoramento mais eficiente do progresso terapêutico. Dessa forma, ferramentas como o Medme demonstram o potencial da inteligência artificial para ampliar a capacidade de acompanhamento clínico e aprimorar a qualidade das intervenções fisioterapêuticas.

I.4. Objetivo da pesquisa

Dado o exposto, o presente trabalho propõe o desenvolvimento de um sistema baseado em inteligência artificial capaz de interpretar os dados obtidos durante o processo de avaliação fisioterapêutica. A partir da análise dessas informações, o sistema tem como finalidade gerar recomendações para a elaboração de planos de tratamento e sugestões de exercícios terapêuticos, atuando como ferramenta de suporte à decisão clínica. Espera-se que tal abordagem contribua para aprimorar a qualidade dos diagnósticos fisioterapêuticos, além de oferecer suporte ao profissional durante o planejamento terapêutico, favorecendo uma recuperação mais eficiente e personalizada para cada paciente.

II. METODOLOGIA

Um sistema computacional foi desenvolvido com o objetivo de gerenciar registros de diagnósticos fisioterapêuticos e comunicar os resultados obtidos ao profissional responsável pelo acompanhamento clínico. A arquitetura do sistema foi estruturada utilizando tecnologias modernas de desenvolvimento web. Para o gerenciamento do back-end, responsável pela lógica de negócio e pelo processamento das informações, foi utilizado o framework NestJS. Já a implementação da interface do usuário (front-end), responsável pela interação com o sistema, foi realizada utilizando a biblioteca ReactJS. Para o armazenamento e gerenciamento dos dados clínicos, foi adotado o sistema de gerenciamento de banco de dados PostgreSQL. Todas as ferramentas empregadas no desenvolvimento do sistema são de código aberto (open-source), o que contribui para maior flexibilidade, escalabilidade e possibilidade de reprodução da solução em outros contextos de pesquisa.

Paralelamente ao desenvolvimento da plataforma, foi elaborado um formulário estruturado destinado à coleta de informações clínicas utilizadas no processo de avaliação e diagnóstico fisioterapêutico. O formulário foi desenvolvido com o auxílio e validação de uma profissional fisioterapeuta, garantindo que os campos contemplassem os principais elementos clínicos necessários para a anamnese e para o exame físico do paciente. Após o processo de validação, o formulário foi integrado ao sistema, permitindo que os dados inseridos pelos profissionais fossem armazenados e posteriormente utilizados na análise realizada pela inteligência artificial. Adicionalmente, foi realizada uma busca por registros reais de diagnósticos e tratamentos fisioterapêuticos com o objetivo de construir uma base de dados capaz de alimentar o processo de treinamento dos modelos de aprendizado de máquina.

Durante o desenvolvimento do sistema, também foi conduzido um levantamento bibliográfico e técnico sobre diferentes abordagens de treinamento de modelos de inteligência artificial aplicáveis ao problema proposto. A partir dessa análise, foram identificadas duas abordagens principais com potencial de apresentar resultados satisfatórios para a aplicação pretendida. A primeira abordagem envolve técnicas de Natural Language Processing (NLP), que utilizam métodos de processamento de linguagem natural para interpretar e analisar informações textuais presentes em registros clínicos. A segunda abordagem considerada foi baseada em modelos generativos fundamentados na arquitetura transformer, conhecidos como modelos generativos pré-treinados (Generative Pre-trained Transformers). Esses modelos utilizam mecanismos de atenção para processar grandes volumes de dados e gerar respostas ou inferências a partir das informações fornecidas. Um exemplo amplamente conhecido dessa classe de modelos é o ChatGPT, que emprega arquiteturas baseadas em transformers para processamento avançado de linguagem natural.

III. RESULTADOS

As seções a seguir trazem as informações e recursos obtidos a partir do desenvolvimento da solução.

III.1. Página de Login

A página de login constitui o ponto inicial de acesso ao sistema, sendo responsável pela autenticação dos usuários cadastrados. Por meio dessa interface, o usuário fornece suas credenciais de acesso para ingressar na plataforma. Adicionalmente, o sistema disponibiliza mecanismos de recuperação de senha, permitindo a redefinição das credenciais em casos de esquecimento ou necessidade de atualização, garantindo assim maior segurança e continuidade no acesso ao sistema.

III.2. Página Inicial

Após a autenticação bem-sucedida, o usuário é direcionado para a página inicial do sistema. Essa interface centraliza informações relevantes para o acompanhamento clínico, apresentando, por exemplo, registros dos pacientes recentemente cadastrados ou atendidos. Além disso, a página inicial atua como um painel de navegação que permite o acesso às diferentes funcionalidades do sistema, facilitando a gestão das informações clínicas e o fluxo de trabalho do profissional. A Fig. 1 é uma representação de como a sequência de pacientes é cadastrada.

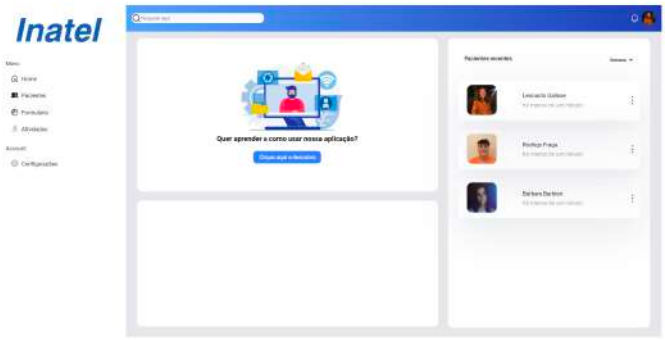


Fig. 1. Página inicial com o cadastro dos pacientes.

III.3. Página de Cadastro de Paciente

A página de cadastro de paciente tem como finalidade registrar as informações iniciais do paciente no sistema. Esse procedimento é fundamental para a organização dos dados clínicos, uma vez que o registro do paciente constitui um pré-requisito para a realização das etapas subsequentes do processo de avaliação fisioterapêutica. Sem a criação do cadastro do paciente, não é possível associar registros de anamnese ou exames físicos, impossibilitando, consequentemente, o acesso ao módulo responsável pela elaboração do diagnóstico.

III.4. Página de Inserção de Anamnese

A página de inserção de anamnese é destinada ao registro das informações clínicas obtidas durante a entrevista inicial com o paciente. Essa interface foi estruturada com base nos campos presentes no formulário de avaliação clínica utilizado no processo de diagnóstico fisioterapêutico. O preenchimento adequado dessas informações é essencial para a continuidade do fluxo de avaliação, uma vez que a conclusão da etapa de anamnese habilita o acesso ao módulo de registro do exame físico.

III.5. Página de Inserção de Exame Físico

A página de inserção do exame físico é destinada ao registro das informações obtidas durante a avaliação funcional do paciente. Os campos presentes nessa interface foram estruturados com base nos parâmetros clínicos utilizados no formulário de análise e diagnóstico fisioterapêutico. O preenchimento dessas informações permite caracterizar de forma mais precisa as condições físicas do paciente, sendo um passo fundamental para o acesso à etapa de elaboração do diagnóstico e planejamento terapêutico.

III.6. Página de Diagnóstico

A página de diagnóstico representa a etapa final do processo de avaliação clínica no sistema. Nesse ambiente, o fisioterapeuta registra suas conclusões sobre o estado clínico do paciente, incluindo o diagnóstico fisioterapêutico, o plano de tratamento proposto, os objetivos terapêuticos e eventuais observações relevantes. Essas informações consolidam os dados coletados nas etapas anteriores e servem de base para a condução do processo de reabilitação.

III.7. Página de Atividades

A página de atividades apresenta ao paciente os exercícios terapêuticos que deverão ser realizados ao longo do tratamento. Esses exercícios são definidos pelo fisioterapeuta com base na análise das informações clínicas e no suporte fornecido pelo sistema de inteligência artificial. Dessa forma, a interface possibilita ao paciente acompanhar as orientações terapêuticas prescritas, contribuindo para a execução adequada das atividades e para o monitoramento do progresso durante o processo de reabilitação. A Fig. 2 é um registro da página de atividades, em que demonstra as informações indicadas ao paciente.

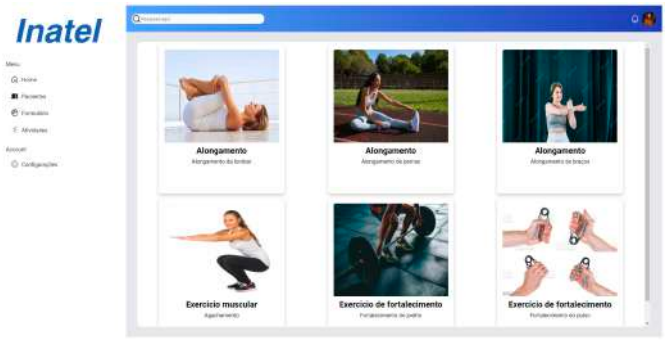


Fig. 2. Página de atividades indicadas para os pacientes.

IV. Discussão

Os modelos de inteligência artificial analisados neste estudo apresentam abordagens metodológicas distintas para o processamento e interpretação dos dados clínicos. Entre as alternativas avaliadas, destaca-se o modelo baseado em Natural Language Processing (NLP), cujo objetivo principal consiste em treinar algoritmos capazes de interpretar informações textuais e extrair padrões relevantes a partir da análise de linguagem natural. Nesse tipo de abordagem, o processo de aprendizado depende da identificação e classificação de palavras-chave presentes nos registros clínicos, permitindo que o sistema associe determinados termos ou expressões a possíveis diagnósticos ou condutas terapêuticas. Entretanto, para que esse tipo de modelo alcance desempenho satisfatório, é necessária a disponibilidade de um grande volume de registros clínicos estruturados e padronizados. Na prática, muitos diagnósticos fisioterapêuticos são registrados de forma heterogênea e sem um padrão formal de documentação [8], o que pode resultar em perda de informações relevantes durante o processo de treinamento do modelo, dificultando sua aplicação no contexto específico deste trabalho.

Diante dessas limitações, optou-se por investigar a utilização de um modelo generativo treinado a partir de dados estruturados provenientes do formulário de diagnóstico desenvolvido para o sistema. Nesse cenário, o modelo utiliza um conjunto fixo de variáveis clínicas previamente definidas, permitindo que a análise seja realizada com base em um número constante de parâmetros. Como os dados coletados referem-se exclusivamente ao contexto do diagnóstico fisioterapêutico, o modelo pode ser caracterizado como um sistema generativo especializado, direcionado a um domínio específico de aplicação [9]. Essa abordagem tende a produzir resultados mais consistentes, uma vez que o modelo é treinado utilizando informações diretamente relacionadas ao problema clínico analisado.

Outro aspecto relevante do desenvolvimento do sistema refere-se ao tratamento e à padronização dos dados utilizados no treinamento do modelo. Informações de caráter sigiloso ou que permitam a identificação direta do paciente, como nome, data de nascimento e número de documentos pessoais, não são utilizadas no processo de treinamento. Os dados clínicos selecionados são convertidos em representações quantitativas, permitindo sua utilização em algoritmos de aprendizado de máquina. Para isso, variáveis clínicas são normalizadas em escalas numéricas padronizadas, variando entre 0 e 10, nas quais valores mais baixos representam condições clínicas desfavoráveis e valores mais altos indicam melhores condições funcionais. Em casos de variáveis binárias, utiliza-se uma codificação simples com valores 0 ou 1, representando respostas negativas ou positivas, respectivamente. Por exemplo, uma amplitude de movimento classificada com valor 3 pode indicar uma limitação significativa da mobilidade articular.

A quantificação e padronização dos dados clínicos são etapas fundamentais para viabilizar o treinamento eficiente do modelo

generativo, uma vez que algoritmos de aprendizado de máquina operam de forma mais eficiente quando os dados estão estruturados em formatos numéricos consistentes. Além disso, essa estratégia contribui para reduzir ambiguidades na interpretação das informações clínicas e facilitar o processamento computacional dos dados. Com base na consulta a especialistas da área, estimou-se que um conjunto de aproximadamente mil registros de diagnósticos fisioterapêuticos seria suficiente para fornecer uma base inicial adequada para o treinamento do modelo de inteligência artificial. Considerando a existência de um formulário padronizado de coleta de dados clínicos e a estrutura organizada das variáveis utilizadas, a adoção de um modelo generativo especializado apresenta-se como uma alternativa mais adequada em comparação ao modelo baseado em processamento de linguagem natural para o contexto deste estudo.

V. Conclusão

O sistema de gerenciamento de diagnósticos desenvolvido oferece funcionalidades essenciais para o funcionamento e integração do módulo de inteligência artificial. Entre essas funcionalidades, destaca-se a interface de cadastro de pacientes, responsável pelo registro inicial das informações do indivíduo no sistema. A partir desse cadastro, torna-se possível associar formulários clínicos específicos destinados à realização das avaliações fisioterapêuticas. Esse processo constitui a etapa inicial do fluxo de análise clínica dentro da plataforma.

Após o cadastro do paciente, o sistema permite o preenchimento de formulários estruturados referentes à anamnese e ao exame físico. Nessas etapas, o profissional insere informações clínicas obtidas durante a avaliação, incluindo dados sobre histórico clínico, sintomas relatados e parâmetros funcionais observados. Essas informações estruturadas constituem a base de dados utilizada pelo sistema para a análise automatizada realizada pela inteligência artificial. A partir desse conjunto de dados, o sistema tem como objetivo auxiliar o profissional na elaboração do diagnóstico fisioterapêutico e na definição de um plano de tratamento adequado às condições clínicas do paciente. Adicionalmente, a plataforma disponibiliza um módulo destinado à apresentação das atividades terapêuticas, no qual o paciente pode visualizar os exercícios que deverão ser realizados ao longo do processo de reabilitação.

A análise das diferentes abordagens de inteligência artificial indicou que a utilização de um modelo generativo especializado representa a alternativa mais adequada para o contexto do sistema proposto. Essa escolha está associada à existência de um formulário padronizado para a coleta de dados clínicos, o que favorece a consistência e a uniformidade das informações utilizadas no treinamento do modelo. A padronização das variáveis reduz a variabilidade dos dados de entrada e contribui para um processo de aprendizado mais eficiente por parte da inteligência artificial. Dessa forma, espera-se que a construção de uma base de dados estruturada possibilite, em etapas futuras do projeto, o treinamento robusto do modelo e a implementação de um sistema de apoio à decisão clínica capaz de auxiliar o fisioterapeuta no processo de diagnóstico e planejamento terapêutico.

IV. Agradecimentos

Os autores agradecem ao Instituto Nacional de Telecomunicações (INATEL) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio e fomento concedidos a este projeto de pesquisa.

Referências

1. K. R. Chowdhary. “Introducing Artificial Intelligence”. Em: Fundamentals of Artificial Intelligence. New Delhi: Springer India, 2020, pp. 1–23. ISBN: 978-81-322-3972- 7. DOI: 10.1007/978-81-322-3972-7 1. URL: https: //doi.org/ 10.1007/978-81-322-3972-7 1.
2. TOPOL, Eric J. High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, v. 25, p. 44–56, 2019.
3. SENDERS, Jeffrey T. et al. Machine learning and neurosurgical outcome prediction: a systematic review. World Neurosurgery, v. 109, p. 476–486, 2018.
4. VAN DIJK, H. et al. Artificial intelligence in physiotherapy: a systematic review of applications and challenges. Journal of Medical Systems, v. 45, 2021.
5. BARTLETT, J. et al. Detecting pain and physical exertion using facial expression recognition with deep learning. IEEE Journal of Biomedical and Health Informatics, 2020.
6. WANG, R.; BLACKFORD, S.; WU, Z. Machine learning approaches for monitoring physical exertion and fatigue using wearable sensors. Sensors, v. 19, n. 21, 2019.
7. Fanni Zsarnoczky-Dulhazi, Solt Agod, Steve Szarka, Kornelia Tuza e Bence Kopper. “AI BASED MOTION ANALYSIS SOFTWARE FOR SPORT AND PHYSI- CAL THERAPY ASSESSMENT”. Em: Revista Bra- sileira de Medicina do Esporte (2023). DOI: https: / / doi . org / 10 . 1590 / 1517 - 8692202430012022 0020p. URL: https://www.scielo.br/ j/rbme/a/ nZdcCvptq99Qmzp7FF6ds5F/?format=pdf&lang=en.
8. MONIQUE DE SOUZA PUTTINI CALZA´. “Adequac¸a˜o de diagno´sticos e metas fisioterapeˆuticas na pra´tica cl´ınica”. Em: (). URL: https://bdm.unb.br/ handle/10483/26836.
9. [14] Generalists vs. Specialists in Data Science and Analy- tics. URL: https://towardsdatascience.com/generalists- vs - specialists - in - data - science - and - analytics - fe5d8b55f1e6 (acesso em 2023).

Autores



João Carlos Rodrigues Franco é graduando em engenharia de Software. É bolsista de iniciação cientifica em 2023 no eHealth Innovation Center em IA em fisioterapia. Estagiou no Instituto de Previdência do Município de Pouso Alegre. Atualmente é desenvolvedor de back-end e possui interesses na área de engenharia de software.



Filipe Bueno Vilela é engenheiro biomédico pelo Instituto Nacional de Telecomunicações (INATEL). Possui mestrado e doutorado em Engenharia de Materiais, com ênfase em biomateriais, pela Universidade Federal de Itajubá (UNIFEI). Atualmente, atua como professor, pesquisador e desenvolvedor no INATEL.

Desenvolvimento de um Simulador Interativo de Máquina de Anestesia Baseado em ESP32 para a Educação em Engenharia Biomédica

Marcelo Fabian Cruz Guizada, Mauricio Marcelo Peredo Claros, Rodrigo Martinez Severich
Universidad Privada del Valle – UNIVALLE
cgm0031216@est.univalle.edu, peredo.mauri@gmail.com, rmartinezs@univalle.edu
https://orcid.org/0009-0003-6573-7732, https://orcid.org/0000-0002-5884-3321, https://orcid.org/0009-0003-5588-7753

***Abstract*—**The development of physical simulators represents a valuable tool for understanding complex medical systems during biomedical engineering education. This work presents the design and development of an interactive anesthesia machine simulator capable of reproducing, at a basic structural and functional level, the main subsystems of these clinical devices. The system integrates flow and pressure sensors, electromechanical actuators, electronic control circuits, and a graphical interface for variable visualization and the execution of different operational scenarios. Its architecture is organized into power supply, sensing, control, and actuation modules, implemented through microcontrollers for signal acquisition and system management. The simulator enables the recreation of gas flow dynamics, system tests, and common faults, providing a controlled environment to analyze the functional behavior of an anesthesia machine.

***Index Terms*—**Anesthesia machine, biomedical engineering, embedded systems, sensors, simulation.

***Resumo*—**O desenvolvimento de simuladores físicos constitui uma ferramenta relevante para a compreensão de sistemas médicos complexos durante a formação em engenharia biomédica. Este trabalho apresenta o projeto e desenvolvimento de um simulador interativo de máquina de anestesia capaz de reproduzir, em nível estrutural e funcional básico, os principais subsistemas desses equipamentos clínicos. O sistema integra sensores de fluxo e pressão, atuadores eletromecânicos, circuitos eletrônicos de controle e uma interface gráfica para visualização de variáveis e execução de diferentes cenários operacionais. Sua arquitetura é organizada em módulos de alimentação, sensoriamento, controle e atuação, implementados por microcontroladores para aquisição de sinais e gestão do sistema. O simulador permite recriar dinâmicas de fluxo de gás, testes de sistema e falhas comuns, proporcionando um ambiente controlado para analisar o comportamento funcional de uma máquina de anestesia.

***Palavras-chave*—**Engenharia biomédica, máquina de anestesia, sensores, simulação, sistemas embarcados.

I. INTRODUÇÃO

As máquinas de anestesia constituem um dos equipamentos mais complexos no ambiente cirúrgico, uma vez que integram múltiplos subsistemas responsáveis pelo fornecimento e controle de gases médicos, ventilação mecânica, monitoramento de parâmetros e mecanismos de segurança. Esses sistemas combinam componentes pneumáticos, eletrônicos e de controle que devem operar de forma coordenada para garantir o funcionamento seguro do equipamento.

A análise e compreensão do funcionamento desses dispositivos representam um desafio do ponto de vista técnico, devido à integração de diversos sensores, atuadores e sistemas de controle. Por essa razão, o desenvolvimento de plataformas de simulação permite reproduzir o comportamento básico desses sistemas em ambientes controlados, facilitando o estudo de sua arquitetura e dinâmica de operação sem a necessidade do uso direto de equipamentos clínicos.

Nesse contexto, este trabalho apresenta o projeto e desenvolvimento de um simulador interativo de máquina de anestesia orientado a reproduzir, em nível estrutural e funcional básico, os principais subsistemas presentes nesses equipamentos. O sistema integra sensores, atuadores, módulos eletrônicos de controle e uma interface gráfica que permite visualizar variáveis do sistema e executar diferentes cenários operacionais.

O artigo descreve a arquitetura do simulador, o projeto eletrônico do sistema, a integração de hardware e software embarcado e o desenvolvimento da interface gráfica para supervisão e controle do sistema.

II. ENGENHARIA DO PROJETO

A. Arquitetura do Sistema

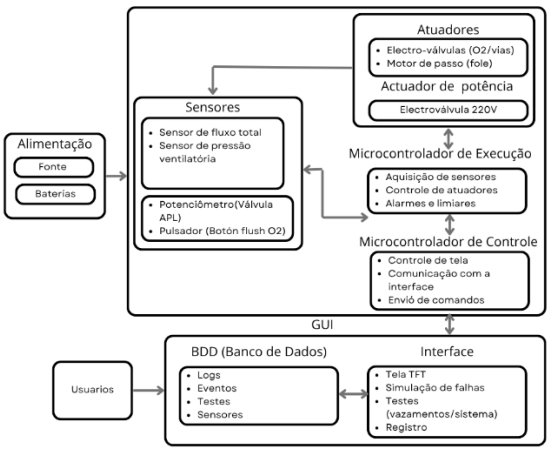


Fig. 1. Diagrama de blocos do simulador de máquina de anestesia.

Na Figura 1 mostra a arquitetura geral do simulador da máquina de anestesia. O sistema está organizado em módulos interconectados que reproduzem as funções essenciais do equipamento clínico real, incluindo detecção, atuação, controle, visualização e interação com o usuário. Essa estrutura modular facilita a compreensão da relação entre hardware, software e mecanismos de segurança operacional dentro de um ambiente educacional seguro, uma abordagem amplamente utilizada em simuladores médicos projetados para o treinamento em anestesia [1], [2], [3].

O módulo de alimentação é composto por uma fonte principal e um sistema de baterias, devido à presença de diversos sensores e atuadores que requerem diferentes níveis e características de potência. Além de fornecer as tensões necessárias para o funcionamento do sistema, esse módulo permite simular condições de fornecimento de energia de emergência e reproduzir cenários associados a falhas no fornecimento elétrico, os quais constituem situações relevantes no treinamento clínico por meio de simulação [1].

O bloco de detecção coleta as principais variáveis do sistema de ventilação simulado. Esse módulo inclui um sensor de fluxo total e um sensor de pressão ventilatória. Além disso, incorpora-se um potenciômetro que simula o funcionamento da válvula APL e um botão que reproduz a função de descarga de O_2. Em conjunto, esses elementos permitem medir as condições relevantes do circuito ventilatório e reproduzir tarefas de monitoramento típicas de uma máquina anestésica, nas quais a medição de fluxo e pressão é fundamental para compreender o comportamento do sistema respiratório durante a ventilação anestésica [4].

A arquitetura do sistema incorpora duas unidades de processamento distintas. O microcontrolador de execução é responsável por adquirir os sinais dos sensores, controlar os atuadores e gerenciar os limites de alarme. Essa separação funcional permite a distribuição de tarefas entre ambos os microcontroladores, evitando a sobrecarga de processamento e assegurando um funcionamento estável do sistema. Por outro lado, o microcontrolador de controle gerencia a comunicação com a interface gráfica, controla a tela TFT e envia os comandos necessários ao módulo de execução, permitindo a implementação de uma estrutura modular semelhante à utilizada em equipamentos anestésicos comerciais e em sistemas de controle automatizado aplicados à anestesia [5], [6], [7].

O módulo atuador é composto por válvulas solenoides responsáveis por direcionar o fluxo de gases, um motor de passo que controla o movimento do fole e uma válvula solenoide de 220 V, utilizada para simular o fluxo em alta pressão. A interação desses atuadores permite reproduzir a dinâmica de pressão, fluxo e ventilação, facilitando a interpretação do comportamento de um sistema anestésico real, no qual múltiplos componentes mecânicos e pneumáticos interagem para controlar a ventilação do paciente [4], [8].

O sistema também inclui uma interface de usuário que integra a tela TFT, ferramentas de simulação de falhas, testes do sistema e um módulo de registro de eventos. Essas funcionalidades permitem ao usuário interagir com funções críticas do simulador, visualizar parâmetros do sistema, realizar testes de vazamento e ativar condições anômalas para fins de análise. Todas essas informações são armazenadas em um banco de dados (BDD) que registra logs, eventos, testes e valores dos sensores, proporcionando rastreabilidade do desempenho durante a operação do simulador. O uso de interfaces interativas e registro de eventos é uma prática comum em plataformas de simulação médica utilizadas para a formação clínica e a avaliação de competências [9], [10].

Em conjunto, esses módulos permitem a interação direta do usuário com a interface gráfica e o ambiente de simulação. Essa organização facilita a compreensão do fluxo completo do sistema anestésico, bem como a interpretação de alarmes, a resposta a falhas e a realização de testes operacionais, aspectos fundamentais no treinamento baseado em simulação aplicado à anestesia [1], [11].

B. Projeto de Hardware

b.1) Projeto da fonte de alimentação

O sistema requer múltiplos níveis de tensão para alimentar os diferentes módulos eletrônicos e atuadores do simulador. Para dimensionar a fonte de alimentação, foi realizada uma análise de consumo considerando os principais componentes do sistema, incluindo o microcontrolador ESP32, eletroválvulas de 12 V e 6 V e o servomotor MG996R. Os requisitos elétricos identificados (Tabela 1) resultaram em um consumo total estimado de 2,29 A.

TABELA I
Tensões e correntes utilizadas no projeto

Componente	Tensão	Corrente
ESP-32 38 pinos	5V	0,24 A
Electroválvula 12V	12 V	0,30 A
Electroválvula 6V	6 V	0,25 A
Servo motor MG-996r	6 V	1.5 A
	CORRENTE TOTAL	2,29 A

Devido ao fato de que alguns atuadores apresentam picos transitórios de corrente durante a partida ou mudanças de estado, foi incorporada uma margem de segurança de 30 %, elevando a corrente de projeto para aproximadamente 2,97 A, a fim de garantir estabilidade e evitar quedas de tensão no sistema de alimentação, prática comum no dimensionamento de fontes para sistemas eletrônicos com cargas indutivas [12].

$$I_{Fonte} = I_{Total} + (I_{Total} * 0.3) \quad (1)$$

$$I_{Fonte} = 2{,}29\ A + (2{,}29 * 0.3) \quad (2)$$

$$I_{Fonte} = 2{,}97\ [A] \quad (3)$$

O circuito de alimentação foi projetado como uma fonte múltipla de 3 A, capaz de fornecer as tensões requeridas por cada módulo do sistema, assegurando um fornecimento elétrico estável para o controle, sensoriamento e atuação do simulador [12].

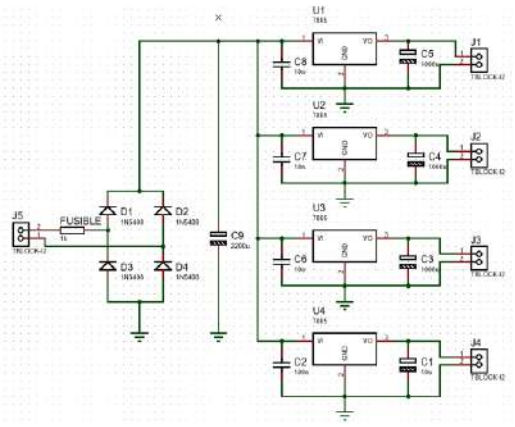


Fig. 2. Circuito de Fonte de Alimentação Múltipla de 3 A.

b.2) Projeto do circuito de execução e controle de atuadores

O circuito de execução e controle de atuadores (Figura 3) tem como objetivo reproduzir o comportamento dinâmico dos componentes de uma máquina de anestesia por meio do controle eletrônico de válvulas, motores e atuadores. Este módulo é baseado em um microcontrolador ESP32, que atua como unidade central responsável por processar os sinais provenientes dos sensores e executar as ações correspondentes nos atuadores [13].

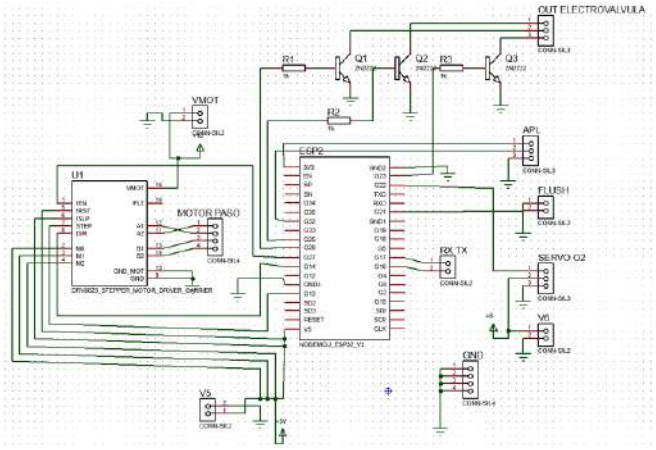


Fig. 3. Projeto esquemático do circuito de controle de atuadores.

Para simular o movimento do fole do sistema respiratório foi utilizado um motor de passo, controlado por meio do driver DRV8825, o qual permite micropassos de até 1/32 e suporta correntes de até 2,2 A por fase. Essa característica permite gerar movimentos suaves e precisos do mecanismo, o que é fundamental para simular ciclos respiratórios dentro do sistema [14].

O controle das eletroválvulas é realizado por meio de transistores 2N2222, que funcionam como etapa de comutação entre o microcontrolador e um módulo de relés de quatro canais. Esse projeto permite isolar a lógica de controle das cargas indutivas, evitando que as correntes geradas pelas bobinas das válvulas afetem o funcionamento do microcontrolador, abordagem comum em circuitos de controle de atuadores eletromecânicos [12].

Adicionalmente, o sistema incorpora um servomotor MG996R, responsável por representar o movimento de uma válvula auxiliar do sistema. Este atuador foi selecionado devido ao seu alto torque mecânico e à sua compatibilidade com sinais PWM gerados pelo ESP32, permitindo controlar sua posição com precisão [13].

O circuito também inclui um potenciômetro que simula a válvula APL (Adjustable Pressure Limiting), permitindo ajustar manualmente o limite de pressão do sistema simulado. Da mesma forma, foi implementado um botão de ativação do O_2 Flush, utilizado para simular a injeção rápida de oxigênio, uma função importante em equipamentos anestésicos reais [15].

Em conjunto, a arquitetura deste circuito permite reproduzir diferentes estados do sistema anestésico simulado, facilitando a interação do estudante com os processos de controle e operação da máquina.

b.3) Projeto do circuito de controle geral e sensores

A arquitetura eletrônica do simulador baseia-se em um esquema de controle distribuído por meio de dois microcontroladores ESP32. O primeiro atua como módulo de controle principal (Figura 4), responsável pela gestão da interface gráfica e pela comunicação com o usuário, enquanto o segundo se dedica exclusivamente à execução de tarefas relacionadas a sensores e atuadores [7].

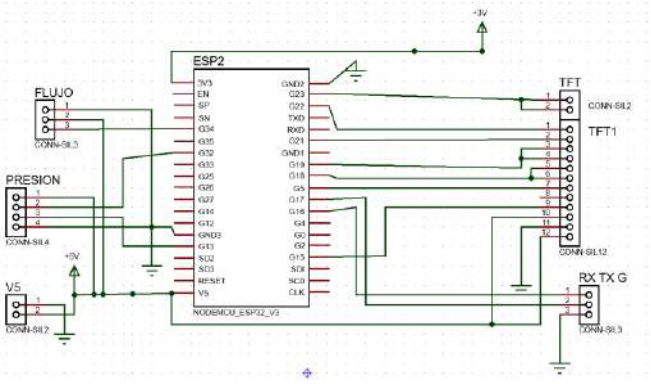


Fig. 4. Projeto esquemático do circuito de controle geral e sensores.

O sistema de sensoriamento permite monitorar variáveis relevantes do circuito respiratório simulado. Entre os elementos implementados estão um sensor de fluxo na linha inspiratória e sensores de pressão nas linhas inspiratória e expiratória, os quais permitem observar o comportamento do fluxo de gases dentro do sistema, parâmetros fundamentais para a análise do funcionamento de sistemas de ventilação anestésica [4].

A comunicação entre os dois microcontroladores é realizada por meio da interface UART RX–TX, garantindo a transmissão confiável de comandos entre o sistema de controle e o módulo de execução, método amplamente utilizado para comunicação entre dispositivos em sistemas embarcados [7].

A interação com o usuário é realizada por meio de uma tela TFT conectada ao ESP32 através da interface SPI, a qual permite visualizar estados do sistema, ativar funções do simulador e enviar comandos ao módulo de execução [13].

b.4) Projeto da estrutura

O projeto físico do simulador replica a organização estrutural e a aparência funcional de uma máquina de anestesia, integrando componentes eletrônicos, mecânicos e pneumáticos em um ambiente didático. A estrutura consiste em um gabinete retangular de aproximadamente 50 × 40 × 70 cm, comparável a equipamentos hospitalares compactos, cujo interior abriga a fonte de alimentação, módulos eletrônicos e a fiação interna de forma organizada.

Na parte superior integra-se uma plataforma de trabalho, além de nichos escalonados para sensores, atuadores e partes do circuito respiratório, e uma carcaça superior (cosy) com aberturas para mangueiras, conexões e ventilação. O sistema inclui ainda um compartimento traseiro para um tanque de oxigênio ou fonte externa de gás, permitindo realizar testes do circuito ventilatório sem interferir na área frontal de interação do usuário [1], [15].

C. Software e lógica de controle

c.1) Lógica de código para o ESP32 de controle

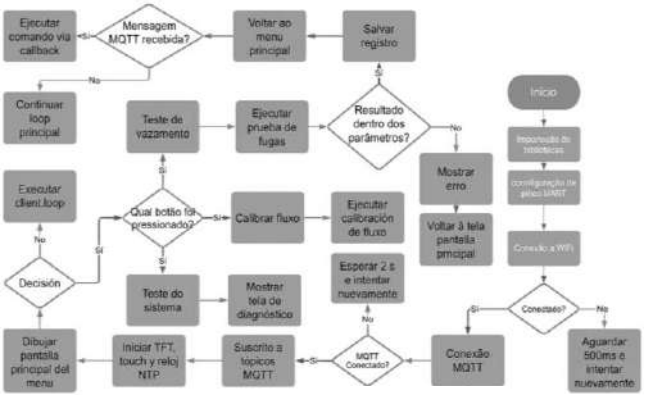


Fig. 5. Diagrama de fluxo do código para o ESP32 de controle.

O ESP32 de controle constitui o núcleo de interação do simulador, pois gerencia a interface gráfica na tela TFT, a comunicação sem fio via MQTT e o enlace serial UART com o ESP32 responsável pela execução dos atuadores. O software é organizado de forma modular, integrando funções de visualização, aquisição de dados, processamento de sensores e gerenciamento de eventos do usuário, permitindo a execução concorrente de tarefas sem comprometer o funcionamento geral do sistema [7].

A interface gráfica é implementada utilizando a biblioteca TFT_eSPI, que permite criar menus, cabeçalhos, painéis de informação e botões táteis a partir dos quais o usuário pode iniciar testes do simulador, como calibrações, diagnósticos ou verificações de funcionamento. Cada procedimento é implementado como uma tela independente com funções específicas para gerenciar a interação e validar os dados recebidos.

A comunicação remota é realizada por meio de conexão WiFi com um servidor MQTT, permitindo receber comandos de uma interface web e publicar periodicamente o estado do sistema. As mensagens recebidas são processadas por uma função de callback, que distingue entre comandos de operação e sinais de erro. Os comandos são encaminhados ao ESP32 de execução via UART, enquanto os erros geram avisos na interface do usuário[13].

As rotinas de teste, como calibração de fluxo e testes de vazamento, são implementadas utilizando estruturas de máquina de estados, que orientam o usuário durante cada procedimento. Essa abordagem permite estruturar processos sequenciais de forma controlada e evitar conflitos entre tarefas simultâneas dentro do sistema [7]. O ciclo principal do programa processa eventos da tela tátil, monitora a conexão MQTT, atualiza as informações do sistema e mantém a sincronização temporal via NTP, garantindo operação contínua do simulador.

c.2) Lógica de código para o ESP32 de execução

O ESP32 de execução é responsável pelo controle direto dos atuadores do simulador, incluindo eletroválvulas, o motor de passo controlado pelo driver DRV8825, o servomotor MG996R e os dispositivos de entrada do sistema respiratório simulado. Diferentemente do módulo de controle, este microcontrolador opera exclusivamente como módulo de execução de comandos recebidos do ESP32 principal por meio da interface UART2.

O software implementa um sistema de recepção de comandos baseado em frames delimitados, no qual cada instrução ativa uma função específica e utiliza flags de estado para evitar a execução simultânea de procedimentos críticos, como testes de vazamento ou calibração.

O controle das eletroválvulas é realizado por relés configurados como low-trigger ou high-trigger, utilizando a função relayWrite() para abstrair a ativação do hardware e controlar válvulas de diferentes tensões. O servomotor MG996R é controlado pela biblioteca ESP32Servo, que gera sinais PWM estáveis. O servo opera em posições predefinidas sincronizadas com o motor de passo que simula o movimento do fole respiratório, permitindo reproduzir dinâmicas de ventilação e condições de teste do simulador.

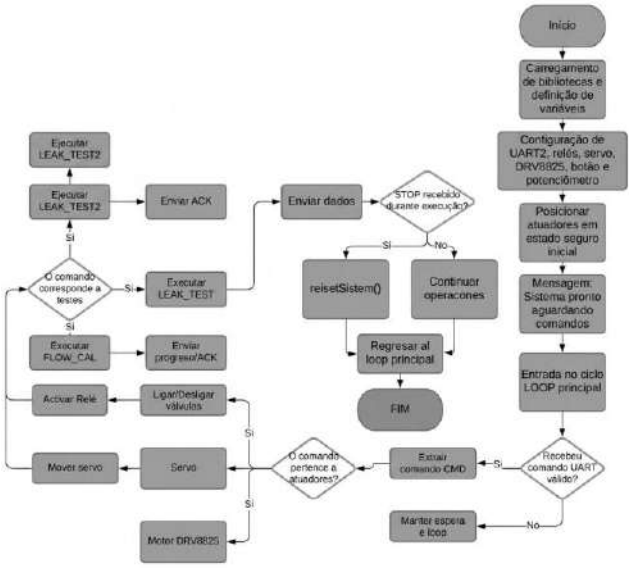


Fig. 6. Diagrama de fluxo do código para o ESP32 de execução.

IV. RESULTADOS

O desenvolvimento do simulador resultou em um sistema funcional que integra módulos eletrônicos, mecânicos e estruturais para a simulação didática de uma máquina de anestesia. A PCB de controle integrou o microcontrolador ESP32, circuitos de comunicação e interface do sistema, com trilhas de alimentação, comunicação e controle organizadas para reduzir interferências. De forma complementar, a PCB de execução foi projetada para controlar eletroválvulas, motor de passo e servomotor, considerando os requisitos de potência e corrente dos atuadores para garantir operação estável.

No aspecto mecânico, o simulador integra gabinete principal, plataforma superior e compartimentos para sensores, atuadores e elementos do circuito respiratório, mantendo organização da fiação e dos módulos internos.

Também foram realizados processos de calibração de sensores, especialmente do sensor de fluxo, utilizando um pulmão de teste de volume conhecido para validar seu comportamento na faixa de operação do sistema. Os resultados demonstram a viabilidade do simulador como ferramenta funcional para ensino e treinamento no uso de máquinas de anestesia.



Fig. 7. Simulador de máquina de anestesia

V. CONCLUSÕES

O desenvolvimento do simulador interativo de máquina de anestesia permitiu integrar componentes eletrônicos, mecânicos e de software em um sistema funcional voltado à formação de estudantes de Engenharia Biomédica. A arquitetura baseada em dois microcontroladores ESP32 constitui um dos elementos centrais do sistema, permitindo a separação das tarefas de controle e execução. O ESP32 de controle é responsável pela interface gráfica, pela comunicação remota e pelo gerenciamento das rotinas do simulador, enquanto o ESP32 de execução realiza o acionamento direto de sensores e atuadores.

A comunicação entre ambos os microcontroladores é realizada por meio de uma interface UART bidirecional, possibilitando a troca contínua de comandos e informações entre os dois módulos. Essa comunicação de ida e volta permite que o ESP32 de controle envie instruções de operação ao módulo de execução e, simultaneamente, receba dados de estado, medições de sensores e confirmações de execução. Dessa forma, garante-se a sincronização adequada entre as tarefas do sistema e uma maior estabilidade operacional do simulador.

O projeto eletrônico, a fabricação das PCBs e a calibração dos sensores de pressão e fluxo demonstraram desempenho adequado para a simulação do comportamento do circuito respiratório. A integração entre os módulos de hardware e software permitiu reproduzir diferentes condições de operação do sistema respiratório simulado, contribuindo para o funcionamento geral do protótipo.

Do ponto de vista educacional, o simulador permite reproduzir testes básicos de funcionamento e diagnóstico de uma máquina de anestesia, facilitando a compreensão da interação entre seus diferentes módulos sem a necessidade de utilizar gases anestésicos reais. Além disso, o sistema foi desenvolvido com custo acessível em comparação com simuladores comerciais, demonstrando a viabilidade do desenvolvimento local de ferramentas didáticas biomédicas por meio de um projeto modular, com potencial para futuras melhorias e ampliação de funcionalidades.

REFERÊNCIAS

[1] M. Green, R. Tariq, and P. Green, “Improving Patient Safety through Simulation Training in Anesthesiology: Where Are We?,” *Anesthesiol. Res. Pract.*, vol. 21, no. 1, p. 30962, 2016, doi: 10.1155/2016/4237523.

[2] A. S. Matveevskii and N. Gravenstein, “Role of simulators, educational programs, and nontechnical skills in anesthesia resident selection, education, and competency assessment,” *J. Crit. Care*, vol. 23, no. 2, pp. 167–172, Jun. 2008, doi: 10.1016/j.jcrc.2007.11.009.

[3] F. Lateef, “Simulation-based learning: Just like the real thing,” *J. Emerg. Trauma Shock*, vol. 3, no. 4, p. 348, 2010, doi: 10.4103/0974-2700.70743.

[4] I. S. Fischler, C. E. Kaschub, D. E. Lizdas, and S. Lampotang, “Understanding of Anesthesia Machine Function Is Enhanced With a Transparent Reality Simulation,” *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 3, no. 1, pp. 26–32, 2008, doi: 10.1097/SIH.0b013e31816366d3.

[5] X. Cai, X. Wang, Y. Zhu, Y. Yao, and J. Chen, “Advances in automated anesthesia: a comprehensive review,” *Anesthesiology and Perioperative Science*, vol. 3, no. 1, p. 3, Jan. 2025, doi: 10.1007/s44254-024-00085-z.

[6] M. J. Khodaei, N. Candelino, A. Mehrvarz, and N. Jalili, “Physiological Closed-Loop Control (PCLC) Systems: Review of a Modern Frontier in Automation,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2968440.

[7] E. A. . Lee and S. A. . Seshia, *Introduction to embedded systems : a cyber-physical systems approach*. MIT Press, 2017.

[8] J. D. Bronzino and D. R. Peterson, *The Biomedical Engineering Handbook*. CRC Press, 2018. doi: 10.1201/b18423.

[9] K. Jenkins *et al.*, “High-fidelity anesthesia simulation in medical student education: Three fundamental and effective teaching scenarios,” *Int. J. Acad. Med.*, vol. 3, no. 1, pp. 10–17, 2017, doi: 10.4103/IJAM.IJAM_45_17.

[10] C. Elendu, “Impact of simulation-based training in medical education: A systematic review,” *J. Educ. Health Promot.*, vol. 13, no. 2, pp. 112–120, 2024, Accessed: Feb. 07, 2026. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11224887/

[11] A. Fleet *et al.*, “Current Clinical and Educational Uses of Immersive Reality in Anesthesia: Narrative Review,” *J. Med. Internet Res.*, vol. 27, p. e62785, Mar. 2025, doi: 10.2196/62785.

[12] Paul. Horowitz and Winfield. Hill, *The art of electronics*. Cambridge University Press, 2024.

[13] Neil Kolban, *Kolban’s Book on ESP32*. 2018.

[14] Texas Instruments, “DRV8825 Stepper Motor Controller IC Datasheet.”

[15] R. A. . Jaffe, C. A. . Schmiesing, and Brenda. Golianu, *Anesthesiologist’s manual of surgical procedures*. Wolters Kluwer Health, 2014.

AUTORES



Marcelo Fabian Cruz Guizada é estudante de Engenharia Biomédica na Universidad Privada del Valle (Univalle), localizada em Cochabamba, Bolívia, onde ingressou em 2022. Também atua como pesquisador na instituição. Atualmente é vice-presidente da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB). Seus interesses estão voltados à pesquisa de novas tecnologias para melhorar a qualidade de vida das pessoas, contribuindo por meio da combinação de conhecimentos em engenharia, tecnologia e medicina.



Mauricio Marcelo Peredo Claros é mestre em tecnologias da informação e docente universitário na área de Engenharia Biomédica. Seu trabalho centra-se na integração de tecnologias emergentes, particularmente na adoção de inteligência artificial em processos educacionais e de pesquisa. Possui diversas publicações científicas em revistas internacionais, relacionadas com tecnologia educacional, inteligência artificial aplicada, telemedicina e desenvolvimento de soluções biomédicas. Participa ativamente em processos de pesquisa aplicada e inovação tecnológica, promovendo o uso de ferramentas digitais, sensores biomédicos e tecnologias de inteligência artificial na formação universitária. Suas principais linhas de interesse incluem inteligência artificial, inovação educacional, instrumentação biomédica e tecnologias para a melhoria da qualidade educacional.



Rodrigo Martinez Severich é Docente de Engenharia Biomédica da Universidade Privada do Vale UNIVALLE (2020), Gerente proprietário da empresa SEVERICH (2017), MBA pela Escola Europeia de Negócios EEN (2016), IMEX pela Escola Europeia de Negócios EEN (2016), Bacharel em Ultrassonografia Abdominal pelo Curso BachelorClass da Universidade Médica de Viena (2017), membro fundador do Colégio de Engenharia Biomédica SIB Cochabamba (2024) e CTO da WIVER, software especializado para gestão e visualização de imagens médicas (2024).

Development of a Low-Cost Elbow Passive Rehabilitation Orthosis with Digital Twin and IoT

Sergio Antonio Vásquez Martínez & Rodrigo Martínez Severich
Universidad Privada del Valle — Biomedical Engineering Program
vms0031149@est.univalle.edu, rmartinezs@univalle.edu
ORCID: 0009-0005-3475-2192, ORCID: 0009-0003-5588-7753

Abstract—The development and technical validation of a low-cost orthosis for passive rehabilitation of the elbow joint is presented. The device performs flexion–extension and pronation–supination movements through two servomotors controlled by an ESP32 microcontroller, with real-time angular monitoring provided by AS5600 and MPU6050 sensors. The architecture integrates a cloud-based web interface with a Three.js-based digital twin, wireless communication through the MQTT protocol, and a relational database for recording therapeutic sessions. Experimental validation, carried out through 50 no-load measurements distributed across five angles for each movement, yielded a mean angular error of 0.80° ± 0.76° in flexion–extension and 0.84° ± 0.85° in pronation–supination, with a maximum error of 2° in both cases, values within the clinical tolerance of ±5°. The safety system demonstrated 100% effectiveness. The total prototype cost was approximately USD 187, representing a reduction of more than 96% compared with commercial CPM devices. These results show that the system constitutes a technically viable and economically affordable alternative for resource-limited settings.

Index Terms— Digital twin, IoT, elbow orthosis, passive rehabilitation.

I. Introduction

Musculoskeletal disorders constitute one of the leading causes of functional disability worldwide [1]. Within this group, the elbow joint plays a critical role in upper limb functionality, enabling the flexion–extension and pronation–supination movements essential for daily activities [2]. Elbow joint stiffness, commonly occurring after surgery, trauma, or prolonged immobilization, compromises the patient's functional independence and requires sustained rehabilitation [3].

Therapeutic management prioritizes, in early stages, dosed passive mobilization under professional supervision [4]. Continuous passive motion (CPM) devices allow programming of range, speed, and duration parameters [5]; however, they present limitations in cost, portability, and availability that restrict their adoption in resource-limited settings. Previously reported academic mechatronic orthoses partially address these limitations, albeit with restrictions in degrees of freedom, connectivity, or automated data logging [6], [7].

In response to this situation, a two-degree-of-freedom ambidextrous orthosis, named REHABITEC, is proposed for passive elbow rehabilitation. The system integrates programmable actuators, angular position sensors with wireless transmission, a web interface with a digital twin, and a relational database. The main contribution lies in the convergence of a real-time digital twin, an MQTT-based IoT architecture, and a two-degree-of-freedom mechatronic system within a prototype whose cost (~187 USD) represents a reduction of over 96% compared to commercial CPM devices.

II. Related work

Commercial CPM systems such as the Artromot-E2 (Ormed GmbH) and Kinetec Centura (Kinetec S.A.) constitute the standard of reference in passive elbow rehabilitation. These devices operate exclusively in the flexion–extension plane, are priced between 5,000 and 15,000 USD, and lack wireless connectivity or integrated digital logging systems [5].

In the academic domain, Arcos Rosero et al. [6] designed a single-degree-of-freedom prototype using a stepper motor with Arduino-based control, demonstrating mechanical feasibility but without remote monitoring or a supervision interface. Gonzales and Furukawa [7] incorporated two articular movements with Bluetooth control, although without persistent data logging or three-dimensional visualization of device behavior.

Other contributions have explored the integration of inertial and magnetic sensors for angular monitoring in upper limb orthoses [8]. In the broader context of digital twins applied to rehabilitation, Wang et al. [9] developed a digital twin system for a lower limb exoskeleton that enables real-time visualization of patient-specific gait and remote monitoring of robot motion, demonstrating improved interaction between the device and the patient. Falkowski et al. [10] investigated the applicability of digital twins for home-based motor rehabilitation, proposing the concept of controlling a rehabilitation exoskeleton via its virtual replica and highlighting the potential of IoT-enabled digital twins for remote therapy supervision. In the orthopedic domain, Dean et al. [11] reviewed the integration of digital twin technologies with artificial intelligence for surgical planning, patient-specific outcome prediction, and simulation-based training, noting that applications in rehabilitation monitoring remain at an early stage. Despite these advances, the convergence of a real-time digital twin with an IoT architecture in a compact, low-cost orthosis oriented toward resource-limited settings remains an open area of development.

TABLE I
COMPARISON WITH RELATED WORK

FEATURE	COMMERCI -AL CPM	ARCOS R. ET AL.	GONZALES FURUKAWA	THIS WORK
DOF	1 (FE)	1 (FE)	2 (FE+PS)	2 (FE+PS)
Connectiv.	No	No	Bluetooth	Wi-Fi
GUI	No	No	App	Web
Database	No	No	Limited	MySQL
Ambidex.	No	No	N/R	Yes
Cost USD	>10 000	N/R	N/R	~187

DOF: degrees of freedom; FE: flexion–extension; PS: pronation–supination; N/R: not reported.

III. System Architecture

A. General description

REHABITEC consists of three interconnected modules: (1) the mechatronic structure of the orthosis, (2) the control and power electronics, and (3) the software platform with digital twin. Fig. 1 illustrates the overall block diagram of the system. The interaction begins with the web interface, where the clinician configures the therapy parameters. These are transmitted to the ESP32 microcontroller in the control stage via the MQTT protocol over Wi-Fi. The ESP32 generates PWM signals for the servomotors and receives feedback from the angular position sensors, whose data are sent back to the server for logging and real-time visualization.

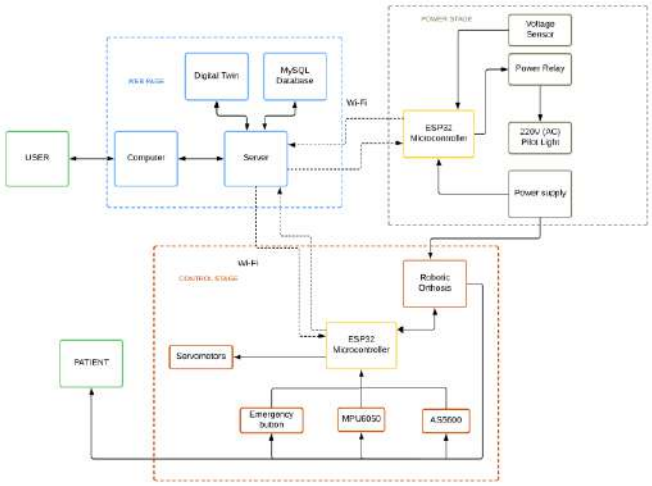


Fig. 1. Block diagram of the REHABITEC system.

B. Hardware stage

The electronic architecture is organized into two complementary subsystems: the power stage and the control stage. The power stage employs a dedicated ESP32 for electrical supply management. This module monitors the line voltage through a voltage sensor and controls a solid-state relay that switches the main power source. Additionally, it drives a 220 V pilot light that serves as a visual indicator of system status and alarm.

The control stage constitutes the logical core of the system. A second ESP32 receives the therapeutic parameters from the server and generates command signals for the servomotors. Simultaneously, it processes data from the AS5600 sensor (angular position in flexion–extension) and the MPU6050 (orientation in pronation–supination). The system incorporates an emergency button whose activation immediately halts the therapy, disables the actuators, and turns on the alert pilot light. The power supply was implemented using a bridge rectifier, a 12 V linear regulator, and a DC-DC step-down converter (XL4016) for 8.4 V, distributing power and signal lines on independent PCBs to minimize electromagnetic interference.

C. Actuator selection

The servomotor selection was based on the torque calculation required for each movement. For flexion–extension, considering a forearm–hand mass of 1.54 kg (2.2% of the body weight of a 70 kg adult [12]), a lever arm of 0.18 m, and a safety factor of 2.5, a required torque of 69.35 kg·cm was obtained. An 80 kg·cm servomotor (model DS5180) with PWM control was selected. For pronation–supination, with a tangential force of 5 N applied at a radius of 0.05 m and a safety factor of 3, the required torque was 7.65 kg·cm, leading to the selection of a 15 kg·cm servomotor (model DS3225). Both actuators integrate motor, gearbox, and position feedback in a compact form factor compatible with the microcontroller.

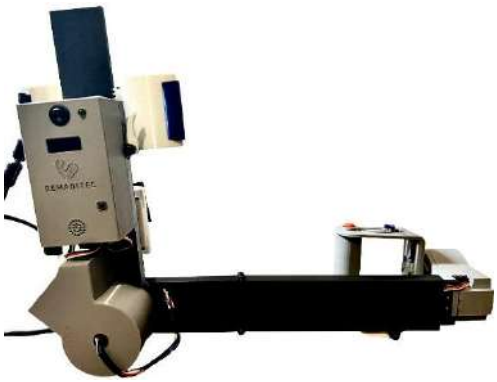


Fig. 2. Fully assembled REHABITEC prototype.

D. Sensing

The AS5600 magnetic sensor, a contactless measurement device, was used for angular position sensing on the flexion–extension axis. Its calibration was performed by comparison with a clinical goniometer, establishing a linear relationship between the I²C value range (0–4096) and the angular range (0°–135°). The MPU6050 inertial measurement unit, featuring a triaxial accelerometer and gyroscope, was used to estimate orientation on the pronation–supination axis. Calibration was carried out by obtaining offset values in a static position, applied directly in the ESP32 firmware.

E. Mechanical design

The orthosis structure was designed in CAD software and manufactured using 3D printing in PLA. The design comprises three main segments: the upper arm fixation module, the forearm support, and the distal hand-holding mechanism. Moving elements are coupled to the servomotors through rigid mechanical arms, while user contact areas incorporate curved surfaces to improve ergonomics. The orthosis includes a length adjustment mechanism in both the upper arm and forearm

segments, as well as a reversing system that enables ambidextrous use. Fig. 2 shows the fully assembled prototype.

F. Software and communication

The graphical interface was implemented as a web application developed with HTML5, CSS3, and JavaScript on the frontend, and Node.js on the backend. The system provides three access levels: physiotherapist (patient management and therapy programming), biomedical engineer (sensor calibration and technical monitoring), and administrator (full user and permission management). From the rehabilitation panel, the clinician selects the arm to be treated and configures the parameters for each movement: start and end angles, angular speed, and session duration. Fig. 3 presents the main page of the web application.

Communication between the web interface and the microcontrollers was implemented using the MQTT protocol, chosen for its low bandwidth consumption, its publish–subscribe model, and its native compatibility with the ESP32. The relational database, implemented in MySQL, stores patient information, sessions, programmed parameters, and angular readings sensed in real time. The system supports report export in PDF and CSV formats, facilitating clinical documentation and longitudinal follow-up of the therapeutic process.

From a usability perspective, the interface was designed following a task-oriented workflow that mirrors the clinician's natural sequence of actions: patient selection, arm configuration, parameter input, session execution, and report generation. Visual indicators such as color-coded connection status, real-time angular readouts, and clearly labeled start/stop controls were implemented to minimize cognitive load during therapy supervision. The three-level access model ensures that each user role only encounters the functions relevant to their responsibilities, reducing interface complexity. Although a formal usability evaluation with standardized instruments (e.g., System Usability Scale) was not conducted in this study, informal feedback from physiotherapists during prototype testing indicated that the workflow was intuitive and the learning curve was minimal.

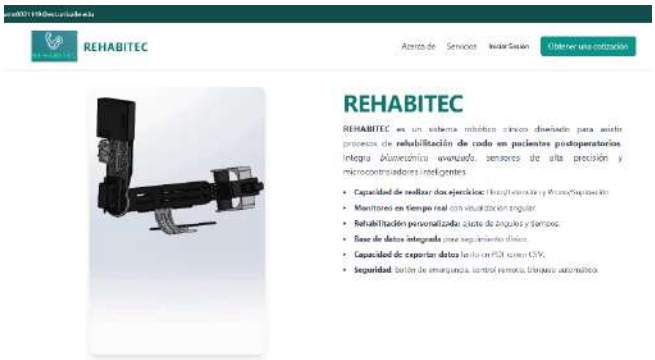


Fig. 3. REHABITEC system main page.

G. Digital twin

A digital twin of the orthosis was implemented using a three-dimensional model rendered with the Three.js library, integrated directly into the web interface. The model reproduces the programmed movements in real time, receiving angular data transmitted via MQTT from the ESP32 and updating the visual representation of the orthosis in synchronization with the physical execution. The 3D scene incorporates orbital camera controls that allow the physiotherapist to inspect the virtual device from different angles during the session. This component serves as a visual validation tool prior to applying the routine on the patient, enabling verification of the angular configuration and detection of potential parameterization errors before starting the session, thereby contributing to procedural safety. Fig. 4 displays the therapy programming interface with the integrated digital twin.

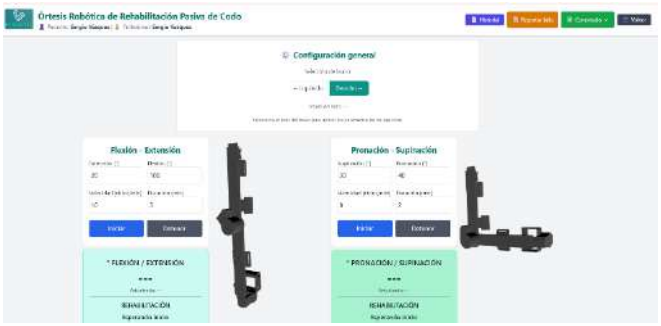


Fig. 4. Therapy programming interface with Three.js digital twin.

In the context of real-world medical applications, the digital twin component addresses a critical gap identified in rehabilitation practice: the lack of visual tools that allow clinicians to anticipate device behavior before applying therapy to the patient. In conventional CPM-based rehabilitation, the clinician programs parameters without a visual preview of the resulting motion, relying solely on numerical inputs. The proposed digital twin mitigates this limitation by providing a synchronized 3D representation that enables the physiotherapist to observe the complete range of motion, identify potential angular conflicts, and adjust parameters accordingly before physical execution. Furthermore, the integration of the digital twin with the MQTT-based IoT architecture and the relational database enables a comprehensive digital record of each session, supporting longitudinal analysis of patient progress and evidence-based clinical decision-making. While the current implementation serves primarily as a visualization and validation tool, its architecture is designed to accommodate future enhancements such as predictive modeling of therapeutic outcomes based on accumulated session data.

IV. Experimental Validation

A. Test protocol

A technical validation protocol was designed to evaluate the angular precision and repeatability of the system under controlled conditions. Tests were performed with the orthosis unloaded (no human load), mounted on a fixed surface, and with previously calibrated sensors. Five target angles were evaluated for each movement: 30°, 60°, 90°, 120°, and 135° for flexion–extension, and 30°, 60°, 90°, 120°, and 150° for

pronation–supination. Each angle was programmed from the web interface and executed five consecutive times, returning to the reference position (0°) between each repetition with a 3 s stabilization interval. The value reported by the corresponding sensor was recorded at each execution, accumulating a total of 50 measurements. Additionally, 5 emergency system tests were conducted.

B. Angular precision results

Table II presents the flexion–extension validation results. A global mean absolute error of 0.80° with a standard deviation of 0.76° was observed. The maximum recorded error was 2°, corresponding to the 30°, 60°, and 120° angles. The lowest mean error was obtained at 90° and 135° (0.40° in both cases).

TABLE II
ANGULAR PRECISION IN FLEXION–EXTENSION (AS5600)

Angle. (°)	Mean.err.	Std.dev.	Max.err.	Rep 2σ
30	1.00	0.71	2	1.41
60	1.40	0.89	2	1.79
90	0.40	0.55	1	1.10
120	0.80	0.84	2	1.67
135	0.40	0.55	1	1.10
Global	0.80	0.76	2	1.53

Table III shows the pronation–supination results. The global mean error was 0.84° ± 0.85°, with a maximum error of 2°. The most precise behavior was observed at 90° (mean error of 0.20°), while the 120° and 150° angles exhibited the highest mean error (1.20°).

TABLE III
ANGULAR PRECISION IN PRONATION–SUPINATION (MPU6050)

Angle. (°)	Mean.err.	Std.dev.	Max.err.	REP. 2σ
30	1.00	1.00	2	2.00
60	0.60	0.89	2	1.79
90	0.20	0.45	1	0.89
120	1.20	0.84	2	1.67
150	1.20	0.84	2	1.67
Global	0.84	0.85	2	1.70

In both movements, all measurements fell within the clinical tolerance of ±5° accepted for passive rehabilitation devices [4], with a maximum absolute error of 2° representing 1.48% of the total flexion–extension range (0°–135°) and 1.33% of the pronation–supination range (0°–150°).

C. Safety system

All five emergency button tests were successful (100% effectiveness). In every case, pressing the button produced immediate servomotor shutdown, power supply deactivation through the solid-state relay, and activation of the 220 V pilot light as an alarm indicator. No communication failures or perceptible delays in the stop system response were recorded. Fig. 5 shows the emergency system components, including the power supply unit and the 220 V pilot light.



Fig. 5. Emergency system with power supply unit and 220 V pilot light.

D. Cost analysis

Table IV presents the prototype cost breakdown, totaling approximately 187 USD. This figure represents a reduction of over 96% compared to commercial CPM devices, whose prices range between 5,000 and 15,000 USD [5]. The most significant cost items correspond to the servomotors (39% of total cost), which were selected based on the torque calculations presented in Section III-C, and the 3D printing material (19%), used for manufacturing the entire mechanical structure. The electronic control components, including the ESP32 microcontroller and both sensors, account for only 17% of the total cost, which reflects the accessibility of IoT-oriented embedded platforms currently available in the market.

It should be noted that the reported cost corresponds to the laboratory prototype and does not include certification, packaging, or serial production expenses. However, the use of open-source hardware platforms, widely available PLA filament, and standard communication protocols such as MQTT suggests that the total cost could be further reduced in a batch production scenario. The deliberate selection of locally available components constitutes a differentiating factor oriented toward system replicability in similar resource-limited contexts, where commercial CPM devices remain economically unfeasible for most rehabilitation centers.

TABLE IV
PROTOTYPE COST BREAKDOWN

COMPONENT	COST (USD)
DS5180 servomotor (80 kg·cm)	50
DS3225 servomotor (15 kg·cm)	22
ESP32 microcontroller	13
Sensors (AS5600, MPU6050)	9
12 V power supply	9
PLA material (3D printing)	35
Voltage regulators	20
Cables, connectors, PCB, hardware	29
Total	187

V. CONCLUSIONS

The development and technical validation of a two-degree-of-freedom orthosis for passive elbow rehabilitation was presented, integrating programmable actuators, angular sensing, wireless communication via MQTT, a web interface with a Three.js-based digital twin, and a relational database. The validation demonstrated a mean angular error below 1° in both movements, with a maximum error of 2°, within the clinical tolerance of ±5°. The safety system achieved 100%

effectiveness, and the prototype cost (~187 USD) represents a reduction of over 96% compared to commercial CPM devices.

The main contributions are: (a) the integration of two movement planes in a low-cost ambidextrous orthosis, with adjustment for right and left arm use and length regulation adaptable to different arm sizes, (b) a Three.js digital twin synchronized in real time with the physical device, (c) an IoT architecture enabling automated logging of therapeutic sessions, and (d) the demonstration of technical feasibility through a validation protocol with quantitative metrics.

Acknowledged limitations include: the unloaded validation without real biomechanical load, open-loop operation, and the absence of clinical trials with patients. Future work includes the implementation of closed-loop control, clinical trials with ethics committee approval, the incorporation of anomaly detection algorithms, and mechanical design optimization.

REFERENCES

[1] WHO, "Musculoskeletal health," 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions

[2] D. A. Neumann, Kinesiology of the Musculoskeletal System: Foundations for Rehabilitation, 3rd ed. Barcelona: Elsevier, 2022.

[3] A. Akhtar, B. Hughes, and A. C. Watts, "The post-traumatic stiff elbow: A review," J. Clin. Orthop. Trauma, vol. 19, pp. 125–131, 2021.

[4] B. Morrey and S. Sanchez, Morrey's The Elbow and Its Disorders, 4th ed. Philadelphia: Elsevier, 2013.

[5] Kinetec S.A., "Kinetec — Continuous passive motion devices," 2021. [Online]. Available: https://www.kinetec.fr/

[6] J. D. Arcos Rosero, D. C. Bolaños, L. F. Alape, A. F. Solis, and E. Roldán, "Mechatronic design of a prototype orthosis to support elbow joint rehabilitation," Bioengineering, vol. 9, no. 7, art. 287, 2022.

[7] A. Gonzales and R. Furukawa, "Development of a CPM device for elbow rehabilitation," in 2024 Latin American Robotics Symposium (LARS), pp. 1–7, 2024.

[8] R. Essaadaoui et al., "Construction of an innovative, modern and affordable teaching tool in practical classroom activities using AS5600 encoder," Phys. Educ., vol. 56, no. 6, art. 063006, 2021.

[9] W. Wang, Y. He, F. Li, J. Li, J. Liu, and X. Wu, "Digital twin rehabilitation system based on self-balancing lower limb exoskeleton," Technol. Health Care, vol. 31, no. 1, pp. 103–115, 2023.

[10] P. Falkowski et al., "Study on the applicability of digital twins for home remote motor rehabilitation," Sensors, vol. 23, no. 2, art. 911, 2023.

[11] B. Dean et al., "Leveraging digital twins for improved orthopaedic evaluation and treatment," J. Exp. Orthop., vol. 11, no. 4, art. e70084, 2024.

[12] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, 4th ed. Hoboken: Wiley, 2009.

VI. ACKNOWLEDGMENTS

The authors would like to thank Universidad Privada del Valle for the institutional support provided during the development of this project, and Ing. Edgar Ramos, Director of the Biomedical Engineering Program, for his guidance and support throughout this research.

AUTHORS



Sergio Antonio Vásquez Martínez is a senior Biomedical Engineering student at Universidad Privada del Valle (Univalle), Cochabamba, Bolivia. He joined the university in 2022, where he also works as a researcher. He serves as president of the Scientific Society of Biomedical Engineering Students (SCEIB). He is co-founder of the startup DocIA, a telemedicine project that integrates IoT kits for vital signs monitoring and artificial intelligence. His interests focus on projects aimed at improving quality of life and well-being through the application of Biomedical Engineering. He actively participates in research related to his field of study, exploring technological solutions that promote advances in health and well-being.



Rodrigo Martínez Severich is a Biomedical Engineering lecturer at Universidad Privada del Valle – UNIVALLE (2020), Owner-Manager of the company SEVERICH (2017), MBA from Escuela Europea de Negocios EEN (2016), IMEX from Escuela Europea de Negocios EEN (2016), Degree in Abdominal Ultrasound from the Medical University of Vienna (2017), founding member of the Higher School of Biomedical Engineering SIB Cochabamba (2024), and CTO of WIVER, specialized software for medical image management and visualization (2024).

Development of a Non-Invasive Hemoglobin Measurement System for Early Anemia Detection

Luis Fernando Arteaga Aldana & Alberto J Armando Aguilar Bustillos
Universidad Privada del Valle – Univalle
aal5000875@est.univalle.edu, aaguilarbu@univalle.edu
https://orcid.org/0009-0004-1695-6856

Abstract— Hemoglobin measurement is fundamental for anemia detection. In this work, two devices were developed for its determination: an invasive system based on spectrophotometry using Beer-Lambert's law, and a non-invasive system based on photoplethysmography using the MAX30102 optical sensor. Both prototypes employ an ESP32 microcontroller for data acquisition, processing, and visualization. The invasive system was validated with 75 blood samples compared to a reference hematology analyzer, obtaining an average error of 5.13%. The non-invasive system estimated hemoglobin from physiological parameters such as SpO_2, heart rate (BPM), perfusion index (PI), and R ratio, showing good agreement with the invasive method in experimental tests.

Index Terms— Hemoglobin, espectrophotometry, non-invasive device, optical sensors.

I. Introduction

Anemia is a global public health problem, mainly affecting vulnerable populations such as children and adolescents, especially in developing countries [1]. This condition is characterized by a decrease in the concentration of hemoglobin, a fundamental protein responsible for transporting oxygen in the blood, whose alteration directly impacts the proper functioning of the body [2], [3].

According to the World Health Organization (WHO), anemia is defined as hemoglobin levels below 12 g/dL in women and 13 g/dL in men [2].

The gold standard method for measuring hemoglobin is a complete blood count, which allows the different blood components to be analyzed using techniques such as flow cytometry. However, this procedure requires blood sampling, specialized infrastructure, and trained personnel, which limits its application in settings with limited access to health services [4], [5].

Given these limitations, non-invasive methods have gained relevance in the biomedical field, as they allow the concentration of hemoglobin to be estimated using optical technologies based on the interaction of light with biological tissues [4], [6]. These methods employ principles such as Beer-Lambert's law, widely used in biomedical optics to relate light absorbance to the concentration of substances in a medium [7].

In recent years, several non-invasive approaches for hemoglobin estimation have been explored, including photoplethysmography (PPG), multi-wavelength spectroscopy, and image-based analysis using digital cameras. These methods rely on the interaction of light with biological tissues to estimate hemoglobin concentration indirectly, offering advantages such as reduced discomfort, real-time monitoring, and applicability in low-resource settings. However, many of these techniques still face limitations in accuracy, calibration, and sensitivity to external factors such as skin pigmentation and motion artifacts.

In this context, this work proposes the design and implementation of two complementary systems for hemoglobin measurement: (i) an invasive system based on spectrophotometry using blood samples, and (ii) a non-invasive system based on PPG using optical sensors. The invasive system serves as a calibration and validation reference, while the non-invasive system provides a rapid and accessible estimation method. Both systems are integrated using an ESP32 microcontroller for data acquisition, processing, and communication, aiming to evaluate their performance and feasibility for early anemia detection.

II. METHODS

A. Device design

The device design is based on a modular architecture that integrates two independent systems for hemoglobin estimation: one invasive and one non-invasive. Both systems share a central processing unit responsible for data acquisition, control, and transmission, allowing for an organized and scalable structure.

As shown in Figure 1, the system incorporates a microcontroller as its control core, the system is based on an ESP32-WROOM-32 microcontroller, selected due to its dual-core architecture, integrated Wi-Fi connectivity, and sufficient processing capability for real-time biomedical signal acquisition, which manages the optical sensors, signal conditioning, and communication with an external platform. This microcontroller allows the synchronization of light signal emission and detection, as well as the initial processing of the acquired data. In addition, amplification and filtering circuits are implemented to improve the quality of the signal from the sensors, ensuring more stable and accurate measurements.

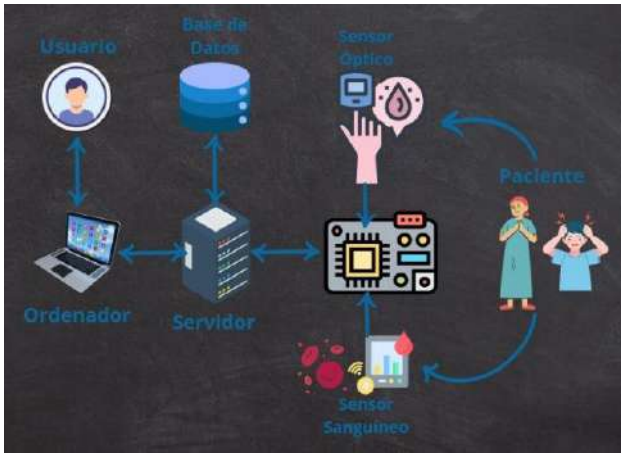


Fig. 1. General diagram of the proposed system.

At the hardware level, the electronic design is complemented by a physical structure that integrates all modules into a compact, 3D-printed casing, facilitating portability and use in clinical settings. The internal distribution of components was optimized to reduce interference and improve accessibility to the main elements of the system.

The physical implementation of the system is shown in Figure 2, where the integration of electronic modules within a 3D-printed enclosure can be observed.

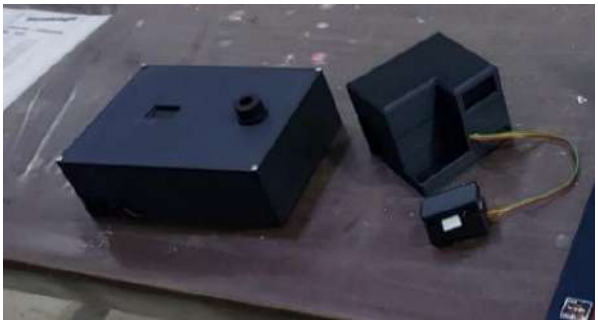


Fig. 2. Physical design of devices

The system is connected to a web-based platform developed using PHP for backend processing and MySQL as the database management system. This platform enables patient registration, structured data storage, and real-time visualization of results. Data transmission is carried out via Wi-Fi using HTTP requests to a local server environment, allowing efficient integration between the hardware device and the software interface.

The selection of the MySQL database was based on its reliability, scalability, and compatibility with web-based applications, allowing efficient structured storage of patient data and measurement records. Additionally, the use of a web platform enables remote access, data traceability, and improved interaction with the system compared to standalone devices.

The MAX30102 optical sensor was selected due to its high sensitivity, low power consumption, and integrated design, which simplifies hardware implementation while providing reliable physiological measurements such as oxygen saturation (SpO_2) and BPM. Similarly, the OPT101 sensor used in the invasive system integrates a photodiode and transimpedance amplifier, reducing external noise and improving signal stability.

The system does not rely on third-party cloud services; instead, a local server architecture was implemented to ensure data privacy, reduce latency, and maintain full control over patient information.

B. Invasive system

The invasive system is based on the principle of spectrophotometry, which allows the concentration of hemoglobin to be estimated by measuring the absorbance of light in a blood sample. This approach is based on Beer-Lambert's law, which establishes a direct relationship between light absorbance and the concentration of the analyte present in the solution [7].

The configuration of the invasive system is presented in Figure 3, illustrating the optical path and signal acquisition process.

In this system, a monochromatic light source is directed through a previously obtained blood sample, the light source operates within the visible spectrum, approximately in the range of 540–580 nm, corresponding to absorption peaks of hemoglobin. The incident radiation is partially absorbed by hemoglobin, and the intensity of the transmitted light is captured by a photodiode, which converts the light signal into a proportional electrical signal [6]. This signal is then conditioned by an amplification circuit, allowing it to be processed by the microcontroller.

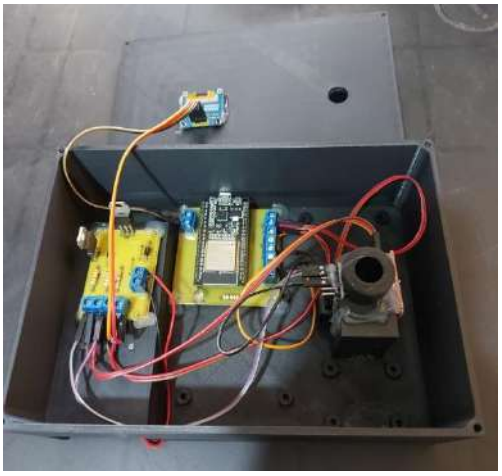


Fig. 3. Configuration of the invasive system based on spectrophotometry.

Hemoglobin concentration is calculated based on absorbance measurements, defined as:

$$A = \log\left(\frac{I_0}{I}\right) \quad (1)$$

Where A is the absorbance, I_0 is the intensity of the incident light, and I is the intensity of the transmitted light. This relationship allows the hemoglobin concentration to be estimated based on the optical absorption of the sample [7].

The system incorporates a light intensity sensor module based on the OPT101 component, which integrates a photodiode and a transimpedance amplifier in a single package, allowing direct conversion of the light signal into an amplified electrical signal. This module simplifies signal conditioning and reduces noise interference, facilitating its acquisition by the microcontroller for further processing.

Finally, the acquired data is processed by the microcontroller, which calculates the corresponding values and sends them to the display and storage system, facilitating subsequent clinical analysis.

b.1) Data acquisition and processing

Data acquisition in the invasive system is performed by a light intensity sensor module based on the OPT101 component [12], which integrates a photodiode and a transimpedance amplifier that converts light radiation into a voltage signal proportional to the detected light intensity. This analog signal is captured by the microcontroller through its analog-to-digital converter (ADC) for subsequent digital processing [8], the signal acquisition was performed at an approximate sampling frequency of 100 Hz to ensure stable measurements and adequate noise filtering.

The principle of spectrophotometry based on Beer-Lambert's law is used for sample analysis, which states that the absorbance of a substance is proportional to its concentration and the optical path of light in the sample [7]. Absorbance is calculated using equation 1.

In order to estimate the hemoglobin concentration, a calibration process was performed using samples with known concentrations. Based on the absorbance values obtained for these samples, a linear relationship between absorbance and hemoglobin concentration was determined. This relationship is expressed by the general equation:

$$Hb = C_1 * A + C_2 \quad (2)$$

where Hb corresponds to the estimated hemoglobin, A is the calculated absorbance, and C1 and C2 are coefficients obtained experimentally through the calibration process.

During the early stages of the system's development, an initial estimation equation was obtained, given by:

$$Hb = 160.31 * A + 0.9186 \quad (3)$$

These coefficients were progressively adjusted as new samples were incorporated into the calibration process, allowing for improved accuracy in hemoglobin estimation. The calibration curve obtained from experimental data is shown in Figure 4, demonstrating the linear relationship between absorbance and hemoglobin concentration.

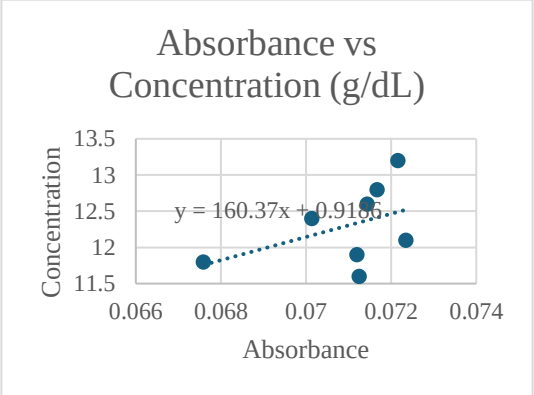


Fig. 4. Calibration curve for hemoglobin estimation.

Finally, the calculated values are processed by the microcontroller and sent to the recording system for storage and subsequent visualization within the web platform developed for the project.

C. Non-invasive system

The non-invasive system developed in this work is based on optical techniques for the indirect estimation of physiological parameters associated with hemoglobin, avoiding the extraction of blood samples. To do this, the MAX30102 sensor is used, an integrated module that incorporates red and infrared LED emitters together with a high-sensitivity photodetector, allowing the measurement of physiological signals using the principle of PPG. This sensor operates using red (660 nm) and infrared (940 nm) wavelengths, which are commonly used for physiological signal acquisition in pulse oximetry.

PPG is an optical technique widely used in biomedical devices to measure variations in blood volume within tissue. These variations produce changes in the intensity of light absorbed or reflected by the tissue, which can be detected by optical sensors and processed electronically [6], [9]. Because hemoglobin has different absorption characteristics depending on the wavelength used, the simultaneous analysis of red and infrared signals allows information related to oxygenation and other hemodynamic parameters to be obtained.

The MAX30102 sensor was integrated into the system through digital communication with the ESP32 microcontroller, which is responsible for acquiring and processing physiological signals in real time. From the PPG signal obtained, different physiological variables relevant to the analysis of the patient's circulatory status were extracted.

Among the variables obtained are:

- SpO_2, calculated from the ratio between the signals captured by the red and infrared channels of the sensor, representing the percentage of oxygenated hemoglobin present in the blood.
- BPM, determined by analyzing the pulsatile component of the PPG signal, corresponding to the number of beats per minute.
- PI, a parameter representing the ratio between the pulsatile component (AC) and the continuous component (DC) of the optical signal, allowing the quality of peripheral perfusion to be estimated.
- R (intensity ratio), defined as the quotient between the signals obtained in the red and infrared channels, used as an indicator of differential absorption in the tissue.

Based on these variables, an experimental model was proposed to estimate hemoglobin concentration using a multivariable linear equation, expressed as:

$$Hb = \beta_0 + \beta_1 SpO_2 + \beta_2 BPM + \beta_3 PI + \beta_4 R \quad (4)$$

where Hb corresponds to estimated hemoglobin and the β coefficients were determined experimentally during the calibration process using invasive measurements as a reference. During the tests carried out, it was observed that the coefficient associated with BPM had a negative value, indicating an inverse correlation between this parameter and the estimated hemoglobin levels in the samples analyzed.

From a physical standpoint, the device was designed with ergonomics, portability, and electronic integration in mind. The casing was modeled in three dimensions with the aim of organizing the internal components and facilitating its use by the patient. Inside, there is a main electronic board that incorporates the ESP32 microcontroller, the voltage regulation system, the OLED display screen, and the notification buzzer. The internal integration of the components is shown in Figure 5, where the main electronic modules are arranged within the device.

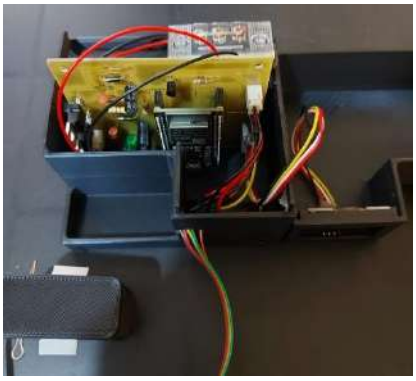


Fig. 5. Internal integration of the components of the non-invasive prototype.

A specific holder was also designed for the MAX30102 sensor, inspired by the operation of commercial pulse oximeters, allowing adequate contact between the user's finger and the optical sensor. This design reduces external interference and minimizes the influence of ambient light, improving measurement stability. The external design of the prototype is presented in Figure 6, highlighting the ergonomic structure developed for user interaction.



Fig. 6. Design of the non-invasive prototype casing.

c.1) Data acquisition and processing

Data acquisition and processing for the non-invasive system were carried out using the ESP32 microcontroller, which was responsible for managing the MAX30102 optical sensor readings, processing physiological signals, and estimating hemoglobin concentration, The sampling rate was configured between 50 and 100 samples per second depending on signal stability conditions. The firmware was designed in a modular way to ensure measurement stability and allow integration with the data display and storage system.

Initially, the system configures the peripherals and establishes a wireless connection with the remote server, allowing the transmission of the data obtained during the measurement. The ESP32 was selected due to its processing capacity and integrated Wi-Fi connectivity, features that facilitate communication with external applications and information storage systems [8], [10].

Once the system is initialized, the firmware proceeds to detect the presence of the finger on the MAX30102 sensor by monitoring the intensity of the infrared signal. When the correct placement of the finger is confirmed for a minimum time interval, the system initiates a stabilization phase in which the raw optical signals from the sensor are collected. During this stage, smoothing filters are applied to reduce noise and external interference, improving the quality of the photoplethysmographic signal obtained. To ensure reliable measurements, the system requires a minimum finger detection interval of approximately 3 seconds before initiating data acquisition. This stabilization period allows the sensor to reach a steady-state signal and reduces the influence of motion artifacts and improper finger placement.

Subsequently, the measurement process is organized into acquisition time windows, in which average values are calculated for the main physiological parameters derived from

the PPG signal: SpO_2, BPM, PI, and the optical intensity ratio R. These parameters are relevant variables for the indirect estimation of hemoglobin concentration, due to their relationship with blood flow dynamics and the optical absorption properties of tissue [6], [9].

To ensure the reliability of the results, the firmware incorporates a stability criterion based on the comparison of multiple consecutive windows. The relative variability of each physiological parameter is evaluated, and the signal is considered stable when the variation between windows remains within predefined limits. This procedure allows measurements affected by user movement, poor finger placement, or interference in the optical signal to be discarded.

Once measurement stability is achieved, the system proceeds to calculate the estimated hemoglobin using a previously calibrated experimental multivariate regression model (equation 4).

Finally, the calculated values are displayed on the device's OLED screen and transmitted to the web server via Wi-Fi connection, allowing them to be stored in a database for later analysis and visualization in the system interface. This process allows the integration of physiological data acquisition, local processing, and remote information management within the same biomedical monitoring architecture.

All procedures involving human participants were conducted following ethical standards. Informed consent was obtained from all participants prior to sample collection, and the data were anonymized to ensure patient confidentiality.

III. RESULTS AND DISCUSSION

The experimental validation of the system was carried out by comparing the measurements obtained with the prototypes developed and clinical reference values. For the invasive prototype, 75 venous blood samples obtained in the clinical laboratory of the Dr. Rubén Zelaya Municipal Hospital were analyzed, using the Zybio Z3 hematology analyzer as reference equipment. The measurements covered a range of approximately 10 g/dL to 16 g/dL, representative of normal clinical hemoglobin values.

The data obtained allowed the absolute, relative, and percentage errors between the prototype measurements and the reference equipment to be calculated. From the analysis of the 75 samples, an average percentage error of 5.13% was obtained, equivalent to an approximate efficiency of 94.87%. It was observed that the greatest discrepancies occurred in the first measurements, due to the initial calibration process of the linear equation used to estimate hemoglobin (Eq. 2). As more samples were incorporated into the experimental process, the coefficients of the equation were progressively adjusted, reducing the difference with respect to the hematology counter.

In Figure 7 shows the correlation graph between the measurements of the invasive prototype and the values obtained with the hematology counter. This figure shows that most of the points are concentrated around the identity line, indicating good agreement between the two methods within the concentration range evaluated.

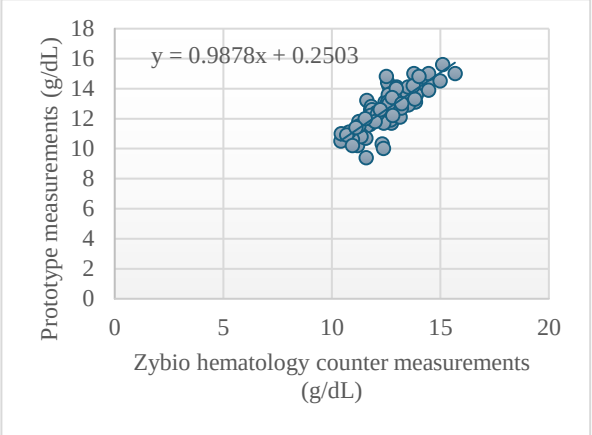


Fig. 7. Correlation graph of prototype measurements and hematology counter measurements.

For the non-invasive prototype, 20 experimental measurements were taken using the MAX30102 optical sensor based on PPG [11]. During each measurement, the physiological parameters SpO_2, BPM, PI, and R ratio were recorded and used as input variables in the multiple regression model defined in Eq. 4 to estimate hemoglobin concentration.

The estimated hemoglobin values obtained using the non-invasive system ranged from approximately 11.95 g/dL to 13.05 g/dL, with a mean of 12.54 g/dL and a standard deviation of 0.32 g/dL, indicating reduced variability in the estimates generated by the model.

To evaluate the agreement between the two systems, a direct comparison was made between the invasive prototype and the non-invasive prototype in a group of 10 participants, the results of which are summarized in Table I. This comparison yielded an average percentage error of 2.27%, equivalent to an approximate efficiency of 97.73%, with absolute errors ranging from 0.18 g/dL to 0.47 g/dL. The invasive prototype measurements were previously validated against a calibrated hematology analyzer (Zybio Z3), which serves as the reference standard. Therefore, the invasive system values presented in Table I can be considered clinically reliable for comparison with the non-invasive prototype.

The absolute error was calculated as the absolute difference between the invasive and non-invasive measurements. The relative error was obtained by dividing the absolute error by the invasive measurement value, which was previously validated against a reference hematology analyzer. The percentage error corresponds to the relative error multiplied by 100. It is important to note that the error values do not represent an average of both systems; instead, they quantify the deviation of the non-invasive measurements with respect to the invasive system, which is considered the reference for comparison.

TABLE I
COMPARISON TABLE OF MEASUREMENTS BETWEEN THE INVASIVE PROTOTYPE AND THE NON-INVASIVE PROTOTYPE

Patient	Value measured in the invasive prototype (g/dL)	Value measured in the non-invasive prototype (g/dL)	Absolute error	Relative error	Percentage error (%)
1	12,34	11,87	0,47	0,04	3,81
2	13,58	13,23	0,35	0,02	2,58
3	14,25	14,05	0,2	0,01	1,40
4	12,9	12,57	0,33	0,02	2,56
5	13,1	12,92	0,18	0,01	1,37
6	14,85	14,54	0,31	0,02	2,09
7	12,67	12,33	0,34	0,03	2,68
8	13,95	13,72	0,23	0,01	1,65
9	14,5	14,18	0,32	0,02	2,21
10	15	14,65	0,35	0,02	2,33
Average error of the prototype (%)			2,27		
Prototype efficiency (%)			97,73		

In Figure 8 shows the correlation graph between the measurements obtained using the invasive prototype and the non-invasive system. The observed upward linear trend shows a strong positive correlation between the two methods, indicating that the non-invasive system consistently reproduces the hemoglobin variations recorded by the invasive method.

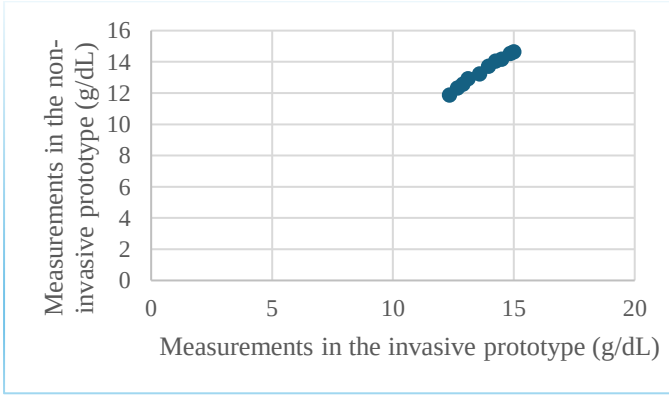


Fig. 8. Gráfica de correlación de las mediciones del prototipo invasivo y no invasivo

Overall, the experimental results show that the developed system performs stably within the evaluated physiological range. The observed agreement between invasive measurements, the reference clinical equipment, and the noninvasive system supports the feasibility of the proposed method for the approximate estimation of hemoglobin concentration using optical techniques.

IV. CONCLUSIÓN

This study presented the design and development of two complementary prototypes for hemoglobin measurement based on invasive and non-invasive approaches, integrating optical sensing, signal acquisition, and processing through an ESP32 microcontroller. The invasive system was experimentally validated using 75 blood samples and compared against a reference hematology analyzer, achieving an average percentage error of 5.13% and an estimated efficiency of 94.87% within a measurement range of 10 to 16 g/dL, demonstrating reliable performance for clinical reference purposes.

In contrast, the non-invasive system estimated hemoglobin concentration from physiological parameters obtained via photoplethysmography (PPG), including SpO_2, BPM, PI, and the R ratio. The estimated values ranged between 11.95 g/dL and 13.05 g/dL, with a mean value of 12.54 g/dL, indicating the feasibility of optical-based indirect hemoglobin estimation as a rapid and patient-friendly alternative.

A key contribution of this work lies in the integration of both approaches within a single framework. While the invasive system provides accuracy and serves as a validation reference, the non-invasive system enhances usability by eliminating the need for blood extraction. This hybrid methodology offers a balance between precision and accessibility, which is not commonly addressed in existing solutions that typically focus on only one measurement approach.

The comparative analysis between both systems yielded an average error of 2.27% and an approximate efficiency of 97.73%, indicating strong agreement and reinforcing the reliability of the proposed approach. These results highlight the potential of the system as a complementary tool for rapid hemoglobin estimation and preliminary anemia screening, particularly in low-resource or point-of-care environments.

Future work will focus on improving optical sensitivity, refining signal processing algorithms, and expanding clinical validation with a larger and more diverse population. These improvements are expected to enhance model accuracy and strengthen its applicability in real-world biomedical scenarios.

REFERENCES

[1] M. Miranda, M. G. Olivares, J. Durán-Pérez y F. Pizarro A., “Prevalencia de anemia y estado nutricional de escolares del área periurbana de Sucre, Bolivia,” 2015. [Online]. Avaliable: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-75182015000400001

[2] M. H. Ahmed, M. S. Ghatge y M. K. Safo, “Hemoglobin: Structure, Function and Allostery,” in Subcellular Biochemistry, vol. 94, Springer, 2020, pp. 345–382. doi: 10.1007/978-3-030-41769-7_14.

[3] N. Brandan, M. V. Aguirre y C. E. Giménez, “Hemoglobina,” 2008. [Online]. Available: https://docs.moodle.org/all/es/images_es/5/5b/Hemoglobina.pdf

[4] L. E. Brando De Angelis, Métodos no invasivos para la detección de hemoglobina en sangre, 2017.

[5] P. M. Torrens, “Cell blood count clinical interpretation,” Revista Médica Clínica Las Condes, vol. 26, no. 6, pp. 713–725, 2015. doi: 10.1016/j.rmclc.2015.11.001.

[6] B. P. Domínguez Capitaine, H. H. Cerecedo Núñez, P. Padilla Sosa y J. I. García Ramírez, “La interacción de la luz con moléculas,” Revista Digital Universitaria, vol. 21, no. 5, 2020. doi: 10.22201/cuaieed.16076079e.2020.21.5.1.

[7] I. Oshina y J. Spigulis, “Beer–Lambert law for optical tissue diagnostics: current state of the art and the main limitations,” Journal of Biomedical Optics, vol. 26, no. 10, 2021. doi: 10.1117/1.jbo.26.10.100901.

[8] Espressif Systems, ESP32 Datasheet, 2019. [Online]. Available: https://www.espressif.com

[9] A. Suhail, “¿Para qué se utilizan los sensores ópticos?,” 2024. [Online]. Available: https://www.azooptics.com/Article.aspx?ArticleID=2329

[10] O. A. Medellín Solís, Servidores en cómputo, 2016.

[11] Maxim Integrated, “MAX30102 Pulse Oximeter and Heart-Rate Sensor for Wearable Health,” Datasheet, 2020.

[12] Texas Instruments, “OPT101 Monolithic Photodiode and Single-Supply Transimpedance Amplifier,” Datasheet.

AUTHORS



Luis Fernando Arteaga Aldana holds a Bachelor's degree in Biomedical Engineering from the Universidad Privada del Valle (UNIVALLE) in Bolivia. His areas of interest include biomedical instrumentation, medical device design, and analog electronics. During his academic training, he participated in the development of projects aimed at the rehabilitation of conditions such as carpal tunnel syndrome, as well as in the study and construction of vacuum systems using the Venturi effect for learning and experimentation purposes. He is currently interested in researching medical imaging technologies and their various areas of application.



Alberto J. Armando Aguilar Bustillos holds a degree in Biomedical Engineering from the Universidad Privada del Valle (2012) and a certificate in Higher Education from the Universidad Pública de El Alto (2023). He currently serves as a professor at the Universidad Privada del Valle, where he teaches courses related to medical instrumentation and bioengineering projects. His professional work, backed by extensive experience in the management and maintenance of medical equipment within the Bolivian healthcare system, has focused on the management of highly complex hospital technology, with a distinguished track record at tertiary care institutions in Cochabamba, such as the “Manuel Ascencio Villarroel” Children’s Hospital and the “Germán Urquidi” Maternal and Child Hospital. His areas of interest include hospital electrical safety and quality management in healthcare services. His continuing education includes specializations in mechanical ventilation, hemodialysis, and hospital quality management under IBNORCA standards, with presentations at national conferences on sterilization and safe surgery, focusing his practice on improving the quality and safety of hospital medical equipment.

Ordem de Volatilidade como Fundamento Metodológico na Preservação de Evidências Digitais

Vinícius Santos Araújo, Guilherme Pedro Aquino, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
vinicius.araujo@inatel.br, guilhermeaquino@inatel.br, evandro.cesar@inatel.br

***Abstract*—This article analyzes the order of volatility as a structuring methodological foundation for the collection and preservation of digital evidence in forensic investigations. Based on the ontological nature of digital evidence and its high susceptibility to alteration or loss, the study articulates established normative frameworks, such as RFC 3227, ABNT NBR ISO/IEC 27037, and NIST guidelines, with technical-legal requirements related to integrity, authenticity, traceability, and chain of custody. The transition from the Dead Forensics paradigm to Live Forensics is examined, particularly in contemporary scenarios involving full disk encryption, fileless malware, and advanced persistent threats. Strategies for prioritizing evidence collection in different operational contexts are discussed, as well as best practices and recurring errors that may affect evidentiary admissibility. It is concluded that the order of volatility is not merely an operational guideline, but an essential technical axiom for ensuring informational preservation and the legal validity of digital evidence.**

***Index Terms*—Order of volatility, digital evidence, digital forensics, chain of custody, incident response.**

***Resumo*—O presente artigo analisa a ordem de volatilidade como fundamento metodológico estruturante na coleta e preservação de evidências digitais em investigações forenses. Partindo da natureza ontológica da evidência digital e de sua elevada suscetibilidade à alteração ou perda, o estudo articula bases normativas consolidadas, como a RFC 3227, a ABNT NBR ISO/IEC 27037 e as diretrizes do NIST, com exigências técnico-jurídicas relacionadas à integridade, autenticidade, rastreabilidade e cadeia de custódia. Discute-se a transição do paradigma de Dead Forensics para Live Forensics, especialmente diante de cenários contemporâneos envolvendo criptografia de disco inteiro, malwares fileless e ameaças persistentes avançadas. São examinadas estratégias de priorização da coleta em diferentes contextos operacionais, bem como boas práticas e erros recorrentes que impactam a admissibilidade probatória. Conclui-se que a ordem de volatilidade não constitui mera diretriz operacional, mas axioma técnico indispensável à preservação informacional e à validade jurídica da prova digital.**

***Palavras chave*—Ordem de volatilidade, evidência digital, computação forense, cadeia de custódia, resposta a incidentes.**

I. Introdução

A expansão das tecnologias de informação e comunicação e a digitalização de infraestruturas críticas e governamentais consolidaram o ciberespaço como domínio estratégico de relevância geopolítica, econômica e militar [1]. Incidentes cibernéticos assumiram caráter estruturado, persistente e multifásico, envolvendo atores estatais e operações coordenadas [2]. Nesse contexto, a identificação, preservação e análise da evidência digital tornou-se central tanto para a persecução penal quanto para a resposta a incidentes [3].

A evidência digital é o único vínculo material capaz de demonstrar autoria, materialidade e extensão do ilícito [3]. Diferente das evidências físicas, os artefatos digitais são representações binárias dependentes de estados transitórios de *hardware* e *software*, sendo altamente suscetíveis a alteração, sobrescrita ou destruição [4]. O princípio de Locard mantém aplicabilidade no ambiente computacional [5], todavia, a própria atividade investigativa pode modificar o estado do sistema examinado [6]. Comandos triviais, a interação com o sistema ou o desligamento de um equipamento comprometido implicam alterações em metadados, encerramento de sessões e descarte do conteúdo em memória volátil [7].

Nesse cenário, o desafio técnico de equipes de *Digital Forensics and Incident Response* (DFIR) e peritos reside em conduzir a coleta de dados sem comprometer a integridade do vestígio [8]. Procedimentos inadequados resultam na perda irreversível de elementos probatórios, comprometendo a admissibilidade judicial e a confiabilidade analítica em investigações [9]. A prevalência de *malwares fileless* e a adoção de criptografia de disco inteiro (*Full Disk Encryption*, FDE) reduziu a janela temporal para a aquisição de dados críticos, como chaves criptográficas e processos maliciosos em memória [10].

Diante dessa problemática, impõe-se a ordem de volatilidade como princípio metodológico da atividade pericial [7]. A ordem de volatilidade não é meramente orientativa, mas um axioma operacional que estabelece a precedência técnica: a coleta deve iniciar pelos componentes de maior transitoriedade e seguir aos elementos de maior persistência [7]. Tal abordagem visa mitigar a perda de informações sujeitas a rápida degradação, preservando dados dependentes de estados específicos de execução [10].

Este artigo analisa a ordem de volatilidade como fundamento da coleta de evidências digitais, articulando bases normativas e exigências operacionais. Busca-se aproximar diretrizes internacionais dos desafios práticos enfrentados por *First Responders* em ambientes heterogêneos. As Seções II a VI apresentam os fundamentos conceituais e normativos, a aplicação metodológica, boas práticas e erros recorrentes, desafios atuais e as considerações finais.

II. Fundamentos da Evidência Digital e Ordem de Volatilidade

A aplicação da ordem de volatilidade exige compreender a natureza da evidência digital e os requisitos técnicos e jurídicos de sua validade [3]. Na computação forense, a evidência digital

é qualquer dado gerado, processado, armazenado ou transmitido por sistemas computacionais com potencial probatório para sustentar ou refutar hipóteses investigativas [3].

Diferente da evidência física, a digital é lógica e intangível, dependendo de suporte eletrônico para manifestar-se [4]. Expressa-se por estados binários em múltiplas camadas, envolvendo *hardware*, *firmware*, sistema operacional e aplicações [4]. Tal estrutura implica que procedimentos de identificação e aquisição podem interferir no ambiente, exigindo rigoroso controle técnico pericial [6].

Normativamente, a ABNT NBR ISO/IEC 27037:2013 estabelece diretrizes para *Digital Evidence First Responders* (DEFRs) e *Digital Evidence Specialists* (DESs), disciplinando a identificação, coleta e preservação de vestígios [8]. Para o dado adquirir relevância probatória, deve observar princípios que assegurem confiabilidade metodológica [8].

A. Princípios básicos para confiabilidade metodológica

a.1) Auditabilidade

A auditabilidade exige documentação integral e cronológica das ações sobre o vestígio [8]. Registram-se instrumentos, versões de software, parâmetros, comandos e justificativas decisórias [9]. Isso permite a reconstrução da cadeia procedimental por terceiro habilitado [8], assegurando conformidade, transparência e controle de qualidade pericial [4].

a.2) Repetibilidade

A repetibilidade é a obtenção de resultados invariáveis pelo mesmo operador, sob idênticos procedimentos e ambiente controlado [4]. Relaciona-se à consistência das ferramentas e estabilidade dos métodos, especialmente via *hashing* para validar a integridade [11]. Tal invariância indica robustez metodológica [11].

a.3) Reprodutibilidade

A reprodutibilidade exige que operadores distintos, com ferramentas diversas, alcancem resultados equivalentes sobre a mesma evidência original [3]. Essa propriedade assegura independência quanto ao perito ou plataforma [8]. Por exemplo, imagens forenses do *FTK Imager* podem ser validadas por ferramentas como *EnCase* ou *Guymager*, preservando valores de *hash* e estruturas extraídas [12], o que reforça o caráter científico da prova [4].

a.4) Justificabilidade

A justificabilidade impõe intervenções tecnicamente fundamentadas e proporcionais ao risco, especialmente em sistemas em execução [8]. Decisões que alterem o estado operacional devem basear-se em critérios objetivos e boas práticas [9]. A validade jurídica depende da **Integridade**, **Autenticidade** e **Rastreabilidade** [3]. A integridade é assegurada por funções como SHA-256 ou SHA-512, cujos valores de *hash* comprovam a ausência de alterações [13]. A autenticidade vincula o dado à sua origem em um marco temporal específico [8]. A rastreabilidade exige registro contínuo da cadeia de custódia, prevenindo questionamentos sobre manipulação [14].

Nesse contexto surge a **Volatilidade**: o grau de suscetibilidade de um dado à alteração ou perda pelo tempo e estado operacional [7]. Registradores, RAM e tabelas de roteamento sofrem sobrescrita constante pelo sistema operacional [7]. A volatilidade impacta a atividade forense, pois artefatos como chaves criptográficas, cargas maliciosas e conexões de rede podem residir apenas em memória ativa [10]. A interrupção da energia implica perda imediata desses dados [7]. Assim, priorizar a coleta conforme a volatilidade decorre da dinâmica computacional [7]. Preservar elementos degradáveis maximiza a retenção informacional e reduz o risco de supressão de evidências essenciais à reconstrução do incidente [12].

B. Ordem de Volatilidade e Base Normativa

O eixo conceitual das práticas contemporâneas de aquisição e preservação de evidências digitais deriva da RFC 3227, intitulada *"Guidelines for Evidence Collection and Archiving"* [7]. Este documento consolidou diretrizes para administradores de sistemas, equipes de resposta e peritos, estabelecendo um referencial metodológico amplamente adotado em DFIR [7].

A RFC 3227 formaliza que a aquisição de evidências deve seguir um critério de precedência baseado na suscetibilidade dos dados à alteração ou perda [7]. Operacionalmente, isso implica iniciar a coleta por artefatos de permanência temporal reduzida e avançar aos elementos de maior persistência [7]. Nesse contexto, a "ordem de volatilidade"é a hierarquização taxonômica de componentes de *hardware* e *software* segundo a expectativa de retenção de dados frente à atividade do sistema, intervenção pericial ou ações de evasão [7]. A taxonomia organiza os artefatos em sete níveis decrescentes de volatilidade, conforme a Tabela I [7].

A observância desta diretriz visa mitigar a sobrescrita e a perda informacional decorrentes da dinâmica sistêmica e da própria intervenção investigativa [7]. Sistemas operacionais são ambientes dinâmicos; comandos rotineiros ou ferramentas de mitigação podem alterar metadados e sobrescrever vestígios recentes [4]. Ao seguir a ordem de volatilidade, prioriza-se a preservação de dados suscetíveis à degradação, postergando a manipulação de elementos persistentes até que as camadas transitórias tenham sido coletadas [7].

III. Aplicação Metodológica da Ordem de Volatilidade

A efetividade da ordem de volatilidade depende de sua integração aos processos formais de resposta a incidentes [9]. As diretrizes do National Institute of Standards and Technology (NIST), especialmente a SP 800-61 Revision 3 (*Incident Response Recommendations and Considerations for Cybersecurity Risk Management*), estabelecem que contenção, erradicação, recuperação e preservação de evidências são atividades interdependentes [9]. A coleta probatória deve ser planejada como componente estruturante da gestão do incidente, proporcional à criticidade do ativo e ao impacto do evento [9].

A. Live Forensics vs. Dead Forensics sob a Ótica da Volatilidade

Historicamente, a prática forense digital foi estruturada sob o paradigma da *Dead Forensics*, que se fundamentava na premissa de que a interrupção imediata do fornecimento de energia ao sistema comprometido preservaria o estado físico

TABELA I
ESTRATOS DA ORDEM DE VOLATILIDADE SEGUNDO A RFC 3227

Nível	Componentes	Características e Implicações Forenses
Extremamente Volátil	Registradores da CPU; Memórias Cache (L1, L2, L3)	Estados momentâneos de processamento. Persistência em nanossegundos ou microssegundos. Aquisição por *software* é praticamente inviável, exigindo instrumentação de baixo nível.
Altamente Volátil	Tabelas de Roteamento; Cache ARP; Tabelas de Processos; Estatísticas de *Kernel*; Memória RAM	Processos ativos, conexões de rede, bibliotecas carregadas e possíveis chaves criptográficas. Essencial em casos *fileless*. Requer coleta ao vivo com ferramentas de *memory acquisition*.
Volatilidade Moderada	TMPFS; Memória Virtual; Swap (`pagefile.sys`, `swapfile.sys`, `hiberfil.sys`)	Dados de paginação e fragmentos de processos descarregados da RAM. Úteis em análises *post-mortem*.
Baixa Volatilidade	Sistemas de Arquivos em HDD/SSD	Arquivos persistentes e metadados. Sensíveis a alterações por leitura. Requer uso de *write-blockers*.
Estável	Logs Remotos; Monitoramento de Rede	Registros externos ao host comprometido. Maior resistência à adulteração local.
Muito Estável	Configurações Físicas; Infraestrutura; Topologia de Rede	Informações estruturais que auxiliam na contextualização técnica do incidente.
Estático	Mídias de Arquivamento; Backups; *Cold Storage*	Maior persistência temporal. Fundamentais para recuperação e comparação com *baseline*.

do meio de armazenamento persistente, evitando alterações subsequentes no disco rígido [12]. O procedimento operacional clássico consistia no desligamento abrupto do equipamento (*pull the plug*), seguido da remoção do dispositivo de armazenamento, sua conexão a bloqueadores de escrita (*write-blockers*) de *hardware* e a geração de imagem forense bit a bit em ambiente laboratorial controlado [4].

Esse modelo mostrou-se adequado em cenários nos quais os artefatos de interesse encontravam-se predominantemente armazenados em mídia não volátil [4]. Contudo, a evolução das arquiteturas de proteção de dados e das técnicas ofensivas tornou essa estratégia insuficiente [15]. A adoção massiva de mecanismos de criptografia de disco inteiro, como BitLocker, FileVault e LUKS, alterou significativamente a dinâmica da preservação probatória [3]. Em sistemas criptografados e desbloqueados, a chave de volume máster (*Volume Master Key*, VMK) permanece residente em memória RAM para viabilizar a operação transparente do sistema de arquivos [15]. A interrupção abrupta da energia provoca o descarte imediato dessa chave, tornando o conteúdo do disco inacessível mesmo diante da posse física do dispositivo [15]. Assim, o desligamento precipitado pode inviabilizar integralmente a análise do conteúdo persistente, convertendo o meio de armazenamento em massa criptografada sem viabilidade prática de descriptografia [10].

Em resposta a essa limitação estrutural, consolidou-se a doutrina da *Live Forensics*, cuja lógica operacional materializa o princípio da ordem de volatilidade [16]. Nessa abordagem, a coleta inicia-se enquanto o sistema permanece energizado e operacional, priorizando a aquisição de dados residentes em memória e outros artefatos efêmeros [10]. Utilizando mídias externas confiáveis e ferramentas forenses especializadas, o perito realiza o *dump* da memória física e a captura estruturada de informações transitórias antes de proceder ao isolamento ou desligamento do equipamento [10].

Reconhece-se que a própria execução da ferramenta de aquisição introduz modificações mínimas no estado do sistema [6]. Contudo, tais alterações são consideradas metodologicamente aceitáveis diante da alternativa de perda definitiva dos dados voláteis [8]. Trata-se de decisão técnica pautada em análise de risco, na qual se privilegia a preservação informacional crítica em detrimento da imutabilidade absoluta do ambiente [9].

B. Priorização em Cenários Comuns de Incidentes

A aplicação prática da ordem de volatilidade não se apresenta de forma rígida ou uniforme [7]. Sua operacionalização deve considerar o contexto técnico, a criticidade do ativo e o impacto organizacional do incidente [9]. Destacam-se, a seguir, cenários recorrentes que demandam estratégias diferenciadas de priorização [12].

Em ambientes de missão crítica, tais como controladores de sistemas SCADA, infraestruturas hospitalares, centrais de processamento financeiro ou roteadores de núcleo, a interrupção do serviço pode produzir efeitos sistêmicos graves [17]. Nesses contextos, a adoção de *Dead Forensics* revela-se operacionalmente inviável [17]. A estratégia adequada consiste na execução de coleta ao vivo com mínima interferência operacional [16]. Soluções de *endpoint detection and response* (EDR) podem ser utilizadas para capturar instantâneos de memória e exportar configurações relevantes sem descontinuar serviços essenciais [12]. Em ambientes virtualizados, recursos do hipervisor permitem a realização de *snapshots* consistentes de memória (vRAM) e de discos virtuais (vDisk), com interrupções controladas e de curta duração [18]. Essa abordagem preserva tanto a disponibilidade quanto a materialidade probatória [18].

Em incidentes envolvendo *malwares* sofisticados e ameaças persistentes avançadas (APTs), a memória volátil assume papel central na reconstrução da cadeia de execução maliciosa [10]. Muitas famílias contemporâneas de código malicioso operam por meio de técnicas *fileless*, realizando injeção de código em processos legítimos e executando cargas úteis exclusivamente em memória, sem persistência em disco [10]. Em consonância com a ordem de volatilidade, procede-se inicialmente à aquisição do *dump* de memória mediante ferramentas

especializadas, como FTK Imager, DumpIt, LiME ou WinPmem [10]. A imagem resultante é posteriormente submetida a análise estruturada por meio de *frameworks* dedicados, como o *Volatility 3* [10].

A análise pode incluir o mapeamento da árvore de processos por meio do módulo `windows.pstree`, visando identificar relações hierárquicas anômalas, por exemplo, instâncias de `cmd.exe` derivadas de processos de serviço como `spoolsv.exe` [10]. Em seguida, a utilização de `windows.netscan` permite enumerar conexões TCP/UDP ativas e identificar potenciais canais de Comando e Controle (C2) [10]. A investigação de injeção de código pode ser conduzida com o módulo `windows.malfind`, que examina descritores de memória virtual (*Virtual Address Descriptor*, VAD) em busca de regiões com permissões simultâneas de leitura, escrita e execução (PAGE_EXECUTE_READWRITE), padrão frequentemente associado a cargas maliciosas injetadas dinamicamente [10]. A extração dessas regiões possibilita posterior engenharia reversa e atribuição técnica do artefato [19].

Em investigações envolvendo possível exfiltração de dados sensíveis, a coleta deve priorizar vestígios transacionais e registros de atividade de usuário [3]. Elementos como tabelas de roteamento, *cache* de navegação, histórico de conexões e listas de arquivos recentemente utilizados (MRU) podem evidenciar padrões de compactação massiva de diretórios e transferência para serviços externos de armazenamento em nuvem [20]. Adicionalmente, estruturas do Registro do Windows associadas ao histórico de dispositivos USB (USBSTOR) fornecem subsídios para identificação de mídias removíveis conectadas ao sistema, com respectivos identificadores de hardware e datas de utilização [20]. Tais artefatos, embora parcialmente persistentes, podem sofrer sobrescrita ou alteração ao longo do uso continuado do sistema, justificando sua coleta prioritária [20].

C. Documentação e Cadeia de Custódia

A coleta tecnicamente adequada, ainda que executada sob os parâmetros rigorosos da *Live Forensics*, carece de validade jurídica se não estiver acompanhada de formalização adequada da cadeia de custódia [21]. A admissibilidade da prova digital depende da demonstração inequívoca de que o conteúdo apresentado em juízo corresponde, de forma íntegra e inalterada, ao vestígio originalmente coletado [14]. No ordenamento jurídico brasileiro, a Lei nº 13.964/2019 introduziu os artigos 158-A a 158-F no Código de Processo Penal, positivando o conceito de cadeia de custódia [14]. O art. 158-A estabelece que esta consiste no conjunto de procedimentos destinados a manter e documentar a história cronológica do vestígio, assegurando a rastreabilidade desde o reconhecimento até o descarte [14].

Em se tratando de evidências digitais voláteis, a documentação deve abranger a totalidade das interações realizadas durante a aquisição: comandos executados, horários exatos, ferramentas utilizadas, parâmetros aplicados e valores de *hash* gerados [8]. A ausência de comprovação técnica de integridade, notadamente mediante aplicação e verificação de funções criptográficas como SHA-256, pode caracterizar quebra da cadeia de custódia [21], ensejando questionamentos quanto à autenticidade e à licitude da prova [14].

A jurisprudência consolidada reconhece que falhas na preservação da integridade e na rastreabilidade comprometem a confiabilidade probatória, podendo resultar na exclusão do elemento de convicção sob fundamentos correlatos à chamada "Teoria dos Frutos da Árvore Envenenada" [21]. Em síntese, a ordem de volatilidade e a cadeia de custódia não se apresentam como institutos concorrentes, mas complementares [14]. A primeira orienta tecnicamente a sequência de aquisição para maximizar a retenção de dados críticos; a segunda assegura que o produto dessa aquisição seja juridicamente válido, auditável e apto a sustentar responsabilização civil ou penal [21].

IV. Boas Práticas e Erros Comuns na Coleta de Evidências Digitais

A maturidade de equipes de resposta e a robustez dos laudos dependem da capacidade dos operadores de internalizar boas práticas e evitar falhas metodológicas [9]. Sob pressão por erradicação e disponibilidade, decisões precipitadas frequentemente comprometem a preservação probatória [9]. A atuação exige disciplina técnica, priorização racional e aderência a protocolos estabelecidos [8].

A. Respeito Hierárquico à Volatilidade

A coleta deve observar a precedência da ordem de volatilidade [7]. Assim, estruturas transitórias (tabelas de *Kernel*, conexões ativas e RAM) devem ser preservadas antes de intervenções em mídias persistentes [7]. O *dump* de memória precede imagens de disco ou inspeções físicas [10]. O planejamento estratégico deve anteceder a execução, com definição de sequência, ferramentas e escopo [9].

Interromper abruptamente a energia é uma das falhas mais graves em incidentes contemporâneos [12]. O desligamento elimina dados dos níveis 1 a 3 da RFC 3227, incluindo estados de execução e artefatos em memória [7]. Em cenários de técnicas *fileless* ou criptografia total, tal conduta pode inviabilizar a reconstrução técnica [15].

B. Execução de Ferramental Confiável

Ferramentas de aquisição requerem mídias externas confiáveis, idealmente protegidas contra gravação [8]. Binários estáticos e verificados minimizam a dependência de bibliotecas do sistema-alvo [6]. Suítes forenses dedicadas, como o *The Coroner's Toolkit* ou utilitários portáteis validados devem operar em isolamento, evitando contaminação cruzada [6].

Comandos nativos (`netstat`, `ps`, `ls`) podem gerar resultados manipulados sob *rootkits* ou interceptação de chamadas [19]. *Hooking* em *Kernel* ou *userland* pode alterar a saída, mascarando processos ou arquivos maliciosos [19]. Deve-se privilegiar ferramentas independentes [8].

C. Integridade da Linha do Tempo (Timelining)

A sincronização temporal é central para reconstruir eventos [4]. No início da coleta, deve-se registrar a diferença entre o relógio do sistema (BIOS/OS) e o Tempo Coordenado Universal (UTC), documentando fuso horário e inconsistências

[4]. Isso permite corrigir desvios e alinhar múltiplas fontes de log [9].

Negligenciar o *clock drift* inviabiliza a correlação entre logs de diferentes dispositivos [12]. Divergências temporais prejudicam a reconstrução sequencial, a análise de causalidade e a identificação de vetores de intrusão [3].

D. Estanqueidade Lógica do Ambiente

A atuação forense baseia-se no princípio da mínima intervenção [8]. Interagir com o sistema altera seu estado, efeito análogo ao "efeito do observador" [6]. A coleta ao vivo deve preceder análises aprofundadas ou remediação [8]. Softwares adicionais ou rotinas não essenciais devem ser evitados, focando na preservação célere de dados voláteis [9].

Comandos de cópia sem parâmetros de preservação modificam atributos críticos (MAC times) [4]. Utilitários como `tar` ou `xcopy` sem flags de retenção de metadados introduzem alterações na linha temporal [4]. Varreduras agressivas de antivírus ou *EDRs* podem gerar escrita massiva, interferindo na integridade do vestígio [9].

E. Controle Formal do Objeto Probante

Imagens forenses requerem aplicação imediata de funções de resumo (*hash*), como SHA-256 ou SHA-512, garantindo a integridade na aquisição [13]. A mídia resultante deve ser selada e protegida conforme a ISO 27037 e normativas aplicáveis [8].

Omitir anotações sobre "quem", "como", "quando"e "por qual meio"os vestígios foram retidos compromete a rastreabilidade [14]. A ausência de comprovação técnica de integridade propicia contestações anulatórias severas nos tribunais [21].

A observância ou deslize frente a estas práticas baseia a segurança jurídica das análises [21]. Metodologias precárias, mesmo com boas intenções, minam a confiabilidade, força e admissibilidade judicial do elemento pericial, invalidando o ressarcimento e a incriminação penal [14].

V. Desafios Modernos e Limitações da Ordem de Volatilidade

As formulações consolidadas pela RFC 3227 em 2002 estabeleceram bases sólidas para a priorização da coleta em ambientes computacionais tradicionais [7]. Contudo, a evolução arquitetural das últimas décadas, marcada pela consolidação da Computação em Nuvem e pela virtualização em larga escala, alterou significativamente o contexto operacional da forense digital [22].

Ambientes *cloud-native*, estruturados sobre microsserviços, contêineres orquestrados e arquiteturas *serverless*, introduzem dinâmicas de execução altamente efêmeras [22]. Tecnologias como Docker e Kubernetes, bem como funções elásticas em plataformas como AWS Lambda e Azure Functions, operam segundo ciclos de vida automatizados e transitórios [23]. Instâncias podem existir por segundos ou minutos, sendo criadas e destruídas conforme demanda [23]. Nesses cenários, artefatos tradicionalmente classificados nos níveis mais voláteis da RFC 3227 podem desaparecer antes mesmo da intervenção humana [22].

A investigação em ambientes de múltiplos inquilinos (*multi-tenancy*) impõe restrições adicionais [22]. O Modelo de Responsabilidade Compartilhada limita o acesso direto ao *hardware* físico do provedor de serviços em nuvem [24]. A captura de memória RAM no nível do *host* ou o acesso a registradores de CPU tornam-se, em regra, inviáveis técnica ou juridicamente [22]. Dessa forma, a aplicação literal da hierarquia clássica encontra obstáculos estruturais [22]. A volatilidade, nesses ambientes, não decorre apenas do desligamento do sistema, mas do próprio ciclo elástico de provisionamento [23]. Contêineres comprometidos utilizados para movimentação lateral ou funções *serverless* exploradas para execução maliciosa podem ser automaticamente encerrados, eliminando vestígios residentes em memória [23]. A escala massiva e a natureza distribuída das infraestruturas também ampliam a complexidade de isolamento e preservação [23].

Diante desse cenário, observa-se deslocamento estratégico da coleta para os estratos tradicionalmente classificados como mais estáveis, especialmente logs remotos, trilhas de auditoria, monitoramento de APIs e telemetria centralizada [25]. Esses registros tornam-se fontes primárias para reconstrução de eventos em ambientes *cloud-native* [22].

Nesse contexto emerge o conceito de *Forensic Readiness*, que pressupõe planejamento prévio de retenção probatória [26]. Soluções de monitoramento contínuo, plataformas de proteção de aplicações nativas em nuvem (*Cloud-Native Application Protection Platforms* - CNAPP) e mecanismos automatizados de auditoria permitem converter dados altamente voláteis em registros persistentes e verificáveis [22]. A ordem de volatilidade permanece válida como referencial metodológico, porém sua aplicação exige adaptação ao modelo distribuído contemporâneo [22]. Mais do que regra rígida, deve ser compreendida como princípio orientador que subsidia o juízo técnico do *Digital Evidence First Responder* [8]. Em ambientes nos quais o congelamento do *host* é impraticável, a coleta imediata de registros de nuvem e eventos de API pode representar a alternativa mais eficaz para preservar a materialidade probatória [25].

VI. Conclusão

A análise desenvolvida ao longo deste trabalho demonstra que a ordem de volatilidade transcende sua formulação original como diretriz operacional e deve ser compreendida como fundamento metodológico estruturante da preservação da evidência digital. Ao articular bases normativas consolidadas, como a RFC 3227, a ABNT NBR ISO/IEC 27037 e as diretrizes do NIST, com requisitos técnico-jurídicos de integridade, autenticidade e rastreabilidade, evidenciou-se que a priorização da coleta segundo o grau de transitoriedade dos dados constitui condição indispensável para maximizar a retenção informacional e assegurar a admissibilidade probatória. A transição do paradigma de *Dead Forensics* para *Live Forensics*, impulsionada por criptografia de disco inteiro, malwares *fileless* e execuções em memória, reforça que decisões técnicas devem ser orientadas por análise de risco. Por fim, os desafios impostos por arquiteturas *cloud-native*, virtualização massiva e modelos de responsabilidade compartilhada impõem adaptações relevantes

à aplicação clássica da hierarquia de volatilidade. Em ambientes distribuídos e efêmeros, a preservação probatória depende de estratégias de *Forensic Readiness*, telemetria contínua e retenção estruturada de logs e eventos de API. Logo, a ordem de volatilidade permanece válida como princípio orientador, desde que aplicada de forma contextualizada e integrada à governança de resposta a incidentes, garantindo equilíbrio entre preservação técnica, viabilidade operacional e solidez jurídica da prova digital.

VII. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] P. W. Singer e Allan Friedman. *Cybersecurity and Cyberwar: What Everyone Needs to Know*. Oxford University Press, 2014.

[2] Verizon. *2024 Data Breach Investigations Report (DBIR)*. Rel. técn. Verizon Business, 2024.

[3] Eoghan Casey. *Digital Evidence and Computer Crime: Forensic Science, Computers, and the Internet*. 3rd. Academic Press, 2011.

[4] Brian Carrier. *File System Forensic Analysis*. Addison-Wesley Professional, 2005.

[5] Edmond Locard. *L'Enquête criminelle et les méthodes scientifiques*. Flammarion, 1920.

[6] Dan Farmer e Wietse Venema. *Forensic Discovery*. Addison-Wesley Professional, 2005.

[7] D. Brezinski e T. Killalea. *Guidelines for Evidence Collection and Archiving*. RFC 3227. IETF, 2002. URL: https://tools.ietf.org/html/rfc3227.

[8] *ISO/IEC 27037:2012 Information technology – Security techniques – Guidelines for identification, collection, acquisition and preservation of digital evidence*. International Organization for Standardization, 2012.

[9] Paul Cichonski, Tom Millar, Tim Grance e Karen Scarfone. *Computer Security Incident Handling Guide*. Special Publication 800-61 Rev. 2. National Institute of Standards e Technology (NIST), 2012.

[10] Michael Hale Ligh, Andrew Case, Jamie Levy e Aaron Walters. *The Art of Memory Forensics: Detecting Malware and Threats in Windows, Linux, and Mac Memory*. Wiley, 2014.

[11] Richard Kissel, Andrew Regenscheid, Matthew Scholl e Kevin Stine. *Guidelines for Media Sanitization*. Special Publication 800-88 Rev. 1. NIST, 2014.

[12] Jason T. Luttgens, Matthew Pepe e Kevin Mandia. *Incident Response & Computer Forensics*. 3rd. McGraw-Hill Education, 2014.

[13] Dang Quynh. *Recommendation for Applications Using Approved Hash Algorithms*. Special Publication 800-107 Rev. 1. NIST, 2012.

[14] Brasil. *Lei nº 13.964, de 24 de dezembro de 2019. Aperfeiçoa a legislação penal e processual penal*. 2019. URL: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/lei/l13964.htm.

[15] Brian Hay, Kara Nance e Matt Bishop. "Live Analysis: Progress and Challenges". Em: *IEEE Security & Privacy* 7.2 (2009), pp. 66–71.

[16] Frank Adelstein. "Live Forensics: Diagnosing the System without Killing it". Em: *Digital Investigation* 3 (2006), pp. 63–66.

[17] Keith Stouffer, Victoria Pillitteri, Suzanne Lightman, Marshall Abrams e Adam Hahn. *Guide to Industrial Control Systems (ICS) Security*. Special Publication 800-82 Rev. 2. NIST, 2015.

[18] Diane Barrett e Gregory Kipper. *Virtualization and Forensics: A Digital Forensic Investigator's Guide to Virtual Environments*. Syngress, 2012.

[19] Michael Sikorski e Andrew Honig. *Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software*. No Starch Press, 2012.

[20] Harlan Carvey. *Windows Forensic Analysis DVD Toolkit*. 4th. Syngress, 2016.

[21] Damásio de Jesus. *Código de Processo Penal Anotado*. 31ª. Saraiva Jur, 2020.

[22] Amith Pichan, Mihai Lazarescu e Sieteng Soh. "Cloud forensics: Technical challenges and solutions". Em: *Digital Investigation* 13 (2015), pp. 38–57.

[23] Murugiah Souppaya, John Morello e Karen Scarfone. *Application Container Security Guide*. Special Publication 800-190. NIST, 2017.

[24] Wayne Jansen e Timothy Grance. *Guidelines on Security and Privacy in Public Cloud Computing*. Special Publication 800-144. NIST, 2011.

[25] Karen Kent e Souppaya Murugiah. *Guide to Computer Security Log Management*. Special Publication 800-92. NIST, 2006.

[26] Robert Rowlingson. "A Ten Step Process for Forensic Readiness". Em: *International Journal of Digital Evidence* 2.3 (2004).

Estudo de Modelo de Comunicação Segura Baseado em Confidencialidade, Integridade, Autenticidade e Irretratabilidade

Vinícius Santos Araújo, Guilherme Pedro Aquino, Evandro César Vilas Boas
Cyber Security and Artificial Intelligence (CS&I Lab)
Instituto Nacional de Telecomunicações - Inatel
vinicius.araujo@inatel.br, guilhermeaquino@inatel.br, evandro.cesar@inatel.br

***Abstract*—Contemporary information security requires a holistic approach that transcends simple protection of data at rest, focusing on the robustness of communications in transit. This article presents an Integrated Secure Communication Model based on four fundamental pillars: confidentiality, integrity, authenticity, and non-repudiation. Through a rigorous analysis of modern cryptographic primitives, such as the Advanced Encryption Standard (AES) and Elliptic Curve Cryptography (ECC), and compliance with international and national standards, such as NIST SP 800-57 and ICP-Brasil, the study proposes a "Sign-then-Encrypt"architecture optimized by hash functions. Performance implications in resource-constrained environments, such as IoT and 5G networks, are discussed, contrasting the efficiency of ECC against RSA. Additionally, the work addresses critical vulnerabilities, specifically the surreptitious forwarding attack, and the naming repairs necessary to ensure legal validity and technical security. The resulting model offers a resilient infrastructure for digital transactions, prepared for the challenges of the transition to the post-quantum era.**

***Index Terms*—Hybrid Cryptography, Digital Signature, Information Security, ICP-Brasil, NIST, Post-Quantum.**

***Resumo*—A segurança da informação contemporânea exige uma abordagem holística que transcenda a simples proteção de dados em repouso, concentrando-se na robustez das comunicações em trânsito. Este artigo apresenta um Modelo Integrado de Comunicação Segura fundamentado em quatro pilares fundamentais: confidencialidade, integridade, autenticidade e irretratabilidade. Através de uma análise rigorosa das primitivas criptográficas modernas, como o Advanced Encryption Standard (AES) e a Criptografia de Curva Elíptica (ECC), e da conformidade com normas internacionais e nacionais, como o NIST SP 800-57 e a ICP-Brasil, o estudo propõe uma arquitetura de "Assinar-então-Cifrar"otimizada por funções de hash. Discutem-se as implicações de desempenho em ambientes de recursos limitados, como IoT e redes 5G, contrastando a eficiência da ECC frente ao RSA. Adicionalmente, o trabalho aborda vulnerabilidades críticas, especificamente o ataque de encaminhamento sub-reptício (surreptitious forwarding), e as estratégias de reparação de identidade necessárias para garantir a validade jurídica e a segurança técnica. O modelo resultante oferece uma infraestrutura resiliente para transações digitais, preparada para os desafios da transição para a era pós-quântica.**

***Palavras chave*—Criptografia Híbrida, Assinatura Digital, Segurança da Informação, ICP-Brasil, NIST, Pós-Quântico.**

I. Introdução

A evolução das tecnologias de rede e a onipresença da computação em nuvem e da Internet das Coisas (IoT) redefiniram as fronteiras da segurança cibernética. O paradigma tradicional de segurança de perímetro, focado em firewalls e isolamento de rede, tornou-se insuficiente perante a complexidade das infraestruturas modernas, onde bilhões de dispositivos interagem de forma dinâmica. Nesse cenário, a Garantia da Informação (*Information Assurance*, IA) evoluiu do conceito clássico de INFOSEC para uma visão multidimensional que protege o dado em todos os seus estados: armazenamento, processamento e transmissão.

A implementação de comunicações seguras não é apenas uma necessidade técnica, mas um requisito regulatório e jurídico. No Brasil, a Infraestrutura de Chaves Públicas Brasileira (ICP-Brasil) e a Lei nº 14.063/2020 [1] estabelecem os critérios para a validade de assinaturas eletrônicas, diferenciando entre níveis de confiança que impactam diretamente a presunção legal de veracidade dos documentos. Para que um sistema de comunicação seja considerado seguro em um nível acadêmico e profissional elevado, ele deve integrar de forma coesa as propriedades de confidencialidade, integridade, autenticidade e irretratabilidade (não repúdio).

A confidencialidade garante que a informação seja acessível apenas a entidades autorizadas, enquanto a integridade assegura que o conteúdo recebido seja idêntico ao enviado, sem modificações espúrias. A autenticidade permite a verificação inequívoca da identidade das partes, e a irretratabilidade impede que um autor negue a autoria de uma mensagem ou transação. O desafio técnico reside em combinar esses pilares sem introduzir latência excessiva ou vulnerabilidades estruturais, como o encaminhamento sub-reptício, onde uma mensagem assinada pode ser redirecionada de forma maliciosa.

Este trabalho detalha um modelo integrado que utiliza criptografia híbrida para balancear desempenho e segurança. A proposta foca na superioridade da Criptografia de Curva Elíptica (ECC) sobre o RSA em ambientes modernos e discute a necessidade de preparo para a computação quântica, que ameaça a base matemática dos criptosistemas atuais. Através desta discussão, busca-se fornecer uma fundamentação teórica e prática sólida para a validação de sistemas de comunicação segura em escala global.

II. Pilares da Segurança

A arquitetura de um modelo de comunicação segura repousa sobre a definição técnica precisa e a implementação rigorosa de quatro serviços fundamentais. A compreensão desses serviços, conforme delineado pelas normas do National Institute of Standards and Technology (NIST) e pela International Organization

for Standardization (ISO), é o primeiro passo para a construção de sistemas resilientes [2].

A. Confidencialidade e Mecanismos de Cifragem

A confidencialidade é a propriedade que impede que a informação seja divulgada a indivíduos, processos ou dispositivos não autorizados. No contexto de redes, ela é alcançada predominantemente via criptografia. Os sistemas criptográficos são divididos em duas categorias principais: simétricos e assimétricos.

Criptosistemas simétricos, como o AES (Advanced Encryption Standard), utilizam a mesma chave para cifrar e decifrar. Eles são altamente eficientes em termos computacionais, permitindo o tráfego de grandes volumes de dados com baixa latência. No entanto, a segurança simétrica depende inteiramente da proteção e da distribuição segura da chave compartilhada. Em contraste, a criptografia assimétrica, ou de chave pública, utiliza pares de chaves matematicamente vinculados: uma chave pública para cifragem e uma chave privada correspondente para decifragem.

Embora a criptografia assimétrica resolva o problema da distribuição de chaves, ela exige maior poder de processamento. Por essa razão, sistemas de comunicação modernos adotam modelos híbridos: a criptografia assimétrica é utilizada para a troca segura de uma ”chave de sessão”simétrica, que então protege o tráfego de dados em massa. A eficácia deste pilar depende da robustez dos algoritmos e do comprimento das chaves, conforme recomendado pelo NIST SP 800-57 [3].

B. Integridade e Funções de Hash

A integridade garante que a informação não foi alterada de forma não autorizada durante o trânsito. Para assegurar essa propriedade, utilizam-se funções de hash criptográfico, que transformam uma entrada de dados de qualquer tamanho em uma saída de tamanho fixo, conhecida como Message Digest ou Integrity Check Value (ICV).

As funções de hash ideais devem ser resistentes a colisões e a ataques de pré-imagem. Atualmente, o padrão ouro inclui a família SHA-2 (como o SHA-256) e o SHA-3 [4], desenvolvido para oferecer resistência mesmo em cenários de alta complexidade computacional. Em um modelo integrado, o hash da mensagem é frequentemente assinado digitalmente para garantir que tanto o conteúdo quanto a autoria permaneçam íntegros e verificáveis.

C. Autenticidade e Infraestrutura de Chaves Públicas (PKI)

A autenticidade é o serviço que permite verificar a identidade de uma entidade ou a origem dos dados. Diferente da simples identificação, a autenticidade requer uma prova de vínculo entre uma identidade digital e uma entidade física ou lógica. Em larga escala, isso é viabilizado por Certificados Digitais e por uma Infraestrutura de Chaves Públicas (PKI).

Um certificado digital, em conformidade com o padrão X.509, é um documento eletrônico assinado por uma Autoridade Certificadora (AC) que atesta a propriedade de uma chave pública por um determinado sujeito. A confiança no sistema é hierárquica, culminando em uma AC-Raiz. No Brasil, a ICP-Brasil desempenha este papel, garantindo que as identidades digitais sejam autenticadas conforme rigorosos processos de validação. A autenticidade protege o sistema contra ataques de disfarce (spoofing) e é a base para o estabelecimento de canais de comunicação seguros.

D. Irretratabilidade e Assinatura Digital

A irretratabilidade, ou não repúdio, assegura que uma entidade não possa negar a autoria de uma mensagem ou a realização de uma transação. Este pilar é alcançado através da assinatura digital, que utiliza a chave privada do emissor para criar um vínculo criptográfico único com o conteúdo da mensagem.

Tecnicamente, o processo envolve gerar um hash da mensagem e cifrar esse hash com a chave privada do remetente. O destinatário utiliza a chave pública do remetente (obtida de um certificado digital autêntico) para decifrar o hash e compará-lo com o hash calculado localmente. Se houver coincidência, a autoria e a integridade são confirmadas simultaneamente. Juridicamente, a irretratabilidade é o que confere validade aos contratos digitais, uma vez que a evidência criptográfica pode ser apresentada a terceiros como prova de consentimento e autoria.

TABELA I
RELAÇÃO ENTRE PILARES E MECANISMOS NO MODELO PROPOSTO

Atributo	Descrição Técnica	Referência
Confidencialidade	Proteção contra acesso não autorizado via cifragem.	AES-256, FIPS 197
Integridade	Detecção de alterações via funções de hash.	SHA-256, FIPS 180-4
Autenticidade	Validação da origem e identidade via PKI.	X.509, ICP-Brasil
Irretratabilidade	Garantia de autoria e não negação via assinatura.	RSA, ECDSA, FIPS 186-5 [5]

III. MODELO INTEGRADO DE COMUNICAÇÃO SEGURA

A integração dos pilares em um fluxo operacional requer a consideração de múltiplos fatores, desde a ordem das operações criptográficas até a eficiência dos algoritmos escolhidos. O modelo proposto baseia-se na premissa de que a segurança deve ser ”by design”, integrando camadas de proteção que se reforçam mutuamente [6].

A. Arquitetura de Cifragem Dupla e Otimizações

O modelo clássico de ”Assinar-então-Cifrar” (Sign-then-Encrypt) oferece um alto nível de garantia. No emissor (A), a mensagem é assinada com a chave privada (PR_A) e o resultado é cifrado com a chave pública do receptor (PU_B). No receptor (B), o processo é invertido: a decifragem com a chave privada (PR_B) expõe a mensagem assinada, que é então validada com a chave pública de A (PU_A).

Contudo, para otimizar o desempenho, especialmente em transmissões volumosas, o modelo proposto utiliza a assinatura do ICV (Integrity Check Value) gerado por um hash. Em vez de assinar toda a mensagem M, assina-se apenas $h = H(M)$. A mensagem M e a assinatura s_h são então cifradas em conjunto. Esta otimização reduz a carga computacional sobre algoritmos assimétricos, que são inerentemente mais lentos que os simétricos.

TABELA II
ANÁLISE COMPARATIVA DE DESEMPENHO: RSA-3072 VS. ECC-256

Métrica de Desempenho	RSA-3072	ECC-256
Tamanho da Chave (Bits)	3072	256
Tempo de Geração de Chave (ms)	2532.8	739.6
Tempo de Decifragem (ms)	317.8	129.2
Consumo Médio de Memória (MB)	15.2	14.0

B. Análise Comparativa de Algoritmos: RSA vs. ECC

A escolha do algoritmo assimétrico é crítica para a viabilidade do modelo em sistemas modernos. O RSA, baseado na dificuldade de fatoração de grandes números primos, dominou a criptografia por décadas. No entanto, para oferecer segurança equivalente a uma chave simétrica de 128 bits, o RSA exige chaves de pelo menos 3072 bits, o que resulta em um consumo de memória e tempo de processamento elevado.

A Criptografia de Curva Elíptica (ECC) [7, 8], baseada no problema do logaritmo discreto em curvas elípticas, oferece segurança equivalente com chaves muito menores (ex.: 256 bits). Esta eficiência traduz-se em menor consumo de energia e menor latência, fatores cruciais para dispositivos IoT e aplicações móveis.

Dados experimentais demonstram que a ECC é aproximadamente duas vezes mais rápida na decifragem e três vezes mais eficiente na geração de chaves em comparação ao RSA para níveis de segurança similares [9]. Para comunicações em tempo real, como as exigidas em sistemas de veículos autônomos (IoV) ou redes 5G, a ECC apresenta-se como a escolha técnica superior.

C. Vulnerabilidades de Protocolo: Encaminhamento Sub-reptício

Um erro comum em implementações de "Sign-then-Encrypt" é a negligência quanto ao contexto da comunicação. O ataque de "encaminhamento sub-reptício"(surreptitious forwarding) [10] ilustra esta falha: um destinatário legítimo (Bob) pode decifrar uma mensagem de Alice e reencaminhá-la para Charlie, cifrando-a com a chave pública de Charlie, mas mantendo a assinatura original de Alice.

Neste cenário, Charlie recebe uma mensagem assinada por Alice e cifrada para ele. Ele pode acreditar erroneamente que Alice enviou a mensagem diretamente a ele, quando na verdade ela foi enviada a Bob. Para evitar este ataque, o modelo deve implementar "Reparos de Identidade"(Naming Repairs):

1) Vínculo do Destinatário: Antes de assinar, o emissor deve incluir o identificador único do destinatário (ID ou chave pública) dentro do conteúdo da mensagem. Exemplo: $S = Sign(PR_A, ID_{Destinatário}||M)$.
2) Verificação de Contexto: O receptor deve obrigatoriamente validar se o seu identificador consta na carga útil assinada. Se houver discrepância, a mensagem deve ser descartada como não confiável.

Estas medidas garantem que a assinatura digital não apenas ateste a autoria, mas também a intenção original do autor em relação ao destinatário, reforçando a autenticidade e a irretratabilidade sob o ponto de vista lógico e jurídico.

TABELA III
RSA VS. ECC

Nível de Assinatura	Requisitos Técnicos	Aplicação Recomendada
Simples	Identificação básica (login/senha).	Atos de baixo risco, agendamentos.
Avançada	Certificados não-ICP ou meios que garantam integridade.	Assinaturas GOV.BR, contratos comerciais.
Qualificada	Uso obrigatório de certificado ICP-Brasil.	Transferência de imóveis, prescrição médica, atos oficiais.

IV. CONTEXTO NORMATIVO E APLICAÇÃO EM SISTEMAS CRÍTICOS

A validação de um modelo integrado de comunicação segura exige conformidade com os marcos legais e técnicos do ambiente onde será operado. No Brasil, isso envolve a compreensão profunda da ICP-Brasil.

A. A ICP-Brasil e a Lei 14.063/2020

A Infraestrutura de Chaves Públicas Brasileira (ICP-Brasil) foi instituída pela Medida Provisória nº 2.200-2/2001 para garantir a autenticidade, integridade e validade jurídica de documentos eletrônicos. A evolução legislativa trouxe a Lei nº 14.063/2020, que modernizou o uso de assinaturas eletrônicas na interação com entes públicos, estabelecendo três níveis de segurança: simples, avançada e qualificada [1].

O modelo discutido neste artigo enquadra-se no nível de Assinatura Qualificada, uma vez que utiliza certificados digitais emitidos por autoridades credenciadas, oferecendo a mais alta presunção legal de veracidade. A recente Instrução Normativa 33/2025 da ICP-Brasil reforça a necessidade de atualização para políticas de assinatura mais recentes e compatibilidade com a AC Raiz v12, visando a resiliência contra ameaças emergentes.

B. Segurança em Redes 5G e IoT

A aplicação do modelo em ambientes de 5G e IoT apresenta desafios adicionais. A natureza descentralizada e a alta densidade de dispositivos exigem protocolos leves, mas robustos. Redes 5G utilizam autenticação mútua e proteção de integridade nas interfaces de rádio, porém a segurança de ponta-a-ponta (E2EE) no nível da aplicação continua dependente de modelos como o proposto.

Esquemas de autenticação baseados em blockchain (EdgeChain) e inteligência artificial para detecção de anomalias estão sendo integrados a modelos de comunicação para oferecer visibilidade em tempo real e resistência a ataques de negação de serviço (DDoS). A integração da ECC nestes ambientes é fundamental, pois permite que sensores com restrição de energia realizem operações criptográficas complexas sem exaurir suas baterias, mantendo a integridade e a autenticidade dos fluxos de dados massivos gerados por cidades inteligentes e indústrias 4.0.

V. ANÁLISE DE SEGURANÇA E VETORES DE ATAQUE

O horizonte da segurança cibernética é marcado pela ameaça da computação quântica. Algoritmos atuais, como RSA e ECC, baseiam-se em problemas matemáticos que podem ser

resolvidos em tempo polinomial por computadores quânticos utilizando o algoritmo de Shor.

A. O Risco Quântico e a Seleção do NIST

A vulnerabilidade dos algoritmos de chave pública atuais significa que a autenticidade e a irretratabilidade de documentos assinados hoje podem ser contestadas futuramente, e a confidencialidade de dados cifrados pode ser rompida assim que um computador quântico funcional esteja disponível. O NIST lidera o esforço global para padronizar algoritmos de Criptografia Pós-Quântica (PQC), focando em abordagens baseadas em redes (lattices), códigos e equações multivariadas. Em 2024, foram anunciados os finalistas para padronização, incluindo o CRYSTALS-Kyber para KEM (Key Encapsulation Mechanism) e o CRYSTALS-Dilithium para assinaturas digitais [11, 12].

B. Estratégias de Transição Híbrida

Para garantir a continuidade da segurança, o modelo de comunicação segura deve evoluir para esquemas híbridos. Estes esquemas combinam um algoritmo clássico comprovado (como ECC ou RSA) com um algoritmo pós-quântico emergente. Se qualquer um dos algoritmos for comprometido, a segurança do outro ainda protege a comunicação. Esta abordagem "dual-layered"garante que a confidencialidade e a irretratabilidade permaneçam válidas a longo prazo, protegendo contra o ataque de "armazenar agora, decifrar depois"(Harvest Now, Decrypt Later). A implementação de PQC em dispositivos IoT continua sendo um campo de pesquisa ativa, devido à maior complexidade computacional e ao tamanho das chaves e assinaturas dos novos algoritmos

VI. Conclusão

A construção de um "Modelo Integrado de Comunicação Segura"é um imperativo técnico para as organizações que buscam operar com confiança no ecossistema digital. A análise apresentada demonstra que a integração harmoniosa de confidencialidade, integridade, autenticidade e irretratabilidade não é apenas uma questão de escolha de algoritmos, mas de design de protocolo e conformidade normativa.

A superioridade técnica da Criptografia de Curva Elíptica (ECC) em relação ao RSA, evidenciada pelos benchmarks de desempenho em tempo de geração de chaves e decifragem, consolida-a como a base para comunicações modernas, especialmente em ambientes restritos como IoT e 5G. Paralelamente, a robustez contra ataques como o encaminhamento sub-reptício exige que os implementadores adotem práticas rigorosas de vinculação de identidade, garantindo que o contexto da transação seja tão seguro quanto o seu conteúdo.

No âmbito jurídico, o alinhamento com os padrões da ICP-Brasil e a classificação de assinaturas da Lei 14.063/2020 assegura que o modelo técnico possua a necessária eficácia legal, permitindo a digitalização de processos críticos com segurança jurídica. Por fim, a iminente revolução quântica impõe uma agenda de atualização contínua, onde a adoção de criptografia pós-quântica híbrida torna-se a próxima fronteira para a preservação do sigilo e da autoria em escala global. Este modelo, portanto, serve como um guia sólido para o desenvolvimento de infraestruturas de comunicação resilientes, prontas para enfrentar as ameaças presentes e futuras.

VII. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Laboratório de Segurança Cibernética e Inteligência Artificial (CS&I Lab) do Instituto Nacional de Telecomunicações (Inatel).

Referências

[1] Brasil. *Lei nº 14.063, de 23 de setembro de 2020*. 2020. URL: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/l14063.htm.

[2] ISO/IEC. *ISO/IEC 27001:2022 - Information security, cybersecurity and privacy protection — Information security management systems — Requirements*. International Organization for Standardization, 2022.

[3] Elaine Barker. *Recommendation for Key Management: Part 1 – General*. Rel. técn. NIST SP 800-57 Part 1 Rev. 5. National Institute of Standards e Technology (NIST), 2020. URL: https://doi.org/10.6028/NIST.SP.800-57pt1r5.

[4] NIST. *SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions*. Rel. técn. FIPS 202. Federal Information Processing Standards Publication, 2015.

[5] NIST. *Digital Signature Standard (DSS)*. Rel. técn. FIPS 186-5. Federal Information Processing Standards Publication (FIPS), 2023.

[6] Yuliang Zheng. "Digital Signcryption or How to Achieve Cost (Signature + Encryption) ¡ Cost (Signature) + Cost (Encryption)". Em: *Advances in Cryptology — CRYPTO '97*. Springer, 1997, pp. 165–179.

[7] Victor S. Miller. "Use of Elliptic Curves in Cryptography". Em: *Advances in Cryptology — CRYPTO '85*. Springer, 1985, pp. 417–426.

[8] Neal Koblitz. "Elliptic Curve Cryptosystems". Em: *Mathematics of Computation* 48.177 (1987), pp. 203–209.

[9] Rebwar Khalid Muhammed et al. "A Comparative Performance Evaluation of Hybrid Encryption Techniques Using ECC, RSA, AES, and ChaCha20 for Secure Data Transmission". Em: *International Journal of Information and Research* 12.2 (2025). URL: https://doi.org/10.53523/ijoirVol12I2ID598.

[10] Don Davis. "Defective Sign and Encrypt in S/MIME, PKCS#7, MOSS, PEM, PGP, and XML". Em: *Proceedings of the 2001 USENIX Annual Technical Conference*. Boston, MA, 2001, pp. 65–78.

[11] NIST. *Module-Lattice-Based Key-Encapsulation Mechanism Standard*. Rel. técn. FIPS 203. Federal Information Processing Standards Publication, ago. de 2024.

[12] NIST. *Module-Lattice-Based Digital Signature Standard*. Rel. técn. FIPS 204. Federal Information Processing Standards Publication, ago. de 2024.

Desenvolvimento de um Sistema Lúdico Baseado em Realidade Virtual Não Imersiva para Reabilitação de Mãos com Artrite Reumatoide

Ervin Emanuel Orellana Pardo, Bianca Camila Garnica Pericon
Universidad Privada del Valle – UNIVALLE
ervin.orellanapardo@alumni.univalle.edu, bianca.garnicapericon@alumni.univalle.edu
ORCID: 0009-0008-3927-0733, 0009-0006-1010-4040

Abstract—This article presents the development of "ExoHand VR", an interactive system designed to promote adherence to physiotherapeutic treatment in patients with rheumatoid arthritis (RA) in stages I–II, implemented at the Pura Vida Salud Fisioterapia center in Cochabamba, Bolivia. The system integrates audiovisual feedback, real-time force monitoring, four non-immersive virtual reality games aimed at stimulating upper limb motor patterns, and an easy-access web platform. The results showed an improvement in grip strength and a greater willingness of patients to continue therapy. Additionally, users reported ease of use and intrinsic motivation during the sessions, aspects corroborated by the responsible physiotherapist. This project represents a relevant advance in the incorporation of innovative technologies in physiotherapy in Bolivia, where the application of emerging technologies in rehabilitation and therapeutic adherence still present limitations.

Index Terms—Rheumatoid arthritis, Physiotherapy, non-immersive virtual reality, Therapeutic adherence.

Resumo— Este artigo apresenta o desenvolvimento do "ExoHand VR", um sistema interativo projetado para promover a adesão ao tratamento fisioterapêutico em pacientes com artrite reumatoide (AR) em estágios I–II, implementado no centro Pura Vida Salud Fisioterapia em Cochabamba, Bolívia. O sistema integra retroalimentação audiovisual, monitoramento de força em tempo real, quatro jogos de realidade virtual não imersiva voltados para a estimulação de padrões motores do membro superior e uma plataforma *web* de fácil acesso. Os resultados evidenciaram uma melhora na força de preensão e uma maior disposição dos pacientes para continuar a terapia. Além disso, os usuários relataram facilidade de uso e motivação intrínseca durante as sessões, aspectos corroborados pelo fisioterapeuta responsável. Este projeto representa um avanço relevante na incorporação de tecnologias inovadoras na fisioterapia na Bolívia, onde a aplicação de tecnologias emergentes na reabilitação e a adesão terapêutica ainda apresentam limitações.

Palavras-chave— Artrite reumatoide, Fisioterapia, Realidade virtual, Adesão terapêutica

I. Introdução

A artrite reumatoide (AR) é uma doença que afeta diretamente as articulações do corpo, incluindo joelhos, quadris, cotovelos, mãos, dedos e ombros, entre outras [1]. Segundo a Organização Mundial da Saúde (OMS) [2], em 2019 havia no mundo 18 milhões de pessoas com artrite reumatoide. A AR é a forma mais comum de artrite autoimune e afeta 5 a cada mil adultos.

Na Bolívia, embora os dados sejam limitados devido ao fato de ainda ser um país em desenvolvimento, estima-se que aproximadamente 60% das pessoas com mais de 50 anos sofram de reumatismo, sendo a artrite reumatoide uma das mais comuns. Sabe-se que, de cada dez pessoas que sofrem da doença, sete são mulheres e três são homens [3]. Em Cochabamba, aproximadamente 18 mil a 25 mil habitantes padecem dessa enfermidade, e cerca de 75% são mulheres [4]. Essa condição gera sintomas como dor, inflamação e rigidez articular, que não apenas limitam a mobilidade, mas também dificultam a realização de atividades cotidianas essenciais.

O projeto ExoHand VR surge como uma proposta que utiliza realidade virtual não imersiva para transformar o enfoque tradicional da reabilitação. Essa plataforma *web* permite que os pacientes realizem exercícios terapêuticos interativos por meio de jogos projetados para melhorar a amplitude de movimento, a força e a funcionalidade das mãos afetadas. Além disso, incorpora um dispositivo de monitoramento de força que oferece retroalimentação em tempo real, tanto para o paciente quanto para o profissional de saúde, permitindo um acompanhamento mais preciso do progresso do tratamento.

Com este projeto, espera-se contribuir de maneira significativa para o bem-estar dos pacientes em Cochabamba, Bolívia, facilitando sua recuperação e melhorando sua qualidade de vida. Da mesma forma, o ExoHand VR representa uma contribuição valiosa no campo da tecnologia aplicada à saúde, abrindo novas possibilidades para a pesquisa e o desenvolvimento de soluções inovadoras no tratamento de doenças articulares crônicas.

II. Trabalhos Relacionados

Nos últimos anos, o uso de realidade virtual e jogos sérios na reabilitação motora tem sido objeto de crescente interesse na literatura científica. Bressler et al. [5] desenvolveram um jogo sério em realidade virtual para reabilitação de mãos e dedos, cujo processo iterativo com usuários confirmou a viabilidade do sistema mesmo em pacientes com função manual severamente comprometida, destacando o papel do feedback visual como fator motivacional; contudo, o estudo não incorporou monitoramento quantitativo de força de preensão em tempo real.

Soluções que integram rastreamento óptico com jogos sérios também foram exploradas para reabilitação do membro superior. Vallati et al. [6] apresentaram um sistema de realidade aumentada projetada para reabilitação do ombro, utilizando rastreamento da mão por *Leap Motion Controller* e métricas de

desempenho em tempo real. Apesar dos resultados positivos, o sistema não contempla o monitoramento da força de preensão nem foi avaliado especificamente em pacientes com artrite reumatoide.

Diante desse cenário, o ExoHand VR se diferencia por integrar, em uma única plataforma web acessível, quatro jogos terapêuticos baseados em visão computacional com um dispositivo de monitoramento de força de preensão em tempo real, desenvolvido especificamente para o contexto clínico boliviano, onde recursos tecnológicos especializados e de alto custo ainda são limitados.

III. Design do Sistema

O sistema fundamenta-se na convergência sinérgica entre *hardware* e *software*, visando otimizar a experiência do usuário e a eficiência operacional. A seguir, apresentam-se as especificações detalhadas de ambos os componentes

A. Arquitetura do Sistema

A arquitetura do sistema é composta por dois módulos principais: (i) módulo de interação, correspondente à plataforma *web* com jogos terapêuticos que utilizam visão computacional para o rastreamento dos movimentos da mão; e (ii) módulo de aquisição e processamento de dados, responsável pela medição da força de preensão por meio de um sensor de força, cujo sinal é processado por um microcontrolador ESP32 e transmitido via protocolo *Message Queuing Telemetry Transport* (MQTT) para a interface *web*. A Figura 1 apresenta a arquitetura principal do sistema, evidenciando a conexão entre o usuário e os diferentes componentes, tanto de *hardware* quanto de *software*.

O sistema permite a interação do paciente com jogos sérios que estimulam movimentos funcionais do membro superior, promovendo maior engajamento e adesão ao tratamento fisioterapêutico. Paralelamente, o sistema registra em tempo real os dados de força de preensão, permitindo o acompanhamento do desempenho do paciente durante as sessões.

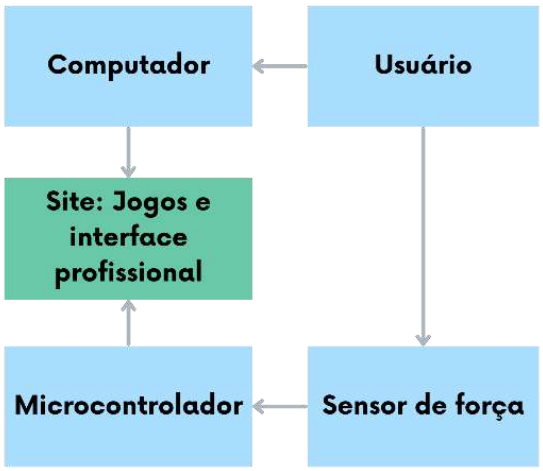
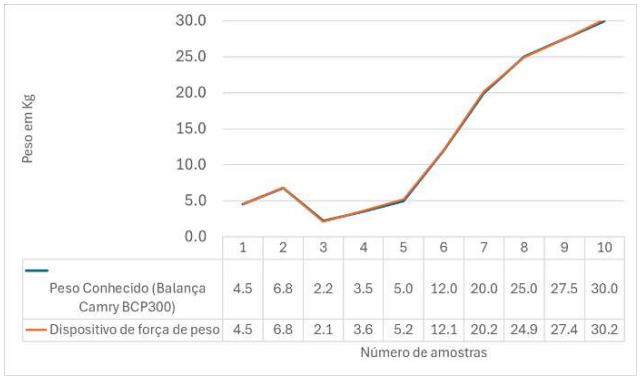


Fig. 1. Arquitetura de sistema que combina *software* e *hardware*

B. Hardware do Sistema

O módulo de *hardware* foi projetado para realizar o monitoramento da força de preensão durante as atividades terapêuticas. O sistema utiliza uma célula de carga de 50 kg em configuração de meio ponte [7], acoplada a um módulo HX711 [8] para amplificação e leitura do sinal, conectada a um microcontrolador ESP32. Para o ajuste do sistema, foi utilizado um peso conhecido de 10 lb, permitindo definir o fator de conversão do sinal. O sistema apresentou um erro absoluto de ±0,1 kg e uma resolução de 0,1 kg. Adicionalmente, foi realizada uma verificação comparativa com uma balança comercial Camry BCP300, obtendo-se um erro relativo inferior a 2%. A Figura 2 apresenta as curvas de comparação entre o dispositivo de monitoramento de força desenvolvido e a balança Camry, evidenciando a correspondência entre as medições.



	1	2	3	4	5	6	7	8	9	10
Peso Conhecido (Balança Camry BCP300)	4.5	6.8	2.2	3.5	5.0	12.0	20.0	25.0	27.5	30.0
Dispositivo de força de peso	4.5	6.8	2.1	3.6	5.2	12.1	20.2	24.9	27.4	30.2

Fig. 2. Curva de comparação entre o dispositivo desenvolvido e a balança de referência

O ESP32 é responsável pelo processamento dos dados e pela comunicação sem fio com o sistema *web* [9] através do protocolo MQTT. O dispositivo inclui ainda indicadores visuais o *Light Emitting Diode* (LED) e sonoros (*buzzer*) para fornecer *feedback* imediato ao usuário durante os exercícios.

A alimentação do sistema é realizada por meio de uma bateria recarregável, permitindo sua utilização em ambientes clínicos sem a necessidade de conexão contínua à rede elétrica.

O protótipo foi desenvolvido utilizando uma estrutura impressa em modelos tridimensionais (3D), garantindo ergonomia e facilidade de uso durante as sessões terapêuticas. A Figura 3 apresenta a estrutura física do dispositivo, incluindo a caixa que abriga a circuitaria do projeto, o *buzzer* e o LED indicador de conexão, bem como a carcaça da célula de carga utilizada para a medição da força.

No que diz respeito ao monitoramento de força como ferramenta clínica, existem soluções comerciais consolidadas, como o *KFORCE Muscle Controller* (KINVENT), dinamômetro portátil com capacidade de até 90 kg, precisão de 0.1 kg [10]. Embora esse tipo de dispositivo ofereça robustez e validação clínica formal, seu design não foi concebido para a mensuração individualizada da força de preensão dedo a dedo, o que representa uma limitação para avaliações funcionais detalhadas em pacientes com comprometimento assimétrico dos dedos, como ocorre frequentemente na artrite reumatoide.



Fig. 3. Caixa que abriga o circuito do dispositivo de monitoramento de força e a carcaça da célula de carga de 50 kg.

No entanto, para sua adoção em contextos clínicos formais, devem ser consideradas algumas limitações práticas do dispositivo. Em relação à calibração, recomenda-se a verificação periódica antes de cada sessão terapêutica utilizando um peso padrão de 10 lb, procedimento que pode ser realizado diretamente pela interface web do sistema, cujos parâmetros de calibração são armazenados na memória do microcontrolador ESP32; embora a variação entre sessões seja mínima, essa prática garante a confiabilidade das medições ao longo do tempo e compensa o desgaste natural dos componentes. Quanto à higienização, o dispositivo requer protocolos de limpeza entre sessões com diferentes pacientes, aspecto que deverá ser contemplado no design de versões futuras com materiais e superfícies adequados para uso clínico.

Por fim, para sua utilização como instrumento de avaliação clínica oficial, o dispositivo necessitaria de aprovação pelos organismos regulatórios competentes, o que representa uma etapa necessária para versões destinadas a uso diagnóstico formal.

C. Software do Sistema

O *software* do sistema foi desenvolvido como uma plataforma *web* interativa, utilizando tecnologias como *JavaScript*, *HyperText Markup Language* (HTML) e *Cascading Style Sheets* (CSS), permitindo sua execução em computadores convencionais sem necessidade de dispositivos especializados.

Para o reconhecimento de movimentos da mão, foi utilizada a biblioteca *MediaPipe Hand Landmarker* [8], que possibilita o rastreamento em tempo real dos pontos-chave da mão através da câmera, conforme ilustrado na Figura 4. Essa abordagem permite a representação gráfica da mão no ambiente virtual e a interação direta com os elementos do sistema.Adicionalmente, foi utilizada a biblioteca Three.js para a criação de elementos gráficos tridimensionais, incluindo o modelado da mão e dos ambientes dos jogos, permitindo a visualização de modelos tridimensionais 3D [11] e a construção dos cenários interativos.

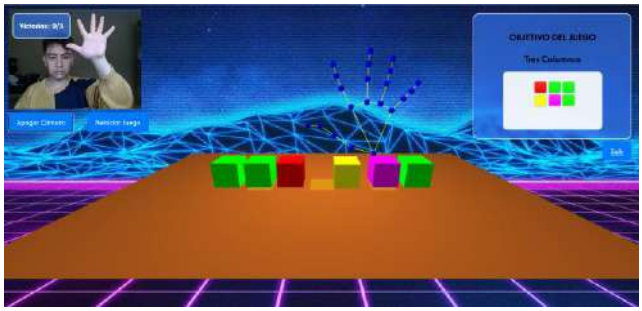


Fig. 4. Mão representada no palco virtual

Foram desenvolvidos quatro jogos terapêuticos com o objetivo de estimular diferentes padrões motores do membro superior, incluindo movimentos de flexão-extensão, abdução-adução e preensão. O desenvolvimento dos exercícios foi baseado no "*Manual de Ayuda al Ejercicio para personas con Artritis Reumatoide*", cujas diretrizes terapêuticas orientaram a seleção e adaptação das atividades propostas, conforme apresentado na Figura 5 [12], garantindo que as atividades propostas estejam alinhadas com recomendações terapêuticas para esta patologia. Os jogos fornecem *feedback* visual e numérico em tempo real, promovendo a motivação do paciente e facilitando o acompanhamento do progresso.

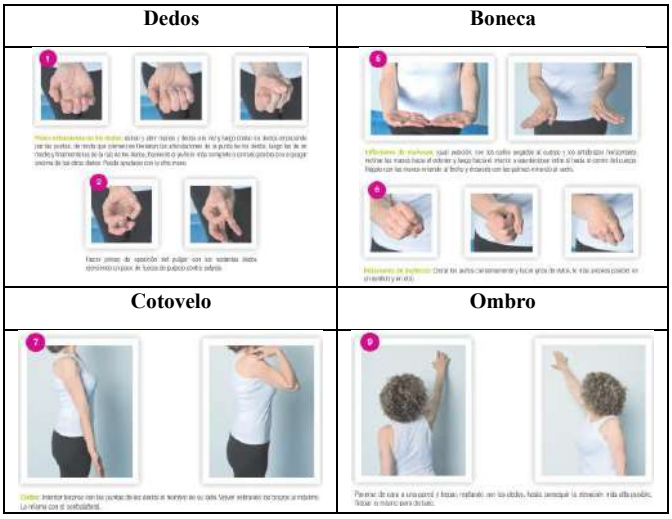


Fig. 5. Exercícios de referência para os jogos desenvolvidos.

O desenvolvimento dos jogos foi baseado no "*Manual de Ayuda al Ejercicio para personas com Artritis Reumatoide*", com orientação do fisioterapeuta Dr. Dery Mérida, garantindo sua adequação clínica.

1. **Empilhe cubos:** Jogo focado na coordenação motora e mobilidade do membro superior, no qual o paciente deve manipular e empilhar objetos virtuais, promovendo movimentos de flexão-extensão e preensão.
2. **Caixas e Cores:** Jogo que estimula o alcance e a coordenação motora, no qual o paciente deve classificar objetos de acordo com cores, favorecendo movimentos de extensão do braço e controle manual.

3. **Guitar Hand:** Jogo rítmico que estimula a mobilidade e resistência muscular por meio da interação com estímulos visuais sincronizados, promovendo movimentos repetitivos dos dedos e da mão.
4. **Recolha as pinças:** Jogo voltado à motricidade fina, no qual o paciente realiza movimentos de pinça para manipular objetos pequenos, fortalecendo a coordenação e precisão dos dedos.

Em conjunto, os jogos promovem a execução de exercícios terapêuticos de forma dinâmica, favorecendo a adesão ao tratamento e a reabilitação funcional dos pacientes. A Figura 6 apresenta os diferentes cenários desenvolvidos para cada jogo implementado na plataforma.

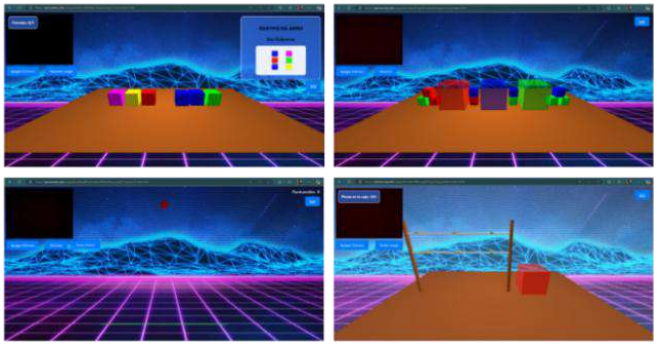


Fig. 6. Configurações do jogo

IV. Resultados e Discussão

O sistema foi avaliado por meio de testes de campo realizados com dois pacientes no Centro de Fisioterapia Pura Vida Salud diagnosticados com artrite reumatoide em estágios iniciais, com idades de 58 e 65 anos. O paciente 1 apresentava AR em estágio I, enquanto o paciente 2 em estágio II. As sessões terapêuticas foram conduzidas ao longo de seis semanas, totalizando 18 sessões. Para a avaliação quantitativa, foram realizadas medições de força de preensão utilizando dinamometria antes e após a intervenção, com foco principal nos dedos dos pacientes. Adicionalmente, de acordo com a avaliação clínica do profissional responsável, foram realizadas medições específicas nas regiões mais afetadas de cada paciente utilizando o dispositivo de monitoramento de força. No paciente 1, as medições foram realizadas em movimentos de flexão do cotovelo, enquanto no paciente 2 foram realizadas em movimentos de abdução do ombro.

Todas as medições foram conduzidas pelo profissional do centro de fisioterapia, sendo posteriormente analisadas e interpretadas conforme critérios clínicos.

A Tabela 1 apresenta os valores de força de preensão obtidos para cada dedo de ambas as mãos, antes e após a realização das sessões terapêuticas. Os resultados permitem observar a variação individual da força e a evolução funcional dos pacientes ao longo do período de intervenção.

TABELA I
Força de Preensão dos Dedos de Cada Mão e Melhoria

Paciente	Mão	Dedo	Antes (kg)	Depois (kg)	Melhoria (%)
P1	Direita	Indicador	3.2	4.2	31.25
P1	Direita	Médio	3.8	6.4	68.42
P1	Direita	Anelar	2.3	3.7	60.87
P1	Direita	Mínimo	0.9	2.7	200.00
P1	Esquerda	Indicador	3.5	5.2	48.57
P1	Esquerda	Médio	5.0	6.3	26.00
P1	Esquerda	Anelar	3.0	3.7	23.33
P1	Esquerda	Mínimo	2.0	2.6	30.00
P2	Direita	Indicador	2.4	2.9	20.83
P2	Direita	Médio	3.7	4.1	10.81
P2	Direita	Anelar	2.9	3.3	13.79
P2	Direita	Mínimo	1.6	2.3	43.75
P2	Esquerda	Indicador	2.1	2.6	23.81
P2	Esquerda	Médio	2.8	3.0	7.14
P2	Esquerda	Anelar	1.9	2.4	26.32
P2	Esquerda	Mínimo	1.9	2.4	26.32

A Tabela 2 apresenta os valores de força medidos nas articulações mais afetadas de cada paciente, de acordo com a avaliação clínica. Observa-se uma maior variação nos resultados do paciente em estágio inicial, evidenciando melhor resposta ao tratamento proposto.

TABELA II
Força nas Articulações dos Pacientes e Melhoria

Paciente	Articulação	Lado	Antes (kg)	Depois (kg)	Melhoria (%)
P1	Cotovelo	Direita	5.1	7.5	47.06
P1	Cotovelo	Esquerda	5.4	7.6	40.74
P2	Ombro	Direita	4.8	5.5	14.58
P2	Ombro	Esquerda	4.3	4.4	2.33

Os resultados indicaram um aumento na força muscular, com incrementos entre 1 e 3 kgf nos membros superiores dos pacientes. Além disso, foram analisadas métricas internas do sistema, como pontuação nos jogos e tempo de execução. Observou-se uma melhoria progressiva no desempenho dos pacientes, com pontuações elevadas e redução no tempo de realização das atividades.

Na avaliação qualitativa, foram aplicados entrevistas e questionários de satisfação, nos quais os pacientes relataram facilidade de uso, maior motivação e maior adesão ao tratamento. O fisioterapeuta responsável também destacou a utilidade do sistema como complemento às terapias convencionais. Os resultados sugerem que o uso de elementos lúdicos associados ao monitoramento em tempo real contribui para melhorar a experiência do paciente e potencializar a adesão terapêutica.

TABELA III
Nível de Motivação dos Pacientes em uma Escala de 5

Paciente	Motivação Antes	Motivação Depois
P1	1	4
P2	2	4

Os resultados obtidos na avaliação de motivação, apresentados na Tabela 3, indicam um aumento significativo no interesse dos pacientes em relação ao processo de reabilitação após a utilização do sistema proposto. O paciente 1 apresentou um aumento mais expressivo, passando de um nível baixo de motivação para um nível elevado, enquanto o paciente 2 também demonstrou melhora relevante, alcançando níveis mais altos de motivação. Esses resultados sugerem que a incorporação de elementos lúdicos e interativos por meio da realidade virtual não imersiva contribui positivamente para o engajamento dos pacientes durante as sessões terapêuticas.

V. Conclusão

O desenvolvimento do sistema ExoHand VR representa um avanço na aplicação de tecnologias de realidade virtual não imersiva como ferramenta complementar na reabilitação de pacientes com artrite reumatoide. A integração de uma plataforma web interativa com um sistema de monitoramento de força permitiu não apenas a execução de exercícios terapêuticos, mas também o acompanhamento do desempenho dos pacientes em tempo real, promovendo uma abordagem mais dinâmica e motivadora.

Os resultados obtidos evidenciaram melhorias na força de preensão em todos os dedos avaliados, com incrementos mais expressivos no paciente em estágio inicial da doença, onde foram observadas variações superiores a 60% em alguns casos. De forma complementar, as medições nas articulações mais afetadas demonstraram aumentos de até 47% na força, enquanto o paciente em estágio mais avançado apresentou melhorias mais moderadas. Esses achados indicam que o sistema pode contribuir de maneira mais significativa nas fases iniciais da doença.

Além dos ganhos funcionais, observou-se um aumento relevante nos níveis de motivação dos pacientes, passando de valores iniciais baixos (1–2) para níveis mais elevados (4) em uma escala de 5 pontos. Esse resultado destaca o impacto positivo do uso de jogos terapêuticos e feedback em tempo real na adesão ao tratamento, um dos principais desafios na fisioterapia de pacientes com doenças crônicas.

Entre os principais desafios enfrentados ao longo do desenvolvimento, destacam-se: a calibração e validação do dispositivo de monitoramento de força para garantir medições confiáveis em contexto clínico; a adaptação dos jogos terapêuticos para atender às limitações motoras dos pacientes sem comprometer a experiência de uso; e a integração em tempo real entre o hardware baseado em ESP32 e a plataforma web via protocolo MQTT, que exigiu ajustes para minimizar latências durante as sessões. Adicionalmente, o reconhecimento de movimentos por visão computacional apresentou sensibilidade às condições de iluminação do ambiente, o que demandou um processo iterativo de ajuste para garantir a estabilidade do rastreamento.

O estudo apresenta limitações relacionadas ao tamanho reduzido da amostra, composta por apenas dois pacientes, e ao período de intervenção de seis semanas, o que restringe a generalização dos resultados. Para publicações futuras, recomenda-se a ampliação do número de participantes, a inclusão de um grupo controle que permita comparações mais robustas com a fisioterapia convencional, e avaliações clínicas complementares com acompanhamento de longo prazo. Esses passos são essenciais para consolidar a evidência sobre a eficácia do sistema e seu potencial adoção em contextos clínicos mais amplos.

Referências

[1] D. Aletaha and J. S. Smolen, “Diagnosis and Management of Rheumatoid Arthritis: A Review,” Oct. 02, 2018, *American Medical Association*. doi: 10.1001/jama.2018.13103.

[2] OMS, “Artritis reumatoide.” Accessed: Feb. 18, 2025. [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/rheumatoid-arthritis

[3] W. Pérez, “60% de personas mayores de 50 años padece de reumatismo - La Razón.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.la-razon.com/sociedad/2012/09/11/60-de-personas-mayores-de-50-anos-padece-de-reumatismo/

[4] A. Trigo, “Prevención: dos perspectivas para la artritis reumatoide | Los Tiempos.” Accessed: Feb. 18, 2025. [Online]. Available: https://www.lostiempos.com/oh/tendencias/20190520/prevencion-dos-perspectivas-artritis-reumatoide?utm_source=chatgpt.com

[5] M. Bressler *et al.*, “A Virtual Reality Serious Game for the Rehabilitation of Hand and Finger Function: Iterative Development and Suitability Study,” *JMIR Serious Games*, vol. 12, p. e54193, Jan. 2024, doi: 10.2196/54193.

[6] R. M. Viglialoro *et al.*, “A Projected AR Serious Game for Shoulder Rehabilitation Using Hand-Finger Tracking and Performance Metrics: A Preliminary Study on Healthy Subjects,” *Electronics 2023, Vol. 12, Page 2516*, vol. 12, no. 11, p. 2516, Jun. 2023, doi: 10.3390/ELECTRONICS12112516.

[7] H. Khan, “Sensor de célula de carga de medio puente de 50 kg - Hoja de datos Hub.” Accessed: Apr. 07, 2025. [Online]. Available: https://www.datasheethub.com/50kg-half-bridge-load-cell-sensor/

[8] P. By ALLDATASHEETCOM, “24-Bit Analog-to-Digital Converter (ADC) for Weigh Scales DESCRIPTION.”

[9] Espressif Systems, “ESP32 Series Datasheet Version 5.2 2.4 GHz Wi-Fi + Bluetooth ® + Bluetooth LE SoC Including,” 2025. [Online]. Available: www.espressif.com

[10] K-Invent, “Technical characteristics of the KFORCE Grip,” 2020.

[11] P. Gutiérrez, “Introducción a Three.js, la librería 3D número uno para HTML5.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.genbeta.com/desarrollo/introduccion-a-three-js-la-libreria-3d-numero-uno-para-html5

[12] A. Peña Arrebola, “Manual de Ayuda al Ejercicio para personas con Artritis Reumatoide,” 2012.

AUTORES



Ervin Emanuel Orellana Pardo é Engenheiro Biomédico, com formação em instrumentação médica, imagens médicas, processamento de sinais biomédicos, biomecânica, inteligência artificial aplicada à saúde e engenharia clínica. Desde 2023, é membro ativo da Sociedade Científica de Estudantes de Engenharia Biomédica, participando de atividades acadêmicas e de pesquisa. Possui experiência no desenvolvimento de páginas *web* e na aplicação de tecnologias voltadas à área da saúde. Além disso, é cofundador da WEXI, uma plataforma focada na economia circular de componentes eletrônicos na Bolívia, orientada à reutilização tecnológica e à redução de resíduos eletrônicos.



Bianca Camila Garnica Pericon é Engenheira Biomédica formada pela Universidad Privada del Valle - UNIVALLE, em Cochabamba, Bolívia (2025). Desde 2023, é membro ativa da Sociedade de Estudantes de Engenharia Biomédica (SCEIB) da UNIVALLE e integra a UMA, unidade de pesquisa tecnológica da universidade. Seu projeto — um robô para o autismo — foi reconhecido com o Primeiro Lugar como Melhor Projeto Internacional na FETIN 2024 (INATEL, Brasil), Primeiro Lugar em Inovação Transformadora nos prêmios KAMAY 2025 da Federação de Entidades Empresariais Privadas de Cochabamba, e o Prêmio de Excelência em Inovação Tecnológica (UNIVALLE, 2024). É cofundadora da Wexi, plataforma dedicada à economia circular de componentes eletrônicos na Bolívia.

Arquitetura inteligente baseada em LoRa para proteção ativa de gado bovino mediante monitoramento de tensão em tempo real

Roly Serrano Quiroz, Eynar Calle Viles & Katerin Deysi Vasquez Benito.
Universidad Privada del Valle – Univalle
sqr0028546@est.univalle.edu, ecallevi@univalle.edu, vbk2018762@est.univalle.edu
https://orcid.org/0009-0003-2767-4275, https://orcid.org/0000-0003-1477-2813,
https://orcid.org/0000-0001-9579-9035

Abstract— Perimeter security in rural livestock farming faces challenges arising from the structural deterioration of wire fences, acts of vandalism, animal predation, and technological infrastructure limitations in remote areas [1], [2]. This work presents the design and implementation of a distributed IoT system for tension monitoring and automatic tensioning of livestock fences based on LoRa communication.

The proposed architecture integrates sensor nodes with ESP32 microcontrollers, high-precision load cells (±0,03 % FS), analog conditioning via the INA125P, and 24-bit digitization, allowing real-time detection of changes in the fence wire. The system includes a web platform with MySQL database storage and visualization through a graphic *dashboard*, as well as an automatic tension correction mechanism using a stepper motor.

Controlled tests demonstrated technical feasibility in rural environments, long-range wireless transmission capability, and stable operation of the alert system. The proposed solution represents an accessible, scalable, and adaptable alternative for livestock protection in regions with limited technological infrastructure.

Index Terms—*LoRa, livestock protection, smart fence, tension monitoring.*

Resumo—A segurança perimetral na pecuária rural enfrenta desafios decorrentes da deterioração estrutural das cercas de arame, atos de vandalismo, predação animal e limitações de infraestrutura tecnológica em áreas remotas [1], [2]. Este trabalho apresenta o projeto e a implementação de um sistema IoT distribuído para o monitoramento de tensão e o tensionamento automático de cercas pecuárias baseado em comunicação LoRa.

A arquitetura proposta integra nós sensores com microcontroladores ESP32, células de carga de alta precisão (±0,03 % FS), condicionamento analógico por meio do INA125P e digitalização de 24 bits, permitindo a detecção em tempo real de alterações no arame da cerca. O sistema inclui uma plataforma web com armazenamento em banco de dados MySQL e visualização por meio de *dashboard* gráfico, bem como um mecanismo de correção automática de tensão utilizando um motor de passo.

Os testes controlados demonstraram viabilidade técnica em ambientes rurais, capacidade de transmissão sem fio de longo alcance e funcionamento estável do sistema de alerta. A solução proposta representa uma alternativa acessível, escalável e adaptável para a proteção do gado em regiões com infraestrutura tecnológica limitada.

Palavras chave— Cerca inteligente, LoRa, monitoramento de tensão, proteção pecuária.

I. Introdução

O setor pecuário representa aproximadamente 40% do valor total da produção agrícola mundial, segundo o International Fund for Agricultural Development (IFAD), desempenhando papel fundamental na segurança alimentar e na geração de empregos em áreas rurais [1]. Nesse contexto, a integridade das cercas perimetrais é um elemento crítico para prevenir perdas associadas ao roubo de gado, ataques de predadores e extravios acidentais [2]. No entanto, os métodos tradicionais de supervisão baseiam-se em inspeções manuais periódicas, o que implica elevados custos operacionais e capacidade limitada de detecção oportuna de falhas estruturais.

O desenvolvimento da Internet das Coisas (IoT) e das redes de comunicação de baixo consumo e longo alcance (LPWAN) abriu novas possibilidades para o monitoramento remoto de infraestruturas rurais [3], [4], [5]. Entre essas tecnologias, LoRa (*Long Range*) destaca-se por sua capacidade de transmitir dados a longas distâncias com baixo consumo energético, sendo uma alternativa viável para ambientes com infraestrutura de telecomunicações limitada [8], [9]. Sensores baseados em extensômetros (*strain gauges*) configurados em ponte de Wheatstone permitem detectar variações mínimas de tensão mecânica com alta precisão [6], [7].

Este trabalho apresenta o projeto e a implementação de um sistema IoT distribuído para monitoramento de tensão e tensionamento automático de cercas pecuárias utilizando tecnologia LoRa. A principal contribuição é uma arquitetura integrada capaz de detectar, comunicar e corrigir falhas estruturais em tempo real, diferenciando-se das soluções existentes ao combinar monitoramento contínuo com mecanismos eletromecânicos de correção automática em ambientes rurais remotos.

II. Arquitetura do Sistema

A. Funcionamento geral do sistema

A arquitetura implementada é modular e distribuída, composta por nós sensores de campo e uma unidade central de monitoramento. Cada nó realiza aquisição de dados, processamento local e transmissão sem fio para a plataforma de visualização, permitindo supervisão contínua e descentralizada do estado estrutural do arame da cerca. A arquitetura geral do sistema e a interconexão entre seus componentes são apresentadas na Figura 1.

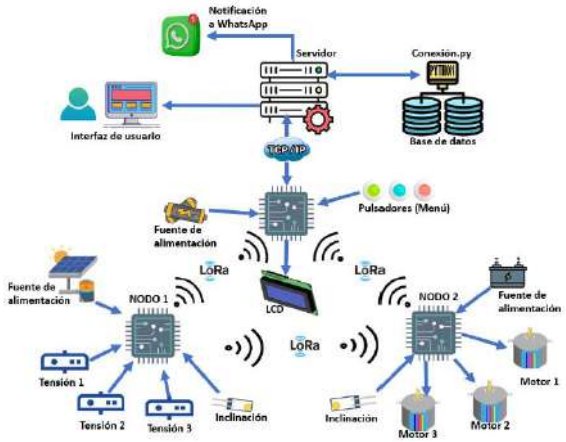


Fig. 1. Diagrama Geral do sistema

A arquitetura fundamenta-se em uma rede distribuída de nós sensores inteligentes, projetados para operar em ambientes rurais caracterizados por condições ambientais variáveis e limitações de infraestrutura tecnológica. Cada nó integra sensores de tensão baseados em células de carga de alta sensibilidade, capazes de detectar variações mínimas na força aplicada sobre os fios da cerca. Esses dispositivos permitem identificar tanto mudanças graduais associadas a fenômenos físicos, como dilatação térmica ou relaxamento mecânico do material, quanto alterações abruptas que poderiam indicar tentativas de manipulação, corte ou impacto do gado contra a estrutura.

O processamento local é executado por microcontroladores, que aplicam critérios de detecção de anomalias antes de transmitir as informações. A comunicação entre os nós e a unidade central é realizada por meio da tecnologia LoRa, selecionada por seu baixo consumo energético e capacidade de cobertura estendida em ambientes rurais.

As informações consolidadas são armazenadas em um servidor e visualizadas por meio de uma plataforma web, a partir da qual são gerenciados alertas e parâmetros de operação. Adicionalmente, o sistema incorpora um módulo de tensionamento automático que permite corrigir desvios de tensão detectados, transformando a infraestrutura tradicional em um sistema ativo de monitoramento e autorregulação.

B. Nó sensor

O nó sensor constitui a unidade fundamental de aquisição e processamento do sistema distribuído. Seu projeto integra componentes eletrônicos orientados a garantir precisão na medição da tensão do alambrado e estabilidade operacional em ambientes rurais. Cada nó incorpora um microcontrolador ESP32, selecionado por sua capacidade de processamento, baixo consumo energético e compatibilidade com múltiplas interfaces de comunicação. As principais especificações técnicas dos componentes empregados são resumidas na Tabela I.

TABELA I
ESPECIFICAÇÕES TÉCNICAS DO NÓ SENSOR

COMPONENTE	MODELO	ESPECIFICAÇÃO
Microcontrolador	ESP32	Dual core 240 MHz
Célula de carga	5 Kg	±0.03% FS
ADC	ADS 1220	24 bits
Amplificador	INA125P	Baixo ruído
Comunicação	SX1278	433 MHz

A medição de tensão é realizada por meio de células de carga do tipo viga em balanço, com capacidade nominal de 5 kg e sensibilidade aproximada de 2 mV/V. Essas operam segundo o princípio de extensômetros (*strain gauges*) configurados em ponte de Wheatstone, no qual a deformação mecânica do elemento gera variações resistivas proporcionais à força aplicada. Devido à baixa amplitude do sinal de saída, implementa-se um amplificador instrumental INA125P para o condicionamento do sinal, garantindo alta relação sinal-ruído e estabilidade térmica.

Posteriormente, o sinal é filtrado por meio de um sistema anti-aliasing de segunda ordem, com frequência de corte de 10 Hz, e digitalizado por um conversor analógico-digital de 24 bits, permitindo alta resolução na detecção de variações mínimas de tensão.

De forma complementar, o nó integra sensores de inclinação baseados em tecnologia MEMS, que permitem detectar desvios angulares nos postes da cerca. Essa dupla medição aumenta a confiabilidade do sistema ao possibilitar a identificação tanto de perda progressiva de tensão quanto de possíveis deslocamentos estruturais. O processamento local inclui algoritmos de detecção de anomalias baseados em limiares calibrados, reduzindo transmissões desnecessárias e otimizando o consumo energético.

C. Sistema de tensionamento automático

O sistema de tensionamento automático constitui o componente ativo de correção estrutural do alambrado. Seu objetivo é restabelecer condições ótimas de tensão quando o sistema detecta valores abaixo do limiar estabelecido. Esse módulo integra um motor de passo 28BYJ-48 acoplado mecanicamente ao sistema de enrolamento do fio, permitindo aplicar ajustes controlados de tensão. O motor de passo 28BYJ-48 opera a 5 V com 2.048 passos por revolução e velocidade

configurada em 10 RPM, controlado por driver ULN2003. A ativação ocorre quando a tensão detectada é inferior a 3,0 N, e o sistema é desativado ao superar esse limiar, evitando sobreajustes.

O controle do motor é realizado por meio de um driver ULN2003, gerenciado pelo microcontrolador do nó sensor. A ativação do mecanismo ocorre somente quando é verificado um desvio persistente fora da faixa operacional, evitando sobreajustes ou ativações desnecessárias. O sistema implementa ciclos de correção controlados em tempo e número de passos, seguidos de uma verificação posterior da nova tensão medida.

A Figura 2 apresenta a integração mecânica das células de carga no poste sensor. A montagem permitiu transmitir a força exercida pelo fio diretamente ao elemento sensorial, garantindo linearidade na medição.

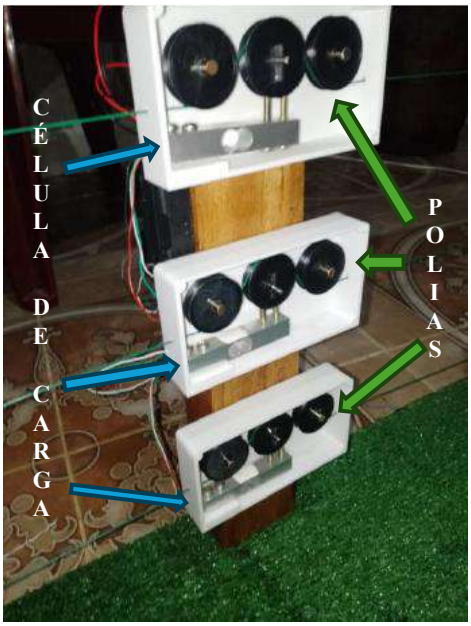


Fig. 2. Integração mecânica do sistema de medição de tensão por meio de células de carga.

Essa abordagem transforma o cercado tradicional em um sistema ciberfísico capaz não apenas de detectar falhas, mas também de intervir automaticamente diante da perda progressiva de tensão, causada por fenômenos como dilatação térmica, relaxamento mecânico do material ou pequenas deformações. Dessa forma, aumenta-se a autonomia operacional do sistema e reduz-se a necessidade de intervenção manual.

D. Plataforma de monitoramento

A plataforma de monitoramento constitui a interface de interação entre o sistema e o usuário final. A arquitetura de software baseia-se em um servidor local implementado no ambiente XAMPP, utilizando Apache como servidor web e MySQL para o armazenamento estruturado de dados históricos.

A comunicação de dados a partir da unidade central é realizada por meio do protocolo MQTT, permitindo uma transmissão eficiente e estruturada das informações.

A interface gráfica foi desenvolvida como um *dashboard* web responsivo, acessível a partir de dispositivos móveis e computadores de mesa (Figura 3). Essa plataforma apresenta visualização em tempo real dos valores de tensão, gráficos históricos, estados de alerta e localização dos nós implantados em campo. Além disso, permite a configuração remota de parâmetros operacionais, incluindo o ajuste de limiares de detecção e a frequência de amostragem.



Fig. 3. *Dashboard* de monitoramento de tensão.

O sistema incorpora um módulo de notificações multicanal que permite gerar alertas automáticos diante de eventos críticos, como queda abrupta de tensão ou inclinação anormal. Essas notificações são enviadas por meio de mensagens via WhatsApp, facilitando uma resposta oportuna por parte do produtor (Figura 4.).

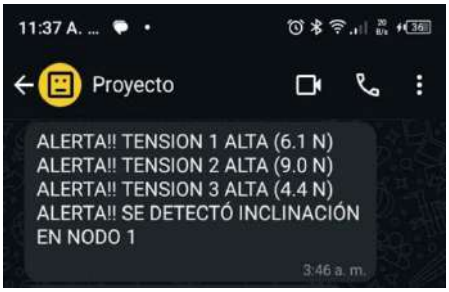


Fig. 4. Mensagens de alerta via WhatsApp.

IV. Resultados

Os testes demonstraram que o sistema detecta alterações na tensão do alambrado com alta sensibilidade, permitindo a geração de alertas em tempo real. A comunicação LoRa apresentou estabilidade em ambientes rurais abertos, alcançando distâncias significativas sem necessidade de infraestrutura adicional. O mecanismo de tensionamento automático conseguiu restabelecer níveis adequados de tensão

em cenários simulados de perda estrutural leve. As principais parãmetros empregados são resumidas na Tabela II.

TABELA II
Parãmetro e Valor medido

PARÂMETRO	VALOR MEDIDO
Tensão T1 (operação normal)	4,15 – 4,97 N
Tensão T2 (operação normal)	8,59 – 12,51 N
Tensão T3 (operação normal)	9,16 – 14,70 N
Limiar de ativação do motor	< 3,0 N
Intervalo de transmissão LoRa	1 seg.
Latência de rede	< 500 ms

O sistema foi implementado em um ambiente controlado que simulou uma configuração real de cercado perimetral. A Figura 6 mostra o protótipo completamente instalado, incluindo postes estruturais, linhas de tensão, nós sensores e a unidade central de monitoramento. Essa configuração permitiu avaliar o comportamento do sistema diante de variações controladas de tensão e alterações mecânicas.

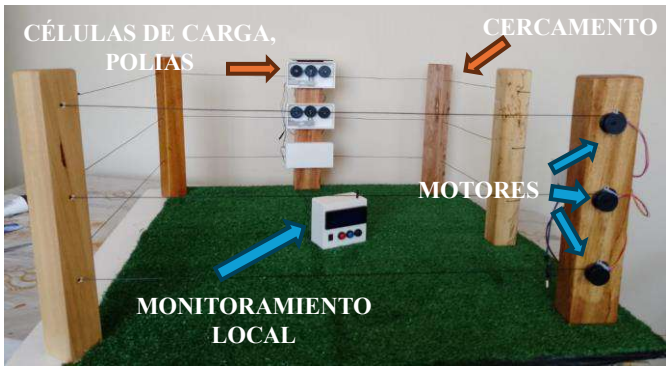


Fig. 5. Protótipo do sistema de monitoramento de tensão e tensionamento automático.

No que se refere à visualização e supervisão das informações, o sistema permitiu duplo nível de monitoramento: local e remoto. Em nível local (Figura 6), a tela LCD instalada na unidade central exibiu em tempo real os valores individuais de tensão (T1, T2 e T3) e o estado do sensor de inclinação, facilitando a verificação imediata do funcionamento do sistema durante os testes experimentais. Essa interface permitiu validar a correta aquisição e processamento de dados diretamente em campo, sem necessidade de conexão externa.



Fig. 6. Visualização em tempo real em nível local.

De forma complementar, o *sa* web apresentou as informações consolidadas por meio de gráficos dinâmicos, tabelas históricas e estados de alerta, permitindo analisar a evolução temporal da tensão em cada linha do alambrado. O armazenamento estruturado em banco de dados possibilitou a rastreabilidade de eventos, identificação de padrões de variação e avaliação comparativa entre períodos de operação normal e situações de perturbação, fortalecendo o caráter preventivo do sistema (Figura 7).

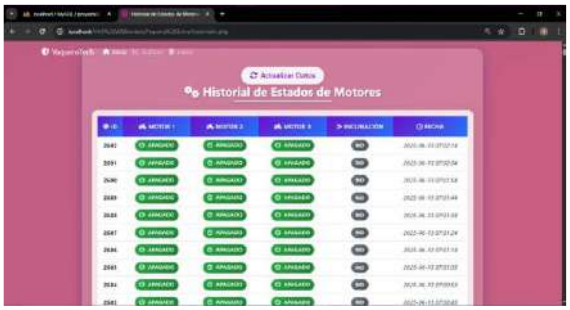


Fig. 7. Dados históricos armazenados.

A análise de custos evidenciou um investimento acessível em comparação com soluções comerciais especializadas, favorecendo sua adoção por pequenos e médios produtores. Em conjunto, os resultados confirmam a viabilidade técnica do sistema como solução integral de monitoramento estrutural ativo, capaz de combinar detecção precoce, transmissão confiável e ação corretiva automatizada em um ambiente rural simulado.

V. Discussão

A solução proposta demonstra uma melhoria substancial em relação aos métodos tradicionais de supervisão manual do alambrado, que dependem de inspeções periódicas e não permitem detecção precoce de falhas estruturais. A

incorporação de monitoramento contínuo e capacidade de resposta automatizada posiciona o sistema dentro do paradigma dos sistemas ciberfísicos aplicados ao setor agropecuário, alinhando-se às tendências de digitalização descritas na literatura sobre Internet das Coisas (IoT) e agricultura inteligente [5], [10].

Do ponto de vista estrutural, a implementação de células de carga como mecanismo de monitoramento transforma o cercado em um sistema de avaliação permanente da integridade mecânica, podendo ser considerada uma aplicação específica do conceito de *Structural Health Monitoring* adaptado à infraestrutura rural. Essa abordagem é coerente com os métodos modernos de sensoriamento distribuído descritos por Fraden [7], nos quais a combinação de alta resolução na aquisição e processamento local melhora a confiabilidade dos dados e reduz falsas detecções.

Quanto à comunicação, a escolha da tecnologia LoRa responde às recomendações de redes LPWAN para ambientes rurais com baixa densidade de infraestrutura [10]. Estudos como os de Vangelista [8] e Adelantado et al. [9] destacam a robustez da modulação *Chirp Spread Spectrum* frente a interferências e sua capacidade de alcançar longas distâncias com baixo consumo energético, características que se refletem no desempenho observado durante os testes experimentais. No entanto, como relatado em diversos artigos científicos, fatores como topografia irregular, vegetação densa e obstáculos físicos podem afetar a qualidade do enlace e o alcance efetivo [8], o que coincide com as limitações identificadas durante a implementação.

Em termos operacionais, a necessidade de calibração inicial das células de carga constitui uma etapa crítica para garantir precisão na medição, conforme descrito na instrumentação baseada em galgas extensométricas [6], [7]. Embora esse requisito implique uma intervenção técnica inicial, não representa uma limitação estrutural do sistema, mas sim uma condição inerente aos sistemas de medição de alta sensibilidade.

A arquitetura modular e distribuída do sistema facilita sua escalabilidade e adaptação a diferentes configurações de cercado, permitindo ampliar o número de nós sem modificar a estrutura central. Esse enfoque é consistente com os princípios de projeto de plataformas IoT escaláveis descritos por Stankovic [5] e com as aplicações de digitalização agrícola revisadas por Wolfert et al. [3], nas quais a interoperabilidade e a expansão progressiva são elementos fundamentais para a adoção tecnológica.

Por fim, o componente de tensionamento automático introduz um elemento diferenciador em relação a soluções exclusivamente de monitoramento. Ao integrar detecção e ação corretiva no mesmo sistema, a proposta ultrapassa o conceito de supervisão passiva e aproxima-se de um modelo de infraestrutura autônoma com capacidade de autorregulação. Essa característica amplia o impacto potencial do sistema, reduzindo o tempo de resposta diante de falhas e diminuindo a dependência de intervenção manual em zonas rurais remotas.

Em conjunto, os resultados respaldam a viabilidade técnica e conceitual do sistema no contexto da pecuária inteligente, evidenciando que a integração de sensoriamento estrutural, processamento distribuído e comunicação LPWAN constitui uma estratégia eficaz para modernizar a proteção perimetral em ambientes rurais.

VI. Conclusão

O projeto permitiu desenvolver e implementar um sistema automatizado de monitoramento e tensionamento de cercas, orientado a melhorar a proteção pecuária em ambientes rurais. A solução integra sensores de tensão baseados em células de carga, microcontroladores ESP32 para processamento local e comunicação sem fio de longo alcance, possibilitando a detecção de variações estruturais no alambrado e a geração de alertas em tempo real.

Durante os testes experimentais, o sistema conseguiu identificar mudanças de tensão em múltiplas linhas da cerca e transmitir essas informações para a unidade central por meio de enlaces sem fio estáveis, com tempos de atualização da ordem de segundos e visualização imediata na interface de monitoramento.

A arquitetura distribuída demonstrou funcionamento estável em cenários rurais simulados, nos quais os nós sensores realizaram aquisição e processamento local de dados antes da transmissão, reduzindo o tráfego de comunicação e otimizando o consumo energético. A integração de sensores de tensão, sensores de inclinação e módulos de comunicação permitiu detectar tanto variações mecânicas nos fios quanto possíveis deslocamentos estruturais dos postes, aumentando a confiabilidade do sistema de vigilância perimetral.

Da mesma forma, o módulo de tensionamento automático, baseado em motores de passo controlados por sinais, permitiu restabelecer níveis adequados de tensão em situações simuladas de perda estrutural leve, demonstrando a viabilidade de incorporar mecanismos de correção automática dentro do sistema.

A plataforma de monitoramento implementada facilitou a visualização dos valores de tensão e dos estados do sistema por meio de uma interface web e uma tela LCD local, permitindo registrar históricos de dados, gerar alertas imediatos e supervisionar o estado do cercado em tempo real.

A análise de custos demonstrou que o protótipo pode ser desenvolvido com componentes eletrônicos de baixo custo e alta disponibilidade, favorecendo sua adoção por pequenos e médios produtores.

Embora tenham sido identificados desafios associados à topografia na propagação do sinal sem fio e à necessidade de calibração inicial dos sensores para garantir medições precisas, os resultados obtidos validam a viabilidade técnica e o potencial

de escalabilidade do sistema. Como trabalho futuro, recomenda-se fortalecer a autonomia energética por meio de sistemas solares otimizados, melhorar a cobertura de comunicação em terrenos complexos mediante repetidores ou redes híbridas, e incorporar algoritmos inteligentes para classificação de eventos e redução de falsos alarmes.

Em conjunto, a proposta demonstra que é possível transformar uma infraestrutura perimetral tradicional em um sistema ciberfísico inteligente, capaz de integrar sensoriamento, comunicação e ação corretiva automática, contribuindo para a modernização tecnológica e a melhoria da segurança pecuária em comunidades rurais.

VII. AGRADECIMENTOS

Este trabalho conta com o apoio do Centro de Produção UMA, pertencente à Universidad Privada del Valle, Cochabamba – Bolívia.

REFERÊNCIAS

[1] «Sector ganadero», FIDA. Acessado: 24 de febrero de 2026. [On-line]. Disponível em: https://www.ifad.org/en/livestock

[2] «DEACERO Blog | Agropecuario». Acessado: 24 de febrero de 2026. [On-line]. Disponivel em: https://blog.deacero.com/tag/agropecuario

[3] S. Wolfert, L. Ge, C. Verdouw, y M.-J. Bogaardt, «Big Data in Smart Farming – A review», *Agric. Syst.*, vol. 153, pp. 69-80, may 2017, doi: 10.1016/j.agsy.2017.01.023.

[4] M. A. Zamora-Izquierdo, J. Santa, J. A. Martínez, V. Martínez, y A. F. Skarmeta, «Smart farming IoT platform based on edge and cloud computing», *Biosyst. Eng.*, vol. 177, pp. 4-17, ene. 2019, doi: 10.1016/j.biosystemseng.2018.10.014.

[5] J. A. Stankovic, «Research Directions for the Internet of Things», *IEEE Internet Things J.*, vol. 1, n.º 1, pp. 3-9, feb. 2014, doi: 10.1109/JIOT.2014.2312291.

[6] E. C. Ortiz, «An Introduction to Measurements using Strain Gages», Accedido: 25 de febrero de 2026. [En línea]. Disponible en: https://www.academia.edu/24444174/An_Introduction_to_Measurements_using_Strain_Gages

[7] «Handbook of Modern Sensors: Physics, Designs, and Applications | Springer Nature Link». Accedido: 25 de febrero de 2026. [En línea]. Disponible en: https://link.springer.com/book/10.1007/978-1-4419-6466-3

[8] L. Vangelista, «Frequency Shift Chirp Modulation: The LoRa Modulation», *IEEE Signal Process. Lett.*, vol. 24, n.º 12, pp. 1818-1821, dic. 2017, doi: 10.1109/LSP.2017.2762960.

[9] F. Adelantado Freixer, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melià-Seguí, y T. Watteyne, «Understanding the limits of LoRaWAN», *IEEE Commun. Mag.*, vol. 55, jun. 2017, doi: 10.1109/MCOM.2017.1600613.

[10] U. Raza, P. Kulkarni, y M. Sooriyabandara, «Low Power Wide Area Networks: An Overview», *IEEE Commun. Surv. Tutor.*, vol. 19, n.º 2, pp. 855-873, 2017, doi: 10.1109/COMST.2017.2652320.

AUTORES



Roly Serrano Quiroz é estudante pesquisador de Engenharia Eletrônica na Universidade Privada del Valle (Univalle, Bolívia). Sua formação acadêmica tem se especializado no desenvolvimento de soluções integrais que combinam hardware, software e redes. Possui experiência prática na integração de sensores e *gateways* LoRaWAN para monitoramento remoto, destacando a implementação de sistemas inteligentes aplicados ao monitoramento de gado bovino. Atualmente, realiza suas atividades de estágio na Nuevatel PCS da Bolívia. Suas áreas de interesse incluem o desenvolvimento de sistemas embarcados, a Internet das Coisas (IoT) e a gestão de projetos utilizando metodologias ágeis (*Scrum*).



Eynar Calle Viles é mestre em Telecomunicações, com especialização em Redes de Comunicações e Telemática, e possui pós-graduação em Internet das Coisas (IoT) pelo Instituto Nacional de Telecomunicações (INATEL), Brasil.

Com mais de 10 anos de experiência, é especialista em projeto e implementação de redes, FTTH, IoT e sistemas embarcados. Desenvolveu diversas inovações tecnológicas, incluindo propostas de patentes, como o "Robô Móvel para Detecção de Metais" e o "Sistema de Localização para o Satélite Tupac Katari".

Além disso, possui várias publicações e participações em revistas e congressos internacionais. Suas principais áreas de interesse incluem antenas e propagação, segurança da informação e IoT, contribuindo ativamente para o desenvolvimento de soluções tecnológicas inovadoras.

Atualmente, atua como coordenador de pesquisa na Faculdade de Informática e Eletrônica da Universidad Privada del Valle (Univalle) e é presidente da IEEE Computer Society Bolivia.



Katerin Deysi Vasquez Benito é Engenheira Biomédica pela Universidade Privada del Valle (Univalle, Bolívia). Sua formação acadêmica tem se concentrado no desenvolvimento e implementação de sistemas eletrônicos inteligentes aplicados à área da saúde e a ambientes produtivos. Possui experiência em pesquisa e desenvolvimento de sistemas embarcados, integração de sensores, aquisição e processamento de sinais, bem como na arquitetura de plataformas IoT para monitoramento remoto. Suas principais áreas de interesse incluem sistemas de monitoramento inteligente, Internet das Coisas, instrumentação biomédica e industrial, e análise de dados aplicados.

Mobile Application for Medicinal Plant Leaf Species Identification and Disease Detection Using Deep Learning

Maria Laura Nuñez Jaillita, Juan Abiel Iriarte Colque & Eynar Calle Viles
Universidad Privada del Valle – UNIVALLE
njm0031533@est.univalle.edu, icj0026532@est.univalle.edu, ecallevi@univalle.edu
https://orcid.org/0009-0006-3400-1353, https://orcid.org/0009-0001-9123-0287,
https://orcid.org/0000-0003-1477-2813

Abstract— Medicinal plants play an important role in traditional medicine and pharmacological research, particularly in regions where access to conventional treatments is limited. However, the correct identification of plant species and the early detection of diseases in leaves remain complex tasks for non-specialized users. Advances in computer vision and deep learning have enabled the development of systems capable of recognizing plant characteristics from digital images.

This work presents a system for the identification of medicinal plant species and the preliminary assessment of their health status using deep learning techniques. To achieve this, a curated dataset containing images of five plant species was constructed. Different convolutional neural network architectures were trained and compared, including a baseline CNN and transfer learning–based models such as MobileNetV2 and EfficientNetB0. Additionally, a binary classification model was explored to detect the presence of visible disease symptoms on leaves.

The experimental results show that transfer learning–based approaches significantly outperform the CNN trained from scratch. In particular, EfficientNetB0 achieved an accuracy of 79.17% in species classification. Finally, the best-performing model was integrated into a client–server architecture that enables near real-time predictions through a mobile application.

Index Terms— Computer Vision, Deep Learning, Medicinal Plants, Plant Disease Detection.

I. Introduction

Medicinal plants play a fundamental role in traditional medicine and pharmacological research, especially in regions where access to conventional treatments is limited [1]. However, the correct identification of plant species and the early detection of diseases in leaves remain complex tasks for non-specialized users. Factors such as visual similarity between species, environmental conditions, and the presence of symptoms on leaves can make this identification process more difficult [2].

In recent years, advances in computer vision and deep learning have enabled the development of systems capable of identifying plant species and detecting diseases from digital images [3]. In particular, convolutional neural networks (CNNs) have demonstrated strong performance in image classification tasks, allowing the automatic extraction of relevant visual features [4]. Modern architectures such as MobileNet and EfficientNet have been widely used due to their ability to achieve high levels of accuracy with adequate computational efficiency [5].

Despite these advances, many studies focus mainly on the evaluation of classification models without considering their integration into practical tools for end users. This limitation reduces their impact in real-world contexts, where farmers, students, or interested users could benefit from automated systems for plant identification and preliminary assessment of plant health.

In this context, this work proposes an integrated system for the identification of medicinal plant species and the preliminary evaluation of their health status using deep learning techniques. The approach includes the construction of a dataset containing five plant species, the training and comparison of different neural network architectures for species classification and leaf health status detection, and the implementation of a predictive system accessible through a mobile application based on a client–server architecture.

II. Dataset and model training process

A. Dataset

The dataset used in this study was constructed through the integration and reorganization of images from multiple public medicinal plant datasets [6] – [12]. Subsequently, the dataset was curated and published on Kaggle to facilitate the reproducibility of the study and its reuse in future research [13].

The dataset contains 13,167 images corresponding to five plant species: Aloe barbadensis Miller, Azadirachta indica, Hibiscus rosa-sinensis, Moringa oleifera, and Ocimum sanctum. The images include healthy leaves, leaves showing disease symptoms, natural backgrounds, and complete plants captured under different environmental conditions, which supports the training of more robust models. This dataset was used for two tasks: species classification and plant health status detection.

a.1) Species Dataset

The species dataset was constructed from the previously curated dataset by organizing the images according to species for the classification task. Each species is stored in an independent directory containing images of leaves and plants

captured under different environmental conditions and orientations, allowing the model to learn more robust visual features. The distribution of images per species is shown in Table I.

TABLE I
DISTRIBUTION OF IMAGES IN THE SPECIES DATASET

SPECIES	IMAGES
Aloe barbadensis Miller	3017
Azadirachta indica	3608
Hibiscus rosa-sinensis	2228
Moringa oleifera	2595
Ocimum sanctum	1719
Total	13167

a.2) Health Dataset

The health dataset was created for plant health status detection by reorganizing the same image collection used in the species dataset into two categories: healthy and unhealthy. However, not all images were suitable for this task, as some source datasets lacked reliable health condition labels. To avoid subjective labeling and potential misclassification, only images with clearly identifiable conditions were selected.

The dataset maintains the species-based structure with this binary subdivision, where the unhealthy class represents visible symptoms of disease or damage. This decision prioritizes label reliability over dataset size, which is critical for supervised health classification tasks.

This binary classification approach simplifies the problem and allows the model to focus on detecting the presence of plant conditions affecting its health. The distribution of images by species and health category is shown in Table II. In total, the dataset contains 9,393 images, of which 2,734 correspond to healthy plants and 6,659 to plants showing signs of disease or damage.

TABLE II
DISTRIBUTION OF IMAGES IN THE HEALTH DATASET

SPECIES	HEALTHY	UNHEALTHY	TOTAL
Aloe barbadensis miller	780	1720	2500
Azadirachta indica	519	2025	2544
Hibiscus rosa-sinensis	473	652	1125
Moringa oleifera	464	1764	2228
Ocimum sanctum	498	498	996
Total	2734	6659	9393

B. Data Preprocessing and Augmentation

Before training, the images were subjected to a preprocessing and data augmentation process to improve input quality and the models' generalization capability. All images were resized to a uniform resolution of 224 × 224 pixels, and pixel values were normalized using a 1/255 factor, transforming the original range [0,255] to [0,1], which improves stability during training.

To increase dataset diversity and reduce the risk of overfitting, data augmentation techniques were applied using ImageDataGenerator from Keras, including rotations of up to 20°, random zoom, horizontal flipping, shear transformations, and horizontal and vertical shifts.

The dataset was divided into training and validation sets (80/20) using the validation_split parameter, while an additional independent test set was reserved for final evaluation. During training, a batch size of 16 images was used, balancing training stability and efficient use of computational resources.

C. Training Configuration

The models were trained using TensorFlow/Keras. For optimization, the Adam algorithm was employed, using categorical cross-entropy for species classification and binary cross-entropy for plant health status detection.

During training, a batch size of 16 images was used with inputs resized to 224 × 224 pixels. The models were trained for 15 epochs, using the validation set to monitor performance. The experiments were conducted in Google Colab with GPU support, which allowed the training process to be accelerated.

D. Species Classification Models

For the species classification task, three convolutional neural network architectures were evaluated in order to analyze their ability to identify the five species present in the dataset. A baseline Convolutional Neural Network (CNN) model designed for this study was implemented and compared with two transfer learning–based architectures, MobileNetV2 and EfficientNetB0.

In the MobileNetV2 and EfficientNetB0 models, weights pretrained on ImageNet were used, replacing the final classification layer with a dense layer with Softmax activation adapted to the five classes. This comparison made it possible to evaluate the performance of a simple convolutional network against pretrained models widely used in image classification tasks.

d.1) Baseline CNN

As an initial baseline, a basic convolutional neural network was implemented, composed of multiple convolutional layers followed by normalization, pooling operations, and dense layers for final classification. The model includes three convolutional blocks with filters of 32, 64, and 128 channels, respectively, each followed by Batch Normalization and MaxPooling, allowing the extraction of relevant visual features from the images.

Subsequently, the extracted features are flattened and processed through a dense layer of 128 neurons with ReLU activation. To improve model regularization and mitigate the risk of overfitting, a Dropout layer with a rate of 0.5 was incorporated.

d.2) MobileNetV2 (Transfer Learning Model)

As a second approach, MobileNetV2 was used, a lightweight convolutional neural network architecture optimized for computer vision tasks and devices with limited computational resources. The model was applied through transfer learning, loading weights pretrained on the ImageNet dataset and removing the original classification layer.

The convolutional base of the model was kept frozen during training to preserve the previously learned features. On top of this base, a custom classifier was added, consisting of a GlobalAveragePooling layer, followed by Batch Normalization, a dense layer of 128 neurons with ReLU activation, a Dropout layer with a rate of 0.5, and finally a Softmax output layer for species classification.

d.3) EfficientNetB0

Finally, the EfficientNetB0 architecture was evaluated, a convolutional neural network model designed to improve scaling efficiency across the network's depth, width, and resolution. As in the case of MobileNetV2, transfer learning was applied by loading weights pretrained on ImageNet and removing the top layer of the original model.

The base of the model was kept frozen, and a classification head was added consisting of a GlobalAveragePooling layer, Batch Normalization, a dense layer of 128 neurons with ReLU activation, and a Dropout layer with a rate of 0.5 before the final Softmax layer. This approach allows the model to leverage visual representations learned from large datasets while adapting it to the specific task of species classification.

E. Disease Detection Model

In addition to species classification, the developed system also addresses the detection of plant health status. For this task, the health dataset was used, which reorganizes the same images from the original dataset into a binary classification composed of the healthy and unhealthy categories.

The objective of this model is to detect the presence of visible symptoms of disease or damage on the leaves without differentiating the specific type of pathology. This approach simplifies the classification problem and allows the system to quickly identify whether a plant shows signs of deterioration, which can serve as an initial diagnostic tool for farmers or non-specialized users.

The health detection model uses an architecture similar to that employed for species classification, adapting the output layer to produce a binary prediction through a Softmax layer with two neurons corresponding to the healthy and unhealthy categories.

F. Model Evaluation

To evaluate the performance of the models, several metrics commonly used in image classification tasks were employed. These metrics include accuracy, precision, recall, and F1-score, which allow the analysis of model behavior from different perspectives.

Additionally, confusion matrices were generated to visualize the performance of the models in the classification of each class. These matrices make it possible to identify error patterns and analyze which classes present greater classification difficulty. Comparing these metrics across the different architectures allows the determination of which model offers the best balance between accuracy and generalization capability.

G. Implementation of Predictive System

With the objective of facilitating the practical use of the developed models, a predictive system based on a client–server architecture was implemented, allowing a mobile application to capture images of leaves and send them to a server for automatic analysis. The system integrates deep learning models for species identification and a preliminary plant health detection module, along with informational functionalities aimed at supporting users in the recognition and care of medicinal plants.

g.1) Inference Server

The inference server was developed in Python using the Flask framework, exposing a REST service responsible for receiving images from the mobile application and processing them using the previously trained models. The server loads into memory the EfficientNetB0 models, selected due to their superior performance in the experimental evaluation, to perform both species classification and plant health status detection.

When an image is sent to the system, the server applies the same preprocessing used during training, including RGB conversion, resizing to 224 × 224 pixels, and pixel value normalization. The inference process is then executed, obtaining the prediction of the plant species and its health status.

The results are returned to the client in JSON format, including the predicted class and its confidence level, allowing the mobile application to present the information to the user in a clear and immediate manner.

g.2) Mobile Application

The mobile application acts as the main interface for interaction with the system. Through the device's camera, the user can capture an image of a plant leaf or select a stored image for analysis. The application is responsible for sending the image to the inference server and receiving the response generated by the model. Fig. 1 shows the main interface of the application, where the user can capture or upload an image to initiate the identification process.

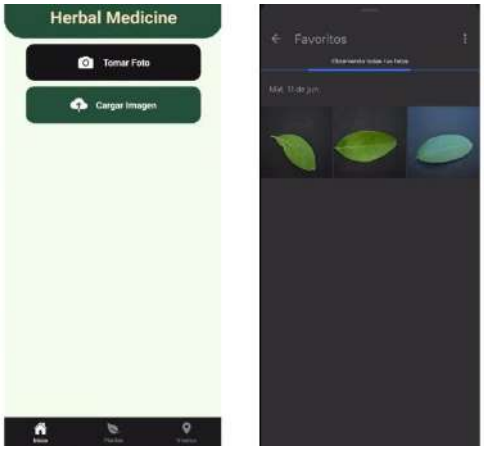


Fig. 1. Mobile application interface for capturing and selecting medicinal plant leaf images.

Once the response from the server is received, the application displays to the user the identified species, the health status of the leaf, and the confidence level of the prediction. Fig. 2 presents the visualization of the results within the application, including examples of leaves classified as healthy and unhealthy. This interaction enables near real-time results, facilitating plant identification and the preliminary detection of possible plant conditions.



Fig. 2. Mobile application displaying plant health diagnosis results (healthy and diseased leaves).

g.3) Information and Location Services

In addition to the automatic identification functionalities, the application incorporates complementary informational services aimed at improving the user experience and providing useful information about the analyzed plants. To manage this data, Firebase was used as a cloud database system, enabling the storage and retrieval of information related to plant species, medicinal properties, and usage recommendations.

III. Results and Discussion

A. Species Classification Results

To evaluate the performance of the species classification models, the metrics accuracy, precision, recall, and F1-score were calculated using the test set. Three different architectures were compared: a baseline CNN, MobileNetV2, and EfficientNetB0. Table III presents the results obtained by each model.

TABLE III
PERFORMANCE COMPARISON FOR SPECIES CLASSIFICATION

MODEL	ACCURACY	PRECISION	RECALL	F1-SCORE
CNN	0.4246	0.5271	0.4246	0.4332
MobileNetV2	0.7807	0.8727	0.7807	0.7729
EfficientNetB0	0.7917	0.8819	0.7917	0.7878

The results show that transfer learning–based models significantly outperform the CNN model trained from scratch. While the baseline CNN achieved an accuracy of 42.46%, MobileNetV2 and EfficientNetB0 reached values close to 79% accuracy, highlighting the advantage of using pretrained models that have already learned visual representations from large datasets such as ImageNet.

Among the evaluated models, EfficientNetB0 achieved the best overall performance, reaching an accuracy of 79.17% and the highest F1-score, indicating a better balance between precision and recall in species classification. This result suggests that the EfficientNet architecture is capable of capturing more discriminative visual features, enabling more accurate identification of the plant species present in the dataset.

B. Confusion Matrix Analysis

To analyze the behavior of the model in greater detail, confusion matrices were generated, allowing the observation of performance in the classification of each species and the detection of possible confusions between visually similar classes. Fig. 3 shows the confusion matrix obtained using the EfficientNetB0 model for species classification.

Additionally, the matrix makes it possible to identify certain error patterns among species with similar visual characteristics. Although most correct predictions are concentrated along the main diagonal, some confusions can be observed, which may be related to variations in lighting, leaf orientation, or similarities in shape and texture. This analysis helps identify potential future improvements for the model.

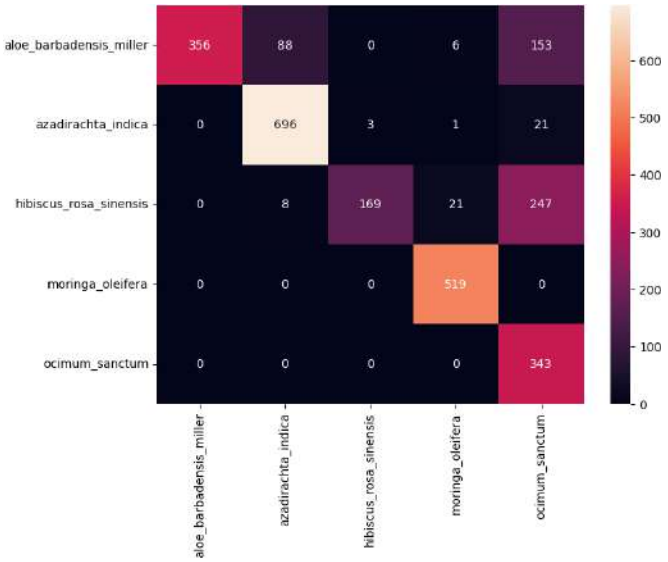


Fig. 3. Confusion matrix for species classification using EfficientNetB0.

The results show that some species are easier to classify than others. In particular, Moringa oleifera and Ocimum sanctum exhibit higher recognition rates, while Hibiscus rosa-sinensis shows greater confusion with other species, likely due to visual similarities in leaf shape and texture.

This analysis highlights that the classification of plant species can be affected by the visual variability of leaves and the morphological similarities between species, which

represents a common challenge in computer vision applications in agriculture.

C. Plant Health Detection Results

In addition to species classification, the performance of the model was evaluated for the plant health status detection task using the health dataset. Table IV presents the results obtained by the different evaluated models.

TABLE IV
PERFORMANCE COMPARISON FOR HEALTH DETECTION

MODEL	ACCURACY	PRECISION	RECALL	F1-SCORE
CNN	0.7093	0.5031	0.7093	0.5886
MobileNetV2	0.7087	0.8173	0.7087	0.7216
EfficientNetB0	0.7093	0.5031	0.7093	0.5886

Although the accuracy values are similar among the models, MobileNetV2 shows a better balance between precision and F1-score, indicating a greater ability to distinguish between healthy leaves and leaves showing signs of disease.

It is worth noting that the CNN and EfficientNetB0 models present identical performance metrics in Table IV. While this result may seem unusual, it is consistent with their final evaluation outputs. However, during training, both models exhibited different learning dynamics. The CNN showed greater variability in validation performance across epochs, whereas EfficientNetB0 demonstrated a more stable and consistent convergence behavior. This suggests that, despite reaching similar quantitative results, the internal learning processes and generalization patterns of the models differ.

It is important to note that the health dataset presents a significant class imbalance, with 2,734 healthy images and 6,659 unhealthy images, which may influence the behavior of the models and favor predictions toward the majority class. Fig. 4 shows the confusion matrix obtained using the MobileNetV2 model for plant health status detection.

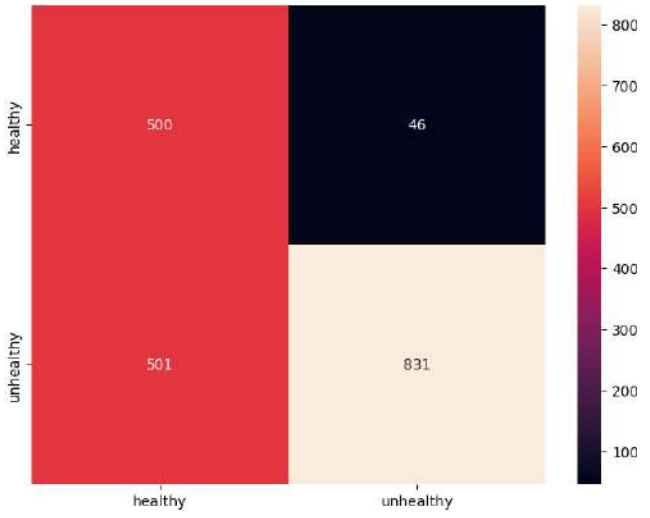


Fig. 4. Confusion matrix for plant health detection using MobileNetV2.

D. Discussion

The obtained results show that transfer learning improves performance in species classification compared to a CNN trained from scratch, which is consistent with recent studies demonstrating that pretrained models enhance accuracy by leveraging representations learned from large datasets [14]. Among the evaluated models, EfficientNetB0 achieved the best performance, aligning with research that highlights its ability to capture complex visual features with high computational efficiency [15].

In the case of plant health status detection, the results indicate that this task is more complex due to the visual variability of symptoms on leaves and differences in lighting, background, and image orientation. These variations have been identified as one of the main challenges in automatic plant disease diagnosis systems based on deep learning, since environmental factors such as lighting changes, background complexity, and seasonal variations can significantly affect the accuracy and robustness of classification models [16].

Additionally, the class imbalance present in the health dataset may affect model behavior, favoring predictions toward the majority class. This issue has been widely reported in recent studies on artificial intelligence applications in agriculture, where datasets often present unequal class distributions, which can influence evaluation metrics and the model's ability to correctly detect minority classes [17]. In this study, no specific balancing techniques such as class weighting or resampling were applied, which may limit generalization across classes. Although data augmentation was used to improve generalization, it does not address the underlying imbalance. Future work could incorporate strategies such as class weighting or oversampling to improve performance in imbalanced scenarios.

Overall, these results suggest that the proposed system shows promising performance for the automatic identification of plant species, while the plant health detection module could benefit from more balanced datasets or additional training techniques.

IV. CONCLUSION

This work demonstrates the feasibility of using deep learning techniques for the automatic identification of medicinal plants from leaf images. The comparison of different neural network architectures showed that transfer learning–based approaches outperform a CNN trained from scratch, with EfficientNetB0 achieving the best results in the species classification task.

The integration of the trained models into a mobile application highlights the practical potential of these techniques for real-world use, bringing plant identification tools closer to students, farmers, and general users.

For plant health status detection, the results indicate greater challenges due to symptom variability and class imbalance, and this module is therefore considered a preliminary component of the system. Its performance could be improved through more balanced datasets and specialized training techniques, such as class weighting or resampling.

Future work includes expanding the dataset with more species and samples, as well as exploring specific disease classification approaches. Additionally, the current client-server architecture based on a local Flask server may limit real-world deployment, and future implementations could consider on-device inference or cloud-based solutions to improve accessibility in field conditions.

REFERENCES

[1] C. C. Davis and P. Choisy, "Sciencedirect," 26 February 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960982223017384. [Accessed 10 March 2026].

[2] D. Chetia, S. K. Kalita, P. P. Pratim Baruah, D. Dutta and T. Akhter, "Arxiv," 16 June 2025. [Online]. Available: https://arxiv.org/abs/2501.09363. [Accessed 10 March 2026].

[3] A. Upadhyay, N. Singh Chandel, K. Pratap Singh, S. Kumar Chakraborty, B. M. Nandede, M. Kumar, A. Subeesh, K. Upendar, A. Salem and A. E. Elbeltagi, "Springer Nature Link," 17 January 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-11100-x. [Accessed 10 March 2026].

[4] Y. Kaya and E. Gürsoy, "Pubmed," 9 September 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/41246232/. [Accessed 10 March 2026].

[5] U. Mishra, A. Pandey, L. G and T. K, "Scientific Reports," 24 December 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-32607-5. [Accessed 10 March 2026].

[6] A. Jannat, I. Jahan Preety and M. Uddin Mojumdar, "Dataset of Medicinal Plant Leaves Hibiscus, Aloe Vera, Moringa, Tulsi, and Neem," 21 November 2024. [Online]. Available: https://data.mendeley.com/datasets/d297rvr3r6/1. [Accessed 2 3 2026].

[7] A. Shah, "Dataset For Classifying Indian Medicinal Plants And Leafs," 2023. [Online]. Available: https://www.kaggle.com/datasets/aryashah2k/indian-medicinal-leaves-dataset/data. [Accessed 2 March 2026].

[8] M. M. Billah, S. Sagor, M. A. Rahman and M. Shorif Uddin, "Hibiscus Leaf Diseases Classification Dataset," 26 January 2026. [Online]. Available: https://data.mendeley.com/datasets/k8pp8cbp8k/3. [Accessed 2 March 2026].

[9] M. A. Hasnat Zilan, N. Biswas and M. Uddin Mojumdar, "Moringa Tree Leaves Dataset for Botanical and Agricultural Analysis," 25 11 2024. [Online]. Available: https://data.mendeley.com/datasets/zc4vdgph9b/1. [Accessed 2 March 2026].

[10] M. Manzurul Islam, R. Islam, S. Jaman Omi, M. Billah Shefa, M. R. Hasan, A. Das, N. Tasnim Niloy, S. Ali, M. H. Ferdaus and F. Mridha, "Image dataset of Neem (Azadirachta Indica) leaf disease," 2 12 2024. [Online]. Available: https://data.mendeley.com/datasets/sjtxmcv5d4/3. [Accessed 2 March 2026].

[11] A. Moin Khan, M. Moinur Reza and M. Uddin Mojumdar, "Aloe Vera Leaf Disease Detection Dataset," 13 November 2024. [Online]. Available: https://data.mendeley.com/datasets/7w6t4zx33n/1. [Accessed 2 March 2026]

[12] s. d. nirmal dhani and T. Sarkar, "Mendeley Data," 1 July 2024. [Online]. Available: https://data.mendeley.com/datasets/gwf2rtc9z2/1. [Accessed 9 March 2026].

[13] L. Nuñez, "Kaggle," 5 March 2026. [Online]. Available: https://www.kaggle.com/datasets/lauranun/medicinal-plant-leaves-species-and-health-dataset. [Accessed 9 March 2026].

[14] M. I. Hossen, M. Awrangjeb, S. Pan and A. Al Mamun, "Springer Nature Link," 25 January 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-11081-x. [Accessed 10 March 2026].

[15] [15] G. Rivera, R. Porras, R. Florencia and J. P. Sánchez-Solís, "Sciencedirect," April 2023. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0168169923001254. [Accessed 10 March 2026].

[16] M. Shoaib, A. Sadeghi-Niaraki, F. Ali, I. Hussain and S. Khalid, "PubMed Central," 21 February 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11885274/. [Accessed 10 March 2026].

[17] J. Richetti, F. I. Diakogianis, A. Bender, A. F. Colaço and R. A. Lawes, "Sciencedirect," February 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169923000303. [Accessed 10 March 2026].

AUTHORS



Maria Laura Nuñez Jaillita is a Systems Engineering student at Universidad Privada del Valle (UNIVALLE). CTO of DocIA, a project focused on integrating artificial intelligence models into software systems for digital health applications. Member of the IEEE Bolivia Section, participating in conferences and academic activities related to technology and engineering.



Juan Abiel Iriarte Colque is a Systems Engineering student at Universidad Privada del Valle (UNIVALLE), Bolivia. His research interests include artificial intelligence, machine learning, computer vision, and mobile application development. His current work focuses on the application of deep learning techniques for medicinal plant identification and plant health detection using image analysis.



Eynar Calle Viles holds a Master's degree in Telecommunications, with a specialization in Communication Networks and Telematics, and a postgraduate degree in Internet of Things (IoT) from the National Institute of Telecommunications (INATEL), Brazil.

With more than 10 years of professional experience, he is a specialist in the design and implementation of networks, FTTH, IoT, and embedded systems. He has developed several technological innovations, including patent proposals such as the "Mobile Robot for Metal Detection" and the "Localization System for the Tupac Katari Satellite."

In addition, he has multiple publications and participations in international journals and conferences. His main research interests include antennas and propagation, information security, and IoT, actively contributing to the development of innovative technological solutions.

He currently serves as Research Coordinator at the Faculty of Informatics and Electronics at Universidad Privada del Valle (Univalle) and is President of the IEEE Computer Society Bolivia.

Desenvolvimento de um protótipo para detecção e avaliação de risco de queda em ambiente hospitalar

Fernando Puebla Stein, Matheus Vieira Honório de Souza, Rita Elizabeth Santos de Almeida
Instituto Nacional de Telecomunicações - Inatel
fernando.puebla@ges.inatel.br, matheus_vieira@ges.inatel.br, r.elizabeth@inatel.br

***Abstract*—This work presents the development of a wearable prototype for fall-risk monitoring based on the ESP32-C3 SuperMini microcontroller and the MPU6050 inertial sensor. The embedded algorithm operates as a finite state machine that detects impacts and abrupt rotations through thresholds, evaluates user posture, and confirms fall events via temporal persistence, reducing false positives. Processed data are transmitted to a remote server through a REST API. Preliminary tests with a volunteer in a controlled environment showed that the system correctly classifies motion states and generates reliable alerts in fall-compatible scenarios. The proposed solution provides a low-cost and replicable approach for prevention and early warning in hospital settings.**

***Index Terms*—Embedded systems; fall detection; inertial sensors; wearable devices.**

***Resumo*—Este trabalho apresenta o desenvolvimento de um protótipo vestível para monitoramento de risco de queda, baseado no microcontrolador ESP32-C3 SuperMini e no sensor inercial MPU6050. O algoritmo embarcado opera como uma máquina de estados finitos que detecta impactos e rotações bruscas por limiares, verifica a postura do usuário e confirma eventos de queda por persistência temporal, reduzindo falsos positivos. Os dados processados são transmitidos a um servidor remoto via API REST. Testes preliminares com voluntário em ambiente controlado demonstraram que o sistema classifica corretamente os estados de movimentação e gera alertas confiáveis em situações compatíveis com queda. A solução proposta constitui uma abordagem de baixo custo e replicável para prevenção e alerta precoce em ambientes hospitalares.**

***Palavras chave*—Detecção de quedas; dispositivos vestíveis; sensores inerciais; sistemas embarcados.**

I. Introdução

De acordo com a Organização Mundial da Saúde (OMS), as quedas são consideradas um problema de saúde global, sendo responsáveis pela segunda principal causa de óbito por lesões acidentais, estima-se que anualmente ocorram cerca de 684.000 mortes em todo o mundo, ficando atrás somente dos acidentes de trânsito. Cerca de 80% das mortes causadas por quedas ocorrem em países de baixa e média renda, sendo as principais vítimas fatais, adultos com 60 anos ou mais. Além disso, estima-se que ocorram cerca de 37,3 milhões de quedas suficientemente graves para demandar atenção médica a cada ano. De acordo com a OMS diversos fatores podem influenciar a gravidade de uma queda, como por exemplo sexo, idade e ocupações desempenhadas pelo indivíduo. Dentre os motivos pontuados pela organização, pode-se destacar: condições médicas, como doenças neurológicas, cardíacas ou outras condições incapacitantes, bem como, efeitos colaterais de medicamentos e perda de equilíbrio, principalmente entre os idosos, fatores comumentes encontrados em hospitais e clínicas médicas.[1]

No Brasil, a incidência desses acidentes em ambiente hospitalar varia de 1,4 a 13 quedas por 1.000 pacientes, dependendo das práticas de prevenção adotadas por cada instituição. Nesse contexto, a queda em ambiente hospitalar configura-se como um evento adverso de alta relevância, não apenas pelo alto risco de morbidade, mas também pelo potencial de agravar o estado clínico do paciente. [2]

Alguns hospitais já dispõem de algumas tecnologias em ambientes hospitalares com o intuito de prevenir e alertar possíveis quedas, entre as quais destacam-se: camas inteligentes com alarmes de saída do leito, detecção de postura e painéis de *status* que auxiliam na segurança do paciente no leito e na intervenção de quedas e acidentes [3]; sensores de pressão para leito e tapetes de piso com alarme integráveis ao sistema de chamada de enfermagem, para identificar transferências entre leitos não assistidas e reduzir o tempo de resposta da equipe médica[4]; plataformas de observação remota com vídeo e apoio de Inteligência Artificial (IA), capazes de monitorar múltiplos pacientes e emitir alertas proativos [5]; e sensores *touchless* baseados em radar, que detectam presença, postura e quedas sem imagem, úteis em quartos e banheiros [6].

Apesar do potencial, a adoção de tais tecnologias enfrenta barreiras, como: resultados inconsistentes ao se utilizar essas tecnologias separadamente, risco de fadiga de alarmes por parte da equipe médica, custos de aquisição e de infraestrutura, falta de privacidade quando utilizado sistemas de vídeos e dificuldade de envio das informações para o sistema de chamada de enfermagem. [7].

Diante desses desafios mencionados, foi proposto um dispositivo *wearable* composto por um *button* capaz de monitorar a deambulação do paciente que se encontra internado, permitindo a detecção precoce de instabilidade da marcha, antecipando a queda. Optou-se pela não utilização de vídeo, a fim de baratear a solução e permitir maior privacidade do paciente, além de ampliar a cobertura do monitoramento.

II. Metodologia

A metodologia adotada para o desenvolvimento do projeto foi dividida em quatro etapas principais: seleção de componentes, desenvolvimento do *firmware*, fabricação do *hardware* e prototipagem mecânica.

A. Materiais e Componentes Eletrônicos

Para a construção dos protótipos, priorizou-se a utilização de componentes que garantissem uma maior autonomia energética e conectividade eficiente para o envio de informações, de forma rápida e *wirelles*. Os principais componentes utilizados foram:

- **Unidade de Processamento:** Microcontrolador ESP32-C3 SuperMini. A escolha deste módulo justificou-se pelo seu tamanho compacto, o que favorece a sua utilização em um dispositivo *wearable*, além de suporte a protocolos de comunicação recentes, como Wi-Fi 6 e Bluetooth 5 LE.
- **Sensores Inerciais:** Módulo MPU6050. Este componente integra um acelerômetro e um giroscópio de 3 eixos, sendo essencial para a captura dos dados de aceleração linear e velocidade angular utilizados no algoritmo de detecção de quedas.
- **Sistema de Alimentação:** Para garantir a autonomia durante o monitoramento contínuo, utilizou-se uma bateria de 2600 mAh. O gerenciamento de energia é realizado através de um módulo de carregador dedicado, acoplado a um suporte para bateria que facilita a manutenção.
- **Regulador de Tensão:** Módulo Conversor *Step-Up* MT3608. Este componente foi aplicado para elevar a tensão nominal da bateria para os 5V constantes exigidos pelo microcontrolador, assegurando a estabilidade do sistema mesmo com a descarga progressiva da célula de energia.

B. Desenvolvimento de Firmware e Software

Para o desenvolvimento do projeto, foi utilizado o microcontrolador ESP32-C3 SuperMini e para o desenvolvimento do algoritmo, foi selecionado o ambiente de desenvolvimento, IDE Arduino. A escolha desta plataforma deve-se à sua robustez e à disponibilidade de bibliotecas necessárias para o funcionamento dos sensores utilizados, como a *Adafruit MPU6050* que permitiu a integração com o sensor incercial.

b.1) Configuração do Sensor e Algoritmo

No *firmware* desenvolvido, o sensor MPU6050 foi configurado com uma escala de fundo de 8g para o acelerômetro e $\pm 500°/s$ para o giroscópio, esses valores foram selecionados para garantir que mudanças bruscas de aceleração e velocidades angulares durante a queda não alterem o sensor, mantendo a integridade dos dados. Para amenizar ruídos mecânicos e vibrações irrelevantes, aplicou-se um filtro digital passa-baixa com largura de banda de 21 Hz, frequência que isola movimentos biomecânicos humanos de ruídos espúrios de alta frequência. O sistema inicia com uma rotina de calibração automática, coletando 50 amostras em repouso para definir as referências (*offsets*) de cada eixo do sensor.

Para a detecção de eventos, o algoritmo realiza a leitura e calcula continuamente a magnitude do vetor de aceleração (A_{total}) e a magnitude da velocidade angular (G_{total}), combinando os dados dos três eixos ortogonais (x, y, z). Estas grandezas são obtidas através das Equações 1 e 2:

$$A_{total} = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{1}$$

Sendo: A_x, A_y e A_z representam as componentes da aceleração medidas nos eixos x, y e z, respectivamente, e A_{total} é a magnitude do vetor aceleração.

$$G_{total} = \sqrt{G_x^2 + G_y^2 + G_z^2} \tag{2}$$

Sendo: G_x, G_y e G_z representam as componentes da velocidade angular medidas nos eixos x, y e z, respectivamente, e G_{total} é a magnitude do vetor de rotação.

O algoritmo de detecção de queda opera como uma máquina de estados finitos baseada em três critérios sequenciais:

1) **Impacto ou Rotação Brusca:** O sistema monitora se A_{total} ultrapassa o limiar de impacto ($12m/s^2$) ou se G_{total} excede o limiar de rotação ($150°/s$), caracterizando uma rotação brusca.
2) **Verificação de Postura:** Após detectar um movimento súbito, o algoritmo analisa o componente A_z do acelerômetro para determinar a orientação do usuário. Valores inferiores a $7m/s^2$ neste eixo indicam uma possível posição deitada.
3) **Confirmação Temporal:** Para evitar falsos positivos é requerido que a "posição deitada"persista por uma janela de 2000 ms (2 segundos). Somente após este período o sistema confirma a queda e envia o alerta via API REST utilizando a conectividade Wi-Fi do ESP32-C3.

b.2) Comunicação via API REST

Após a inicialização do sistema e a conexão à rede Wi-Fi, o ESP32-C3 SuperMini realiza a sincronização de horário via *Network Time Protocol* (*NTP*), permitindo o registro de tempo para cada amostra enviada. Em operação contínua, a cada ciclo de aquisição, o dispositivo coleta as leituras do MPU6050, processa esses dados localmente e determina um *status* da posição do usuário por meio de uma máquina de estados, conforme ilustrado no fluxograma da Figura 1, que informa como saída: `PARADO`, `EM MOVIMENTO`, `ANALISANDO MOVIMENTO` e `QUEDA CONFIRMADA`.



Fig. 1. Fluxograma do *software*.

A comunicação com o servidor é realizada por meio de requisições *Hypertext Transfer Protocol* (HTTP) do tipo *POST*, com corpo em formato JSON e cabeçalho `Content-Type: application/json`. O *firmware* implementa dois *endpoints* REST distintos: um *endpoint* de *checkpoint*, utilizado para registrar a sequência temporal das medições e o instante correspondente, e um *endpoint* para medir os dados do sensor, que recebe os valores de acelerômetro e giroscópio, além de métricas derivadas como, por exemplo, a magnitude de variação da aceleração e da rotação, o *status* estimado e o *checkpoint*. Dessa forma, o servidor pode armazenar tanto a trilha temporal do monitoramento quanto os dados brutos associados a cada avaliação do algoritmo.

O envio ocorre apenas quando o dispositivo está efetivamente conectado ao Wi-Fi, caso contrário, as requisições não são disparadas, evitando tentativas de comunicação sem conectividade. No lado do servidor, a API pode validar e guardar os dados recebidos e, quando o *status* indicar condição crítica como, por exemplo, `QUEDA CONFIRMADA`, acionar mecanismos de notificação aos responsáveis, por meio da central de enfermagem ou um aplicativo, reduzindo o tempo de resposta em situações de emergência.

C. Fabricação do Hardware

A materialização do circuito eletrônico foi realizada internamente no FabLab do Inatel.

O desenvolvimento do *layout* da placa de circuito impresso (PCI) foi realizado utilizando o *software* Autodesk Fusion 360, que integra ferramentas de modelagem 3D e *design* eletrônico em uma única plataforma, facilitando a verificação dimensional e a compatibilidade mecânica do projeto. Durante a etapa de elaboração do *layout*, buscou-se manter dimensões reduzidas, visando sua possível aplicação em um dispositivo *wearable* do tipo *button*, ao mesmo tempo em que se garantia espaço suficiente para acomodação de todos os componentes eletrônicos. A PCI foi desensolvida com as dimensões de 64,75 mm × 46,66 mm, após ajustes pontuais no *layout*. O processo de desenvolvimento da PCI no Fusion pode ser observado na Figura 2. A largura das trilhas foi definida em 0,7 mm, a partir de recomendações do FabLab e também para suportar as correntes elétricas previstas no circuito, sem correr o risco de ocorrer o rompimento das trilhas.

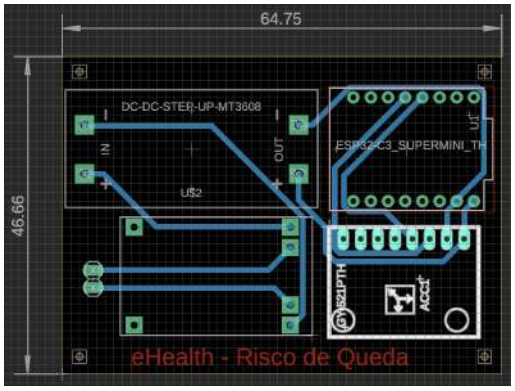


Fig. 2. Layout da PCI com dimensões e roteamento.

A PCI contém os seguintes componentes principais: um microcontrolador ESP32-C3 SuperMini, responsável pelo processamento e comunicação sem fio, um módulo conversor *Dc step-up* MT3608 para regulação da tensão de alimentação a partir da bateria e o módulo sensor GY-521, baseado no circuito integrado MPU-6050, que combina acelerômetro e giroscópio de três eixos para ler os dados de movimento. A integração desses componentes pode ser visualizada no diagrama em blocos da Figura 3.

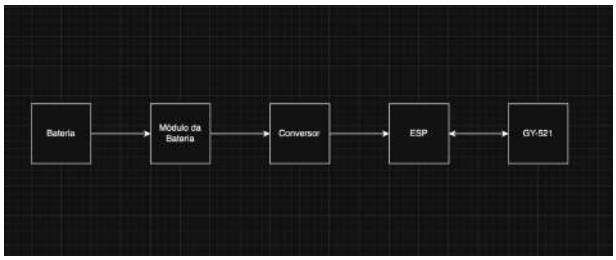


Fig. 3. Diagrama em blocos do hardware.

O processo de fabricação envolveu a validação dos arquivos Gerber gerados pelo Fusion 360 e ajustes na disposição dos furos da placa para compatibilidade com as máquinas de prototipagem do FabLab.

III. Resultados e Discussão

A. Visão geral do dispositivo desenvolvido

O dispositivo final foi composto por três blocos principais: sensoriamento inercial por parte do MPU, processamento embarcado e comunicação sem fio pelo ESP32-C3. O módulo MPU6050 realiza a aquisição contínua de aceleração e velocidade angular em três eixos, cujos sinais são filtrados e processados localmente pelo ESP32-C3 SuperMini. Foi implementado um algoritmo de máquina de estados que identifica eventos de impacto/rotação e confirma a condição de queda por meio da análise de postura e persistência das informações por 2 segundos, reduzindo falsos positivos.

Do ponto de vista energético, o circuito foi alimentado por bateria recarregável de 3.6 Volts com gerenciamento dedicado, garantindo autonomia para monitoramento contínuo, por no máximo 3 dias contínuos. Quando conectado à rede Wi-Fi, o dispositivo envia *checkpoints* para um servidor local via API REST, permitindo o registro histórico das medições e o acionamento de alertas em caso de `QUEDA CONFIRMADA`. Essa arquitetura integrou um *hardware* compacto, processamento local e comunicação remota, possibilitanto futuramente a utilização do projeto em um ambiente real de acompanhamento de pacientes.

B. Avaliação do Hardware Prototipado

A PCI fabricada no FabLab atendeu aos requisitos dimensionais propostos, resultando em um dispositivo compacto e ergonomicamente viável para uma possível utilização como vestível, como pode ser observado na Figura 4. A montagem dos componentes, incluindo o módulo MT3608 e o GY-521, ocorreu conforme o planejado no *design* 3D.

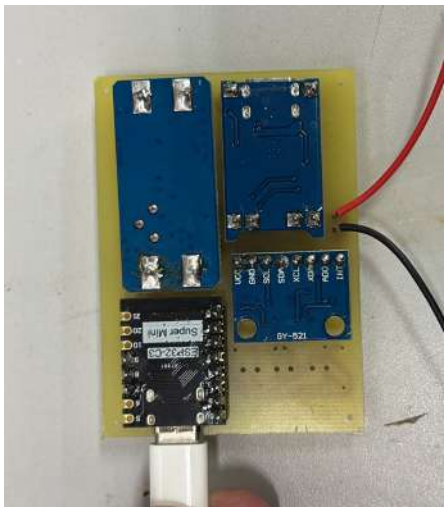


Fig. 4. Protótipo da PCI com módulos integrados.

Contudo, durante os testes de integração, observou-se uma limitação operacional severa referente à conectividade. Embora o microcontrolador ESP32-C3 SuperMini funcione normalmente de forma isolada, estabelecendo conexão com a rede Wi-Fi e comunicando-se sem instabilidades, percebeu-se que, uma vez soldado à placa e alimentado pelo circuito desenvolvido, o dispositivo apresentou falhas constantes ao tentar conectar-se à rede. A principal hipótese para essa instabilidade está na capacidade de fornecimento de energia da PCI. Percebeu-se que o módulo conversor MT3608, em conjunto com o dimensionamento das trilhas, não conseguiu fornecer a corrente necessária para suprir os picos de consumo de até 500 mA exigidos pelo ESP32 durante a conexão, ocasionando quedas de tensão que impedem o funcionamento adequado do microcontrolador ao se comunicar com o Wi-Fi.

C. Validação do Algoritmo e Comunicação

Diante da limitação de conectividade encontrada na PCI final, e com o objetivo de isolar as variáveis de teste para validar o algoritmo desenvolvido e a comunicação com o servidor, optou-se por realizar os testes práticos utilizando uma versão antiga do protótipo. Este protótipo anterior possui equivalência funcional completa em termos de componentes (ESP32-C3, MPU6050 e alimentação), sendo uma versão anterior à PCI desenvolvida, estando montado em uma *protoboard*.

Essa abordagem permitiu validar a máquina de estados implementada no *firmware*. Para isso, foram realizados testes práticos com um voluntário de 1,66 m de altura em um ambiente fechado. O objetivo foi avaliar o comportamento do sistema em duas situações principais: movimentação normal e ocorrência de quedas.

Primeiro, o voluntário simulou ações rotineiras, como caminhar pelo ambiente. Durante essa etapa, foi verificado que o dispositivo conseguiu processar os dados corretamente, mantendo o status classificado como ”EM MOVIMENTO”, exatamente como o esperado, sem disparar alarmes falsos.

Na sequência, foram simulados tombos de diferentes maneiras, como quedas frontais e laterais, sobre uma superfície segura. Nesses testes, o *firmware* conseguiu identificar a mudança brusca de movimento e o impacto, alterando o estado para indicar a queda em todas as tentativas. Esse resultado prático confirmou que a lógica desenvolvida está funcionando como esperado.

D. API e Registro de Dados

Para a etapa de integração, a API REST funcionou como a ponte de comunicação entre o dispositivo e o servidor. Para validar essa camada durante os testes, foi criado um servidor HTTP localmente. Tendo como objetivo verificar se a estrutura seria capaz de receber e armazenar os dados lidos e enviados de forma *wireless*, como o *status* do algoritmo e as leituras do acelerômetro e giroscópio, conforme detalhado na Figura 5.

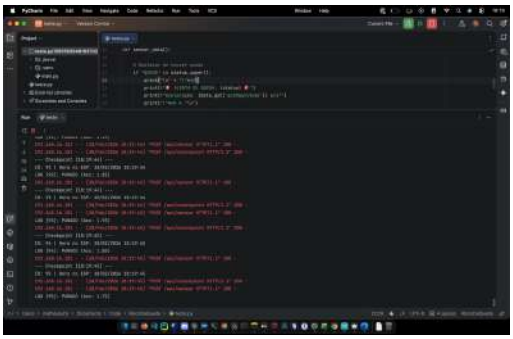


Fig. 5. Funcionamento da API.

Uma escolha importante no desenvolvimento foi a utilização do ESP para a realização de todo o processamento dos dados. Dessa forma, o microcontrolador ficou responsável por executar toda a lógica da máquina de estados e o cálculo das métricas localmente, enviando para o servidor apenas as informações consolidadas, evitando sobrecarga de tráfego desnecessário.

Durante os testes práticos, a comunicação ocorreu via requisições *POST*, estruturadas com o cabeçalho `Content-Type: application/json`. O servidor local recebeu os pacotes com os valores esperados. A API se comportou de forma estável, o que confirmou que a montagem do JSON no *firmware* estava correta e garantiu um histórico confiável para analisar o momento exato de cada evento detectado.

IV. Conclusão

O presente trabalho descreveu o desenvolvimento de um dispositivo vestível voltado ao monitoramento de risco de queda em ambiente hospitalar, concebido no contexto das atividades do eHealth Innovation Center do Inatel. A proposta buscou oferecer uma alternativa acessível e prática frente às soluções existentes, que frequentemente esbarram em custos elevados, questões de privacidade ou limitações de cobertura.

O protótipo construído, baseado no ESP32-C3 SuperMini e no sensor MPU6050, mostrou-se capaz de identificar padrões de movimento associados a quedas por meio de um algoritmo embarcado que combina detecção por limiares, verificação de postura e confirmação temporal. Os testes realizados com voluntário indicaram que a lógica implementada distingue adequadamente situações de movimentação normal e eventos de queda, sem a ocorrência de falsos alarmes durante as avaliações conduzidas.

Entretanto, a validação também evidenciou limitações relevantes. A PCI fabricada apresentou instabilidades na alimentação do microcontrolador durante a conexão Wi-Fi, o

que impediu sua utilização nos testes finais e demandou o uso de um protótipo anterior montado em protoboard. Esse ponto reforça a necessidade de revisão do dimensionamento do circuito de potência, em especial do módulo conversor DC-DC, para garantir o fornecimento adequado de corrente nos picos de consumo do ESP32-C3.

Como melhorias futuras, a correção do circuito de alimentação da PCI é considerada a de maior urgência, visto que foi justamente esse o gargalo que impediu o uso da placa nos testes finais. Além disso, é necessária a fabricação do *button*, a fim de prendê-lo no tórax do usuário e utilizá-lo como um *wearable*, bem como repetir os experimentos com mais voluntários e em condições que se aproximem da realidade de uma enfermaria, onde o paciente se movimenta entre leito, banheiro e corredor ao longo do dia.

Outra ação importante é conectar o dispositivo aos sistemas de chamada de enfermagem, já utilizados pelos hospitais, de modo que o alerta de queda chegue diretamente à equipe de plantão. Por fim, a substituição da lógica baseada em limiares fixos por técnicas de aprendizado de máquina pode trazer ganhos na precisão da detecção, embora isso demande a coleta de um volume de dados significativamente maior.

Ainda que o protótipo atual tenha suas limitações, os resultados obtidos mostram que a abordagem é viável e pode, com os devidos aprimoramentos, ser aplicada no acompanhamento real de pacientes internados.

Agradecimentos

Gostaríamos de registrar nosso apreço ao Instituto Nacional de Telecomunicações (Inatel), pela disponibilização dos recursos e do ambiente acadêmico essenciais à nossa pesquisa e formação. Igualmente, agradecemos à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo imprescindível fomento por meio de seu programa de bolsas. Deixamos também um agradecimento especial à nossa orientadora, Rita Elizabeth Santos de Almeida, por sua valiosa supervisão e constante encorajamento em todas as etapas de desenvolvimento do trabalho.

Referências

[1] World Health Organization. *Falls*. Fact sheet updated 26 April 2021. Accessed: 2025-09-15. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls.

[2] Donna Guillaume, Sybil Crawford e Patricia Quigley. "Characteristics of the middle-age adult inpatient fall". Em: *Applied Nursing Research* 31 (2016), pp. 65–71. ISSN: 0897-1897. DOI: https://doi.org/10.1016/j.apnr.2016.01.003. URL: https://www.sciencedirect.com/science/article/pii/S0897189716000197.

[3] Stryker. *ProCuity Bed Series*. https://www.stryker.com/br/en/acute-care/products/procuity/index.html. Accessed: 2025-10-22. 2025.

[4] TIDI Products. *Posey Bed and Chair Sensors*. https://www.tidiproducts.com/falls-prevention/posey-falls-prevention-alarms-and-sensors/wired-alarms-and-sensors/wired-sensors/posey-bed-chair-sensors. Accessed: 2025-10-22. 2025.

[5] AvaSure. *TeleSitter®*. https://avasure.com/telesitter/. Accessed: 2025-10-22. 2025.

[6] Xandar Kardian. *XK-300 Continuous Vital Sign and Presence Monitoring Sensor*. https://xkcorp.com/products/health-care/xk-300/. Accessed: 2025-10-22. 2025.

[7] Rauland. *Responder 5 Nurse Call System*. https://www.rauland.com/healthcare/responder-5. Accessed: 2025-10-22. 2025.

Autores



Fernando Puebla Stein Graduando em Engenharia de Software no Instituto Nacional de Telecomunicações (Inatel). Atua como pesquisador de Iniciação Científica no eHealth Innovation Center (Inatel), aplicando técnicas de Machine Learning, análise de dados e programação em projetos voltados à área da saúde. Adicionalmente, possui experiência com docência voluntária em tecnologia e programação.



Matheus Vieira Honório de Souza Graduando em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel), onde foi voluntário no laboratório eHealth Innovation Center. Atualmente exerce a função de Consultor Databricks na Datasolutec, atuando em projetos na área de *Professional Services*.



Rita Elizabeth Santos de Almeida Possui formação técnica em eletrônica, com ênfase em equipamentos médicos, na Escola Técnica de Eletrônica Francisco Moreira da Costa - ETE "FMC"(2018) e graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações - Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center, centro de pesquisa de Engenharia Biomédica do Inatel.

Estudo de técnicas de inteligência artificial para aprimoramento de óculos assistivos para pessoas com deficiência visual

Gustavo Almeida Barrios, Lavínia Silva Andrade, Rita Elizabeth Santos de Almeida,
Roberto Luiz Assad Pinheiro, Filipe Bueno Vilela
Instituto Nacional de Telecomunicações - Inatel
gustavo.barrios@gec.inatel.br, lavinia.andrade@ges.inatel.br, r.elizabeth@inatel.br, roberto@amikosolucoes.com, filipe.bueno@inatel.br

***Summary*—This document presents all the research conducted on the development of the VIDI 2.0 project, a pair of glasses whose main objective is to assist people with visual impairments by describing places, colors, objects, and distances.**

***Index Terms*—Diversity, Equity, Inclusion, Accessibility.**

***Resumo*—Este documento apresenta toda a pesquisa realizada sobre o desenvolvimento do projeto VIDI 2.0, um par de óculos cujo objetivo principal é auxiliar pessoas com deficiência visual, realizando descrições de lugares, cores, objetos e distâncias.**

***Palavras chave*—Diversidade, Equidade, Inclusão, Acessibilidade.**

I. Introdução

Thomas Hobbes aborda, em sua filosofia política, que o dever principal a ser alcançado por meio da instituição de um Estado forte é garantir, acima de tudo, o bem-estar da população. Para o autor, a formação do Estado resulta da necessidade de assegurar proteção, ordem e condições mínimas para a convivência social, sendo responsabilidade do poder soberano promover segurança e estabilidade como fundamentos da vida coletiva. Isso inclui dispor de meios suficientes para compreender as necessidades humanas em seus diferentes aspectos: físico, moral, espiritual, psicológico e social [1]. Contudo, a sociedade ainda enfrenta problemas relacionados à falta de inclusão social, especialmente devido à falha na acessibilidade às pessoas que possuem algum tipo de deficiência ou limitação [2], como divulgado pelo IBGE, segundo o qual pessoas com deficiência apresentam menor acesso à educação, ao trabalho e à renda. Segundo ele, apenas 25,6% das pessoas com deficiência tinham concluído pelo menos o Ensino Médio, enquanto 57,3% das pessoas sem deficiência tinham esse nível de instrução [3]. Ao direcionar a atenção especificamente para a Deficiência Visual (DV), o cenário não se mostra diferente. Desde a infância, indivíduos com DV enfrentam desafios tanto para se integrarem ao meio social, o que muitas vezes os leva à dependência de terceiros para a realização de atividades básicas, quanto no desempenho educacional, em razão de falhas didáticas e da ausência de recursos adequados [4].

Muito além da educação, a população com deficiência encara barreiras que limitam sua participação no mercado de trabalho e reduzem seu potencial de geração de renda, o que perpetua um ciclo de dependência e marginalização. As estatísticas indicam que as taxas de ocupação para pessoas com deficiência são dramaticamente inferiores às do restante da população. Essa exclusão econômica é um reflexo direto da falta de investimento em acessibilidade e da falha na inclusão para considerar as necessidades específicas desses indivíduos [5].

Conforme destacado na Revista Educação Especial, "o desenvolvimento da pessoa com deficiência visual apresenta particularidades que precisam ser conhecidas e consideradas. Pesquisas sugerem que, ao chegar à adolescência, a criança com deficiência visual poderá apresentar desenvolvimento normal, desde que não seja privada de experiências sensoriais, motoras, cognitivas e sociais significativas."[6]. Ainda assim, é notória a dificuldade enfrentada diariamente, o que motivou o desenvolvimento do presente projeto.

Soluções tecnológicas vêm sendo desenvolvidas ao longo dos anos com o objetivo de auxiliar esse público. Um exemplo amplamente conhecido é o *OrCam MyEye*, dispositivo que, segundo a empresa responsável, consiste em câmera inteligente capaz de capturar imagens do ambiente e converter informações visuais em áudio em tempo real. O equipamento realizava tarefas como descrição de objetos, leitura de textos e identificação de produtos, por meio de *feedback* sonoro, descrevendo ao usuário elementos que poderiam ser percebidos visualmente [7]. Entretanto, segundo a Organização Nacional de Cegos do Brasil (ONCB), a *OrCam Technologies* anunciou o fim da fabricação dos óculos assistivos *OrCam MyEye* [8]. Além disso, o valor unitário do dispositivo, aproximadamente R$ 14.900,00, torna-o inacessível para grande parte da população, especialmente considerando o contexto socioeconômico brasileiro.

O presente estudo é a continuidade da pesquisa iniciada em 2019 por Roberto Pinheiro, na qual foi proposto o desenvolvimento de um dispositivo assistivo baseado em inteligência artificial e visão computacional, denominado VIDI. A motivação central do trabalho surgiu da observação das diversas dificuldades enfrentadas por pessoas com DV em atividades cotidianas que exigem reconhecimento visual, como identificar objetos, produtos, cores ou até mesmo cédulas de dinheiro. Essas limitações reduzem a autonomia dessas pessoas e aumentam sua dependência de terceiros para a realização de tarefas simples do dia a dia [9].

O protótipo inicial era capaz de capturar imagens por meio

de uma câmera acoplada a óculos e convertê-las em descrições narradas ao usuário, utilizando técnicas de aprendizado de máquina e reconhecimento de imagens. Porém, sua principal limitação estava na necessidade de acessar a internet constantemente, limitando a utilização e restringindo seu uso. Outro ponto crítico da primeira versão eram as limitações de software da época, pois o cenário tecnológico ainda não contava com a expansão significativa da Inteligência Artificial (IA) observada atualmente, existindo diversas limitações para seu desenvolvimento, fato que inclusive foi pauta do estudo de Xie, Zhai e Lu (2025), intitulado ”Evolution of artificial intelligence in healthcare: a 30-year bibliometric study”, publicado na revista Frontiers in Medicine [10].

Neste estudo, inicialmente considerou-se a utilização de modelos de IA baseados em linguagem natural, como o ChatGPT. Esse tipo de solução, contudo, depende de conexão constante com a internet para seu funcionamento, o que representa uma limitação significativa em contextos nos quais o usuário pode perder o sinal de rede móvel ou enfrentar instabilidades de conexão. Diante desse cenário, destaca-se a necessidade de soluções assistivas que não apenas reproduzam funcionalidades já existentes, mas que avancem em aspectos técnicos fundamentais, como acessibilidade, custo e independência de infraestrutura externa. Nesse sentido, a principal contribuição deste trabalho está no desenvolvimento de um sistema embarcado capaz de realizar reconhecimento de imagens e geração de descrições em tempo real de forma totalmente *offline*, eliminando a dependência de conexão com a internet. Além disso, o projeto propõe uma arquitetura otimizada para dispositivos de baixo custo, buscando tornar a tecnologia mais acessível à população, ao mesmo tempo em que explora o uso de técnicas contemporâneas de inteligência artificial embarcada, ampliando a eficiência e a autonomia do usuário em atividades cotidianas.

A. Estrutura geral do Pipeline

O sistema proposto foi estruturado como um *pipeline* de processamento sequencial, composto por etapas de entrada de imagem, extração de características visuais, interpretação semântica e geração de resposta em linguagem natural. Inicialmente, a imagem é fornecida ao sistema por meio de arquivos digitais utilizados nos testes experimentais, sendo prevista, em etapas futuras, a integração com uma câmera acoplada ao dispositivo. Em seguida, são aplicadas técnicas de visão computacional para extração de informações relevantes, incluindo reconhecimento de texto e identificação de objetos. Essas informações são então processadas por modelos de inteligência artificial responsáveis pela interpretação contextual e geração de descrições. Por fim, o resultado é convertido em saída textual, podendo ser posteriormente adaptado para saída auditiva.

No contexto desse pipeline, foram avaliadas combinações entre diferentes ferramentas, incluindo o *EasyOCR*, para a identificação de textos presentes nas imagens; o *BLIP*, para o reconhecimento de objetos; e o *DistilGPT-2*, com o objetivo de integrar as informações extraídas em uma resposta unificada.

A Figura 1 ilustra o fluxo geral do sistema.

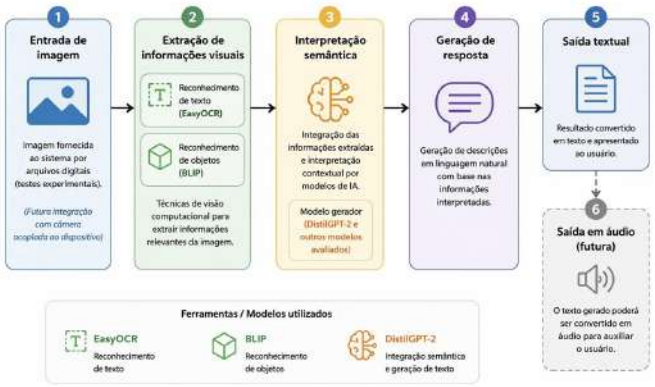


Fig. 1. Fluxo do pipeline do sistema, destacando as etapas de processamento desde a entrada da imagem até a geração da resposta textual.

B. Avaliação de Modelos Alternativos

Foram avaliados diferentes modelos com potencial de execução local, todos quantizados em baixa precisão (4 bits), visando redução do consumo de memória e melhoria do tempo de inferência.

Entre os modelos testados, destacam-se: *DistilGPT-2*, *Phi-2*, *Tiny LLaMA*, *OpenHermes 2.5* e *Nous Hermes 2*.

A seleção desses modelos considerou critérios como capacidade de geração textual, compatibilidade com execução embarcada e possibilidade de integração ao *pipeline* em *Python*.

Além dos critérios mencionados, consideraram-se aspectos arquiteturais dos modelos avaliados. O *DistilGPT-2* é uma versão reduzida do GPT-2 baseada em arquitetura *Transformer*, com menor número de parâmetros, o que favorece sua execução em ambientes com restrições computacionais. O modelo *Phi-2*, por sua vez, apresenta maior capacidade de representação semântica, porém com limitações observadas na geração de texto em língua portuguesa.

Os modelos *Tiny LLaMA*, *OpenHermes 2.5* e *Nous Hermes 2* também são baseados em arquiteturas do tipo *Transformer*, com variações no número de parâmetros e estratégias de treinamento. Enquanto modelos menores, como o *Tiny LLaMA*, apresentam melhor desempenho computacional, modelos maiores tendem a oferecer maior qualidade semântica, ao custo de maior tempo de inferência.

Dessa forma, a escolha dos modelos considerou o equilíbrio entre qualidade das respostas geradas, tempo de processamento e viabilidade de execução em ambiente embarcado.

C. Investigação de Arquiteturas para Image Captioning

As etapas seguintes consistiram na avaliação de modelos de *Image Captioning*. Nesse contexto, a arquitetura *Inception V3* foi analisada como alternativa para extração de características visuais, considerando sua ampla utilização na literatura e compatibilidade com conjuntos de dados consolidados, como os datasets do *Flickr*.

Foram conduzidos testes iniciais com modelos fundamentados nessa arquitetura, visando verificar sua adequação às restrições computacionais do sistema.

D. Implementação em Ambiente Embarcado

A implementação prática foi realizada em uma *Raspberry Pi 4 Model B*, integrada à rede local do laboratório *eHealth*.

Para viabilizar os testes, optou-se pela utilização da versão *Lite* do sistema operacional, permitindo acesso via *SSH* e maior controle dos recursos computacionais.

Para a extração de características, o modelo *Inception V3* foi truncado na camada intermediária *"mixed7"*. Esta camada fornece, originalmente, um mapa de características espaciais de $17 \times 17 \times 768$. Visando a otimização para o ambiente embarcado, este mapa foi redimensionado, via interpolação bilinear, para 7×7 e processado por uma camada convolucional 1×1 para reduzir a profundidade de 768 para 576 canais. Essa sequência de operações resultou em um tensor final de $7 \times 7 \times 576$, que foi linearizado em 49 vetores para alimentação do *Transformer*.

E. Estratégias de Otimização

Com o objetivo de reduzir o tempo de inferência e tornar o sistema viável para uso embarcado, foram aplicadas técnicas de quantização dos modelos avaliados.

Além disso, realizou-se a separação da arquitetura *Inception V3* pré-treinada, visando diminuir o tempo de processamento durante a execução.

F. Arquitetura Inception V3

Inicialmente, foi realizada fundamentação teórica por meio da análise de artigos científicos relacionados à aplicação de *image captioning* em sistemas embarcados. A análise foi baseada em estudos publicados na plataforma *Springer* [11], com destaque para trabalhos que investigam redes convolucionais otimizadas para ambientes com restrições de *hardware*.

Foram realizadas pesquisas sobre a integração do modelo *CLIP* ao pipeline de treinamento, com o objetivo de aprimorar a semântica visual das descrições geradas, considerando que o extrator *Inception V3* foi originalmente treinado com imagens mais simples do *dataset ImageNet*.

Para a otimização do treinamento, implementou-se um escalonador de taxa de aprendizado baseado nas estratégias *Linear Warm-up* e *Cosine Decay*. A fase de *warm-up* promove o aumento gradual da taxa de aprendizado a partir de valores próximos de zero, reduzindo instabilidades iniciais, como explosão de gradientes.

Na etapa subsequente, o *Cosine Decay* realiza a redução progressiva da taxa de aprendizado segundo uma função cossenoidal, permitindo ajustes mais refinados nos parâmetros à medida que o modelo se aproxima da convergência.

II. Resultados

A combinação inicial entre *EasyOCR*, *BLIP* e *DistilGPT-2* apresentou limitações na integração das informações extraídas, especialmente no processamento de *prompts* mais complexos, como aqueles que envolvem simultaneamente texto extraído via *OCR*, descrições visuais e contexto adicional. Esses cenários exigem não apenas a interpretação isolada de cada modalidade, mas também a fusão coerente de dados multimodais.

O modelo *Phi-2* apresentou desempenho insatisfatório na geração de respostas em português, indicando limitações relacionadas ao seu treinamento. Foram observados erros gramaticais, inconsistências semânticas e dificuldades na manutenção de contexto, especialmente em tarefas multimodais.

A Figura 2 ilustra um exemplo de falha semântica gerada pelo modelo *Phi-2*.



Fig. 2. Exemplo ilustrativo de falha semântica e alucinação gerada pelo modelo Phi-2.

O *Tiny LLaMA* apresentou desempenho adequado em imagens simples, porém demonstrou limitações em cenários mais complexos, indicando baixa capacidade de generalização.

Os modelos *OpenHermes 2.5* e *Nous Hermes 2* apresentaram alta qualidade semântica, porém com tempo de resposta elevado (40–50 segundos), inviabilizando seu uso em tempo real.

Os resultados indicam que a arquitetura *Inception V3* apresenta vantagens estruturais relevantes, oferecendo bom equilíbrio entre desempenho computacional e qualidade de extração de características.

A aplicação de técnicas de quantização resultou em redução de 20 a 30% no tempo de inferência. A separação do modelo pré-treinado permitiu alcançar aproximadamente 7,8 segundos por execução.

A Figura 3 apresenta o tempo de execução e a saída do terminal obtidos após a aplicação das técnicas de otimização no modelo Inception V3.

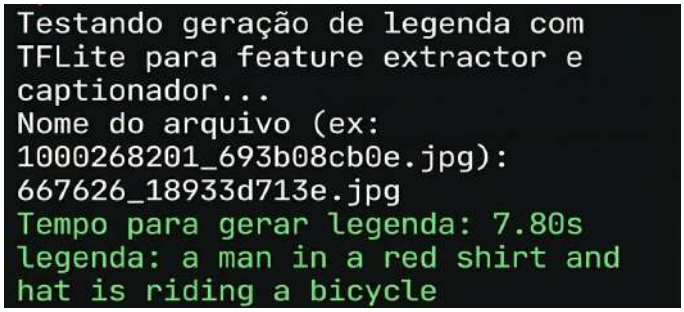


Fig. 3. Desempenho de inferência do modelo Inception V3 após técnicas de quantização e separação estrutural.

Do ponto de vista prático, a solução demonstrou viabilidade em ambiente embarcado. No entanto, aspectos como consumo de energia, latência em uso contínuo e experiência do usuário ainda demandam investigação.

Apesar dos avanços, o sistema ainda apresenta limitações, especialmente na geração de descrições semanticamente con-

sistentes em cenários complexos, evidenciando a necessidade de melhorias na integração multimodal.

Para complementar a análise qualitativa, foram consideradas métricas quantitativas, como tempo médio de inferência e consistência das respostas geradas. O tempo de inferência foi definido como o intervalo entre a entrada da imagem e a geração da resposta textual, enquanto a qualidade semântica foi avaliada com base na coerência, relevância e adequação das descrições.

Embora métricas automáticas como BLEU ou ROUGE não tenham sido aplicadas, devido à natureza exploratória do estudo, foram realizadas análises comparativas entre os modelos com base em desempenho e estabilidade.

As diferenças entre os modelos são apresentadas na Tabela I.

TABELA I
COMPARAÇÃO SINTÉTICA ENTRE MODELOS AVALIADOS

Modelo	**Qualidade**	**Tempo de Inferência**	**Observações**
EasyOCR + BLIP + DistilGPT-2	Baixa	Moderado	Limitações na fusão de dados multimodais e em prompts complexos.
Phi-2	Baixa	Moderado	Erros gramaticais e inconsistências semânticas em PT-BR.
Tiny LLaMA	Moderada	Moderado	Bom desempenho em cenários simples; limitações em imagens complexas.
OpenHermes 2.5 / Nous Hermes 2	Alta	Alto	Alta qualidade semântica, porém latência elevada para uso em tempo real.
Inception V3	**Alta**	**Baixo**	**Melhor equilíbrio entre desempenho e custo computacional.**

*Resultados considerando modelos quantizados.

III. CONCLUSÃO

O presente trabalho apresentou a evolução do sistema VIDI, propondo um pipeline baseado em inteligência artificial embarcada para auxiliar pessoas com deficiência visual na interpretação de informações visuais. Os resultados demonstraram a viabilidade da execução local, com destaque para o uso da arquitetura *Inception V3* e das técnicas de otimização, que permitiram reduzir significativamente o tempo de inferência em ambiente embarcado.

Apesar desses avanços, ainda foram identificadas limitações na qualidade semântica das legendas geradas. Mesmo após treinamentos prolongados e ajustes no modelo, as descrições apresentaram inconsistências e, em alguns casos, desconexão com o conteúdo das imagens. A incorporação do modelo *CLIP* ao pipeline não foi suficiente para superar completamente essas limitações, evidenciando desafios na integração multimodal.

Como trabalhos futuros, destacam-se a ampliação do conjunto de dados, o aumento do número de épocas de treinamento e a investigação de arquiteturas mais robustas para *image captioning*. Além disso, pretende-se evoluir o sistema para suportar entradas multimodais e integrar os componentes de captura e saída auditiva, aproximando o protótipo de uma solução prática para uso cotidiano.

Por fim, os resultados obtidos reforçam o potencial de soluções baseadas em inteligência artificial embarcada como ferramentas de apoio à acessibilidade, contribuindo para o aumento da autonomia de pessoas com deficiência visual.

REFERÊNCIAS

[1] Thomas Hobbes. *Leviatã*. Obra original publicada em 1651. São Paulo: Martins Fontes, 2003.

[2] Brasil. *Lei nº 13.146, de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)*. http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 23 fev. 2026. 2015.

[3] Instituto Brasileiro de Geografia e Estatística (IBGE). *Pessoas com deficiência têm menor acesso à educação, ao trabalho e à renda*. Acesso em: 18 fev. 2026. 2023. URL: https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37317-pessoas-com-deficiencia-tem-menor-acesso-a-educacao-ao-trabalho-e-a-renda.

[4] Organização das Nações Unidas. *Convenção sobre os Direitos das Pessoas com Deficiência*. https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html. Acesso em: 23 fev. 2026. 2006.

[5] World Health Organization e World Bank. *World Report on Disability*. Geneva: World Health Organization, 2011. URL: https://www.who.int/publications/i/item/9789241564182.

[6] Carla Beatris Valentini, Cláudia Alquati Bisol, Luani dos Santos Paim e Ana Paula Fachinetto Ehlers. "Educação e deficiência visual: uma revisão de literatura". Em: *Revista Educação Especial* 32 (2019). Acesso em: 24 fev. 2026. DOI: 10.5902/1984686X33154. URL: https://periodicos.ufsm.br/educacaoespecial/article/view/33154.

[7] OrCam Technologies. *OrCam MyEye*. https://www.orcam.com/pt-pt/orcam-myeye. Acesso em: 18 fev. 2026. 2026.

[8] Organização Nacional de Cegos do Brasil. *Empresa OrCam Technologies anuncia fim da fabricação dos "óculos falantes" OrCam MyEye*. https://pt.linkedin.com/posts/oncbempresa_orcam-technologies-anuncia-fim-da-activity-7224502784081989636-swtC. Publicação no LinkedIn. Acesso em: 24 fev. 2026. 2024.

[9] Roberto Luiz Assad Pinheiro e Filipe Bueno Vilela. "VIDI: Artificial Intelligence and Vision Device for the Visually Impaired". Em: *Proceedings of the XXVII Brazilian Congress on Biomedical Engineering (CBEB 2020)*. Ed. por Teodiano Freire Bastos-Filho, Eduardo M. de Oliveira Caldeira e Antonio Frizera-Neto. Vol. 83. IFMBE Proceedings. Cham: Springer, 2022. DOI: 10.1007/978-3-030-70601-2_200. URL: https://doi.org/10.1007/978-3-030-70601-2_200.

[10] Yaojue Xie, Yuansheng Zhai e Guihua Lu. “Evolution of Artificial Intelligence in Healthcare: A 30-Year Bibliometric Study”. Em: *Frontiers in Medicine* 11 (2025), p. 1505692. DOI: 10.3389/fmed.2024.1505692. URL: https://doi.org/10.3389/fmed.2024.1505692.

[11] Abubeker Kiliyanal Muhammed Kunju, S. Baskar, Sherin Zafar, Bushara A. R., Rinesh S. e Shafeena Karim A. “A Transformer-Based Real-Time Photo Captioning Framework for Visually Impaired People with Visual Attention”. Em: *Multimedia Tools and Applications* 83 (2024), pp. 88859–88878. DOI: 10.1007/s11042-024-18966-7. URL: https://doi.org/10.1007/s11042-024-18966-7.

AUTORES



Gustavo Almeida Barrios possui curso técnico em Informática pelo IFSULDEMINAS – Campus Pouso Alegre, concluído em 2025. Atualmente é estudante de Engenharia de Computação e bolsista do projeto VIDI 2.0 no INATEL – Instituto Nacional de Telecomunicações.



Lavínia Silva Andrade possui curso técnico em Informática pelo IFSULDEMINAS – Campus Pouso Alegre, concluído em 2025. Atualmente é estudante de Engenharia de Software, monitora de apoio e voluntária no Projeto VIDI 2.0 no INATEL – Instituto Nacional de Telecomunicações.



Rita Elizabeth Santos de Almeida possui formação técnica em Eletrônica, com ênfase em equipamentos médicos, pela Escola Técnica de Eletrônica Francisco Moreira da Costa – ETE “FMC” (2018), e graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações – Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center.



Roberto Luiz Assad Pinheiro é engenheiro de Controle e Automação pelo Instituto Nacional de Telecomunicações (Inatel). Atua no desenvolvimento de soluções tecnológicas nas áreas de hardware embarcado, Internet das Coisas (IoT) e inovação aplicada à saúde. É fundador e CEO da Amiko Soluções.



Filipe Bueno Vilela é engenheiro biomédico pelo Instituto Nacional de Telecomunicações (Inatel). Mestre e doutor em Engenharia de Materiais, na área de biomateriais, pela Universidade Federal de Itajubá (UNIFEI). Atua como professor, pesquisador e desenvolvedor no Inatel.

HelpFull: Monitoramento Emocional com Suporte de Inteligência Artificial em uma Plataforma Web

Arthur Pereira da Silva, Otto Simões Miranda, Pedro Henrique Diana Andolpho Evaristo,
Pedro Otávio de Oliveira Carlos, Jéssica Antonio Delgado, Isaque Katahira, Fábio Nogueira de Queiroz
Faculdade de Tecnologia de Ourinhos - FATEC OURINHOS
arthur.pereira1407@gmail.com, otto.miranda@fatec.sp.gov.br, pedrudih@gmail.com, pedro.carlos2803@gmail.com,
jessica.delgado@fatecourinhos.edu.br, isaque.katahira@fatecourinhos.edu.br, fabio.queiroz@fatecourinhos.edu.br

***Abstract*—Mental health has become an increasingly relevant topic in contemporary society, especially due to the growing number of cases of anxiety, stress, and depression. In this context, digital technologies have been explored as tools to support emotional well-being, allowing individuals to monitor their emotional states and develop self-care practices. Web-based applications focused on emotional monitoring can contribute to self-awareness by enabling the daily recording of emotions and personal reflections. This work presents the development of the HelpFull platform, a web-based system designed to help users track their emotional state over time. The application allows users to record emotions, write in a digital diary, and visualize mood patterns through graphical representations. In addition, the platform integrates an artificial intelligence chatbot capable of providing supportive messages related to mental health.The source code of the proposed system is available The source code of the HelpFull platform is publicly available at:[https://github.com/ikatahira/helpfull].**

***Index Terms*—artificial intelligence, emotional monitoring, mental health, self-care.**

***Resumo*—A saúde mental tem se tornado um tema cada vez mais relevante na sociedade contemporânea, especialmente diante do aumento de casos de ansiedade, estresse e depressão. Nesse contexto, tecnologias digitais têm sido exploradas como ferramentas de apoio ao bem-estar emocional, permitindo que indivíduos acompanhem seus estados afetivos e desenvolvam práticas de autocuidado. Aplicações web voltadas ao monitoramento emocional podem contribuir para o autoconhecimento ao possibilitar o registro diário de emoções e reflexões pessoais. Este trabalho apresenta o desenvolvimento da plataforma HelpFull, um sistema web projetado para auxiliar usuários no acompanhamento de seu estado emocional ao longo do tempo. A aplicação permite registrar emoções, escrever em um diário digital e visualizar padrões de humor por meio de gráficos. Além disso, integra um chatbot baseado em inteligência artificial capaz de oferecer mensagens de apoio relacionadas à saúde mental. O código-fonte do sistema proposto está disponível em: [https://github.com/ikatahira/helpfull].**

***Palavras chave*—autocuidado, inteligência artificial, monitoramento emocional, saúde mental.**

I. Introdução

A saúde mental consolidou-se como um dos temas centrais nas discussões sobre saúde pública e bem-estar na sociedade contemporânea. O aumento global nos índices de ansiedade, depressão e estresse tem impulsionado a busca por estratégias de acompanhamento emocional, visto que transtornos mentais comuns já afetam uma parcela significativa da população mundial, gerando impactos sociais e econômicos profundos [1]. Nesse cenário, o autocuidado surge não apenas como uma prática complementar ao suporte profissional, mas como um pilar essencial para a autorregulação e a manutenção da qualidade de vida [2].

Paralelamente, a ubiquidade das tecnologias digitais transformou a forma como os indivíduos interagem com a sua própria saúde. Dispositivos móveis e plataformas web oferecem novas possibilidades para o monitoramento contínuo de estados afetivos, permitindo que práticas tradicionais, como a escrita expressiva, sejam potencializadas por ferramentas de análise de dados. Segundo Pennebaker e Smyth[3], a externalização de sentimentos através da escrita promove benefícios cognitivos e imunológicos, auxiliando na organização de experiências traumáticas ou estressantes. A digitalização desse cuidado facilita o acesso à autoajuda e oferece mecanismos visualmente intuitivos para a compreensão de padrões de humor ao longo do tempo [4].

No entanto, o desenvolvimento de ferramentas que aliem segurança de dados, usabilidade e suporte emocional contextualizado ainda representa um desafio técnico e humanitário. Conforme Bender, Gebru e Shmitchell[5], o uso de modelos de linguagem e inteligência artificial generativa exige um design ético rigoroso para evitar vieses e garantir que a interface promova a reflexão e o autoconhecimento de forma segura.

Este artigo apresenta o desenvolvimento do sistema *HelpFull*, uma plataforma web projetada para auxiliar utilizadores no monitoramento da sua saúde mental. O sistema integra funcionalidades de diário eletrônico, visualização analítica de humor e um assistente virtual baseado em inteligência artificial generativa. O objetivo deste trabalho é descrever a arquitetura técnica da solução, as escolhas de segurança e como a integração entre computação e conceitos de psicologia pode resultar em uma ferramenta eficaz para o suporte ao bem-estar subjetivo.

Além disso, o software que implementa a plataforma *HelpFull* está disponível gratuitamente em https://github.com/ikatahira/helpfull. A versão disponível do repositório inclui o código-fonte completo tanto do sistema de gerenciamento baseado em PHP quanto do microsserviço em Python-Flask. Esta arquitetura permite que a plataforma seja implantada em ambientes locais ou em nuvem, viabilizando a integração do módulo de suporte emocional generativo.

II. Fundamentação Teórica

Esta seção apresenta os fundamentos teóricos que sustentam o desenvolvimento do sistema *HelpFull*. Para compreender a proposta da plataforma, são discutidos conceitos relacionados à saúde mental, ao autocuidado emocional e às estratégias de reconhecimento e monitoramento das emoções. Também são

abordadas técnicas utilizadas no apoio ao processamento emocional, como a identificação de estados afetivos e a escrita expressiva, além do papel das tecnologias digitais no acompanhamento do bem-estar psicológico. A articulação desses elementos teóricos permite compreender como soluções tecnológicas podem contribuir para o registro, análise e visualização de padrões emocionais, favorecendo práticas de autoconhecimento e autorregulação.

A. Saúde Mental e Autocuidado na Sociedade

A saúde mental pode ser compreendida como um estado relacionado ao bem-estar subjetivo e ao funcionamento positivo na vida, envolvendo dimensões emocionais, psicológicas e sociais. Nessa perspectiva, a saúde mental não se limita à ausência de transtornos mentais, mas também inclui a presença de níveis elevados de bem-estar e funcionamento social [6].

De acordo com a Organização Mundial da Saúde (OMS), a saúde mental é um estado de bem-estar no qual o indivíduo reconhece suas próprias capacidades, consegue lidar com as tensões normais da vida, trabalhar de forma produtiva e contribuir para a comunidade. Entretanto, transtornos mentais têm se tornado um importante problema de saúde pública em nível global. Estima-se que aproximadamente 4% da população mundial conviva com depressão, o que corresponde a cerca de 332 milhões de pessoas. Entre adultos, a prevalência é de cerca de 5,7%, sendo mais frequente entre mulheres do que entre homens. Esses dados evidenciam a relevância da depressão como um importante problema de saúde pública global[7].

Esses dados evidenciam a magnitude dos transtornos mentais como um desafio relevante para a saúde pública global, reforçando a importância de estratégias que promovam a conscientização, o acompanhamento e o cuidado contínuo com a saúde mental.

Nesse contexto, o cuidado contínuo com a saúde mental tem sido apontado como uma estratégia importante para a promoção do bem-estar psicológico. Estudos indicam que práticas de autocuidado contribuem significativamente para a prevenção do estresse, da ansiedade e do esgotamento emocional, além de favorecerem maior equilíbrio e qualidade de vida. Essas práticas envolvem ações realizadas pelo próprio indivíduo para reconhecer suas necessidades emocionais e adotar comportamentos que promovam o bem-estar mental de forma regular e preventiva [8].

B. Identificação e Monitoramento das Emoções

Reconhecer e compreender as próprias emoções é um elemento central da inteligência emocional. De acordo com Goleman, a autoconsciência emocional refere-se à capacidade de identificar os próprios sentimentos no momento em que ocorrem, permitindo que o indivíduo compreenda melhor suas reações e comportamentos [9].

De acordo com Gross[10], a regulação emocional envolve processos pelos quais os indivíduos influenciam as emoções que experimentam e a forma como essas emoções são expressas. Nesse contexto, pesquisas recentes indicam que tecnologias digitais possibilitam o monitoramento contínuo do humor e das emoções e contribuindo para a identificação de padrões relacionados ao bem-estar psicológico [11].

Nesse sentido, a identificação e o monitoramento das emoções constituem estratégias relevantes para o desenvolvimento do autoconhecimento e da regulação emocional. Ao acompanhar os próprios estados emocionais ao longo do tempo, os indivíduos podem compreender melhor seus padrões de comportamento e adotar estratégias mais adequadas para promover o bem-estar psicológico.

C. Registro e Visualização de Dados Emocionais

A escrita terapêutica pode ser definida como um processo de exploração de pensamentos e sentimentos pessoais por meio do ato de escrever, com o objetivo de promover autoconhecimento, crescimento pessoal e bem-estar psicológico. Essa técnica tem sido utilizada em diferentes abordagens psicoterapêuticas para auxiliar no tratamento de transtornos mentais e no processamento de experiências emocionais [12].

Além disso, o avanço das tecnologias digitais tem possibilitado o desenvolvimento de ferramentas voltadas ao acompanhamento da saúde mental, permitindo o registro e monitoramento de dados emocionais ao longo do tempo. Essas ferramentas podem auxiliar na identificação de padrões emocionais e apoiar práticas de autocuidado e acompanhamento do bem-estar psicológico [13].

De modo geral, os estudos apresentados evidenciam a importância do reconhecimento, registro e acompanhamento das emoções como estratégias para o desenvolvimento do autoconhecimento e da regulação emocional. A utilização de técnicas como a escrita terapêutica e o monitoramento de estados emocionais pode contribuir para o fortalecimento do bem-estar psicológico. Nesse contexto, o uso de tecnologias digitais surge como uma alternativa promissora para apoiar práticas de autocuidado e acompanhamento da saúde mental.

Dessa forma, a literatura aponta a eficácia do monitoramento emocional como ferramenta terapêutica. A partir destes conceitos, a seção seguinte descreve o desenvolvimento do sistema HelpFull, detalhando a infraestrutura computacional e as tecnologias empregadas para viabilizar a coleta e o processamento seguro de tais dados.

III. METODOLOGIA

O desenvolvimento do sistema *HelpFull* adotou uma metodologia de engenharia de software orientada a ciclos iterativos, priorizando a modularidade e a resiliência no tratamento de dados sensíveis. O projeto baseia-se numa arquitetura de serviços distribuídos, permitindo a separação de preocupações (*Separation of Concerns*) entre a persistência de dados transacionais e o processamento de linguagem natural (PLN).

Esta abordagem, articulada com os conceitos de monitoramento emocional discutidos na fundamentação teórica, viabilizou a definição dos requisitos técnicos necessários para a plataforma. O sistema opera através de três camadas tecnológicas integradas, sintetizadas na Tabela I.

TABELA I
RESUMO DAS TECNOLOGIAS.

Camada	Tecnologia	Função
Front	HTML/JS	Interação
Back	PHP/MySQL	Dados
IA	Python/Flask	Suporte

A. Infraestrutura e Persistência de Dados

A camada de aplicação web utiliza a *stack* LAMP. A persistência de dados no *MySQL* é mediada por *PHP Data Objects* (PDO), assegurando robustez contra ataques de *SQL Injection* por meio de *prepared statements*. A proteção de dados sensíveis é garantida pelo algoritmo de *hashing* BCRYPT, alinhando-se às boas práticas de segurança de informação na gestão de credenciais de utilizadores.

B. Processamento de Linguagem Natural

O suporte emocional é viabilizado por um microserviço em Python (Flask) que atua como um *gateway* para o modelo generativo Gemini 2.5 Flash. A lógica de interação é delimitada por *system instructions* que parametrizam o modelo para atuar estritamente no escopo de suporte ao bem-estar psicológico, garantindo a consistência ética das respostas.

C. Integração e Comunicação

A comunicação entre o *front-end* e os serviços de *back-end* é realizada de forma assíncrona via protocolo HTTP/REST, com trocas de dados em formato JSON. A configuração de políticas de CORS permite a interoperabilidade segura entre os componentes distribuídos.

D. Modelagem do Sistema

A modelagem do sistema foi consolidada através de diagramas UML, essenciais para a rastreabilidade entre requisitos e implementação. A Figura 1 delimita os atores e as fronteiras funcionais do sistema, enquanto a Figura 2 explicita a hierarquia de dados e relacionamentos da persistência.



Fig. 1. Diagrama de Casos de Uso do sistema *HelpFull*.

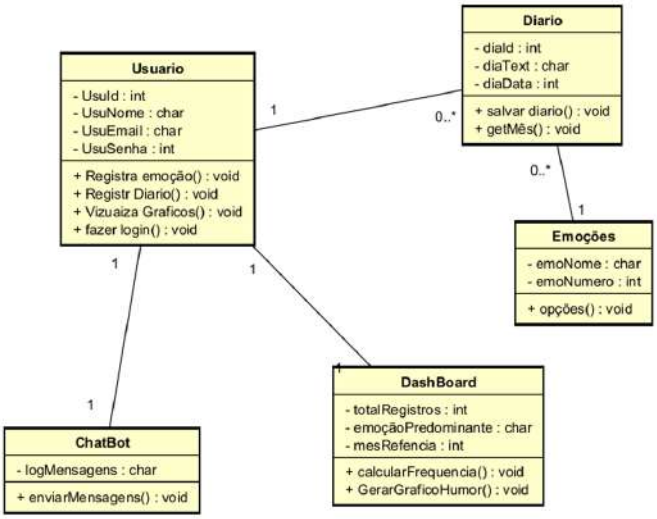


Fig. 2. Diagrama de Classes do sistema.

Com a infraestrutura e a modelagem do sistema devidamente delineadas, o processo de desenvolvimento permitiu a implementação do *HelpFull* como uma plataforma funcional de suporte ao bem-estar emocional. As escolhas tecnológicas e arquiteturais descritas nesta seção serviram como base para a operacionalização das funcionalidades de registro diário, análise de humor e interação assistida por IA. A seção subsequente apresenta os resultados obtidos com a aplicação prática do sistema, demonstrando o comportamento da plataforma em ambiente de uso e a eficácia das ferramentas desenvolvidas para o acompanhamento dos padrões emocionais dos utilizadores.

IV. RESULTADOS

Esta seção apresenta as principais funcionalidades desenvolvidas na plataforma *HelpFull*. Por meio das interfaces do sistema, é possível observar como o usuário pode registrar suas emoções, escrever no diário emocional, visualizar gráficos de acompanhamento e consultar o histórico de registros. As figuras a seguir ilustram as principais telas da aplicação e demonstram o funcionamento das funcionalidades implementadas.

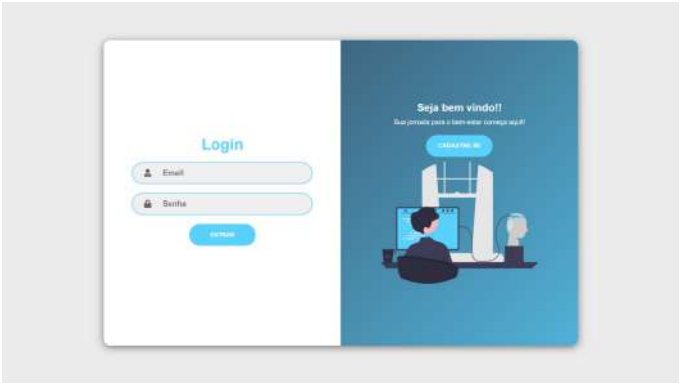


Fig. 3. Tela de login da plataforma HelpFull.

A Figura 3 apresenta a interface inicial de acesso ao sistema *HelpFull*. Nessa tela, o usuário pode inserir seu e-mail e senha para realizar o login na plataforma.

A Figura 4 apresenta a interface do diário emocional do sistema. Nessa tela, o usuário pode registrar um texto sobre seu

dia e selecionar a emoção que melhor representa seu estado emocional no momento, como feliz, triste, neutro, ansioso, bravo, calmo ou grato. Esses registros permitem acompanhar o estado emocional ao longo do tempo.



Fig. 4. Interface do diário emocional e registro de humor.

Com base nos registros emocionais realizados pelos usuários, o sistema também possibilita a visualização de padrões de humor ao longo do tempo. A Figura 5 apresenta um exemplo da distribuição dos humores registrados durante um determinado período de utilização da plataforma.



Fig. 5. Distribuição dos humores registrados pelos usuários ao longo do mês na plataforma *HelpFull*

Observa-se que o sistema permite identificar quais emoções foram registradas com maior frequência, possibilitando ao usuário refletir sobre seu estado emocional ao longo do tempo. Esse tipo de visualização contribui para o acompanhamento do bem-estar emocional e para o desenvolvimento do autoconhecimento.

Além das funcionalidades de registro emocional e diário digital, a plataforma *HelpFull* também conta com um *chatbot* de apoio emocional integrado ao sistema. A Figura 6 apresenta a interface dessa funcionalidade.

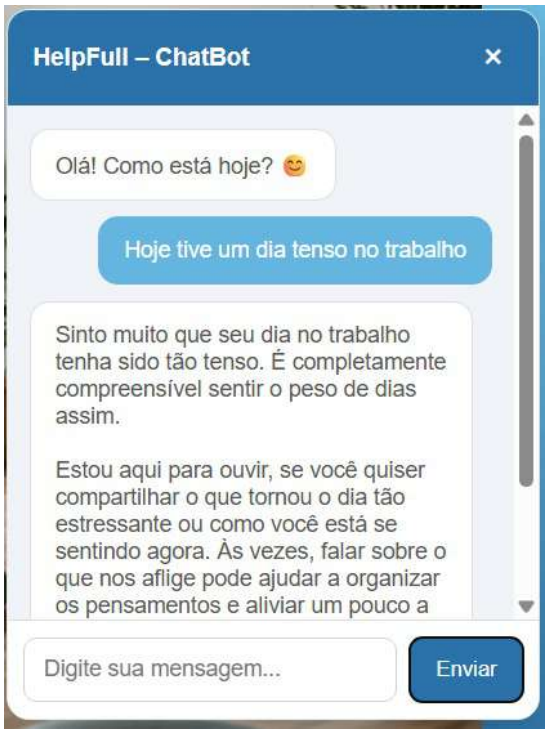


Fig. 6. Interface do chatbot de apoio emocional

O chatbot permite que o usuário descreva o que está sentindo e receba respostas automáticas com mensagens de acolhimento e orientações relacionadas ao bem-estar emocional. Para o desenvolvimento dessa funcionalidade foi utilizada a API do modelo de linguagem Gemini.

Durante a implementação, foi realizado um processo de configuração e treinamento do modelo para que a inteligência artificial responda apenas a conteúdos relacionados à saúde mental, autocuidado e bem-estar emocional. Dessa forma, busca-se garantir que as respostas geradas permaneçam dentro do escopo da plataforma.

O objetivo dessa funcionalidade é oferecer um canal inicial de escuta e apoio emocional ao usuário. O chatbot não substitui acompanhamento profissional, mas atua como um recurso complementar que incentiva a reflexão e o autocuidado.

Essas funcionalidades contribuem para o incentivo ao autoconhecimento e ao acompanhamento do bem-estar emocional.

V. Conclusão

Este trabalho apresentou o desenvolvimento da plataforma *HelpFull*, uma aplicação web voltada ao monitoramento emocional e ao incentivo ao autocuidado psicológico. O sistema permite que os usuários registrem suas emoções, escrevam em um diário digital e acompanhem seus registros ao longo do tempo por meio de visualizações gráficas.

Além disso, a integração de um chatbot baseado no modelo Gemini possibilita oferecer mensagens de apoio emocional dentro do escopo de saúde mental e bem-estar. Dessa forma, a plataforma contribui para incentivar o autoconhecimento e o acompanhamento do estado emocional dos usuários.

Como trabalhos futuros, destaca-se a possibilidade de ampliar os mecanismos de análise de dados emocionais, incorporar algoritmos de aprendizado de máquina para identificação de padrões mais complexos e realizar estudos de usabilidade

com um número maior de usuários, a fim de avaliar o impacto da plataforma no apoio ao bem-estar psicológico.

VI. Agradecimentos

Os autores agradecem à Faculdade de Tecnologia de Ourinhos (FATEC Ourinhos) pelo suporte institucional, pela orientação docente e pelos recursos disponibilizados que possibilitaram a realização deste projeto.

Referências

[1] World Health Organization. *World mental health report: Transforming mental health for all*. Geneva: World Health Organization, 2022. ISBN: 9789240049338.

[2] Martin E. P. Seligman. *Flourish: A Visionary New Understanding of Happiness and Well-being*. New York: Free Press, 2011. ISBN: 9781439190753.

[3] James W. Pennebaker e Joshua M. Smyth. *Opening up by writing it down: How expressive writing improves health and eases emotional pain*. 3rd. New York: Guilford Publications, 2016. ISBN: 9781462524921.

[4] John Torous, Nicholas C. Jacobson, Joseph P. Wright, Elena Rodriguez-Villa, Michael Robert Anderson, Shairi Turner, Atish Choudhury e Isaac Kohane. "Clinical review of user engagement with mental health apps". Em: *Digital Medicine* 3 (2017), pp. 1–10. DOI: 10.1038/s41746-018-0013-1.

[5] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major e Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" Em: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. New York, NY, USA: ACM, 2021, pp. 610–623. DOI: 10.1145/3442188.3445922.

[6] Corey L. M. Keyes. "The Mental Health Continuum: From Languishing to Flourishing in Life". Em: *Journal of Health and Social Behavior* 43.2 (2002), pp. 207–222.

[7] World Health Organization. *Depressive disorder (depression)*. Acesso em: 3 mar. 2026. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/depression.

[8] J. C. da Silva Junior. "Autocuidado e saúde mental: práticas, desafios e perspectivas para a promoção do bem-estar". Em: *Aurum Publicações* (2025).

[9] Daniel Goleman. *Emotional Intelligence: Why It Can Matter More Than IQ*. New York: Bantam Books, 1995.

[10] James J Gross. "The emerging field of emotion regulation: An integrative review". Em: *Review of general psychology* 2.3 (1998), pp. 271–299.

[11] Laurence Astill Wright, J. Monk-Cunliffe, B. Guo e R. Morriss. "Adverse Events of Mood Monitoring and Ambulatory Assessment in Depression and Bipolar Disorder: Systematic Review and Meta-Analysis". Em: *JMIR Mental Health* 12 (2025), e79500. DOI: 10.2196/79500. URL: https://mental.jmir.org/2025/1/e79500.

[12] Chiara Ruini e Cristina C Mortara. "Writing technique across psychotherapies—from traditional expressive writing to new positive psychology interventions: A narrative review". Em: *Journal of Contemporary Psychotherapy* 52.1 (2022), pp. 23–34.

[13] John Torous, Sandra Bucci, Imogen H Bell, Lars V Kessing, Maria Faurholt-Jepsen, Pauline Whelan, Andre F Carvalho, Matcheri Keshavan, Jake Linardon e Joseph Firth. "The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality". Em: *World psychiatry* 20.3 (2021), pp. 318–335.

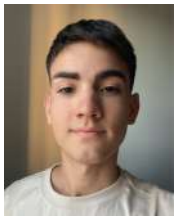
Autores



Pedro Henrique Diana Andolpho Evaristo Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Durante sua formação, tem se dedicado ao aprimoramento de competências na área de computação, com ênfase em programação, modelagem e gerenciamento de bancos de dados, redes de computadores e engenharia de software. Demonstra interesse no desenvolvimento de soluções tecnológicas e na aplicação de boas práticas de desenvolvimento, buscando constantemente aprofundar seus conhecimentos por meio de projetos acadêmicos, estudos complementares e participação em atividades relacionadas à área de tecnologia da informação.



Pedro Otavio de Oliveira Carlos Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Ao longo de sua formação acadêmica, tem se dedicado ao aprimoramento de conhecimentos na área de computação, com ênfase em programação, modelagem e gerenciamento de bancos de dados, redes de computadores e engenharia de software. Demonstra interesse no desenvolvimento de soluções tecnológicas voltadas para a resolução de problemas computacionais, bem como na aplicação de boas práticas de desenvolvimento e na utilização de metodologias que contribuam para a qualidade e eficiência de sistemas de informação.



Otto Simoes MirandaTécnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Durante sua formação acadêmica, tem se dedicado ao aprofundamento de conhecimentos nas áreas de programação, modelagem e gerenciamento de bancos de dados, redes de computadores e engenharia de software. Demonstra interesse no desenvolvimento de soluções computacionais eficientes, buscando aplicar boas práticas de desenvolvimento e ampliar suas competências técnicas por meio de estudos, projetos acadêmicos e atividades relacionadas à área de tecnologia da informação.



Arthur Pereira da Silva Técnico em Desenvolvimento de Sistemas pela Etec Jacinto Ferreira de Sá, atualmente é graduando no curso de Análise e Desenvolvimento de Sistemas pela Fatec Ourinhos. Atua no aprofundamento de conhecimentos em programação, banco de dados, redes complexas e engenharia de software.



Jéssica Antonio Delgado Possui graduação em Matemática pelo Centro Universitário Sagrado Coração (2013), mestrado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2016) e doutorado em Engenharia Elétrica pela Universidade Estadual Paulista Júlio de Mesquita Filho (2021). Tem experiência na área de Matemática, com ênfase em Engenharia Elétrica, atuando principalmente nos seguintes temas: método de barreira, fluxo de potência ótimo, método de barreira modificada, método de penalidade e otimização não linear. Atualmente é Professora na Faculdade de Tecnologia em Ourinhos (FATEC Ourinhos) e professora no UNIFIO.



Isaque Katahira Docente do Centro Estadual de Educação Tecnológica Paula Souza, atuando na Faculdade de Tecnologia de Ourinhos e na Escola Técnica Estadual Professor Mário Antônio Verza, tendo exercido a função de Diretor Substituto na Etec Professor Mário Antônio Verza (novembro de 2023 a abril de 2024). Ingressou na Etec em 2007 e na Faculdade de Tecnologia em 2017.Possui Licenciatura em Sistemas e Tecnologias da Informação (2006), Bacharelado em Desenvolvimento de Sistemas (2007) e Tecnólogo em Projeto de Sistemas de Informação (2008) pela Faculdade de Tecnologia de Ourinhos. Concluiu o Mestrado em Bioinformática pela Universidade Tecnológica Federal do Paraná (2018) e o Doutorado em Ciência da Informação pela Universidade Estadual Paulista UNESP (2022). Participa como revisor da Revista Eletrônica e-F@tec e é membro do Grupo de Pesquisa Representação Temática da Informação.



Fábio Nogueira de Queiroz Mestre em Ensino pela Universidade Estadual do Norte do Paraná, Programa de Pós-Graduação em Ensino (PPGEN/UENP). Possui graduação em Tecnologia em Processamento de Dados pela Faculdade de Tecnologia de Ourinhos - FATEC (2002). É licenciando em Matemática pela Universidade Estácio de Sá (2022), licenciado em Informática pela Faculdade de Tecnologia de Sorocaba - FATEC (2008). Licenciatura Plena em Pedagogia pela FALC (2015). Pós Graduação Latu Sensu em MBA Informática Tecnologia Internet (2004) e Pós Graduação Latu Sensu em Docência do Ensino Superior (2011). Atualmente é Professor II-E, na Fatec Ourinho e Coordenador da Etec Jacinto Ferreira de Sá em Ourinhos, no curso de Técnico em Desenvolvimento de Sistemas. Professor, nos cursos de Técnico em Desenvolvimento de Sistemas, Técnico em Informática para Internet e Técnico em Análise de Sistemas. Possui experiência na área de Informática, Técnica e Didática Pedagógica.

Desenvolvimento de um sistema embarcado baseado em visão artificial para avaliação pupilar e oculomotora

Brayan Armando Zeballos Rios, Maria Rene Ledezma Flores, Eynar Calle Viles
Universidad Privada del Valle – Univalle Cochabamba – Bolivia
zrb0030380@est.univalle.edu , mledezmaf@univalle.edu , ecallevi@univalle.edu
ORCID: 0009-0002-8814-8626 , ORCID: 0000-0002-6141-7075, ORCID:0000-0003-1477-2813

Abstract— This paper presents a non-invasive system for pupillary and oculomotor evaluation intended for outpatient consultation, based on embedded processing, binocular capture, and controlled light stimulation. The architecture integrates a processing unit, a dedicated controller for an addressable matrix, and an illuminance verification subsystem. The digital pipeline enables pupillary segmentation, center estimation, and the generation of quantifiable time series for both eyes. Spatial adjustment yielded an average conversion factor of 0.190 mm/px with a standard deviation of 0.002 mm/px, and errors of MAE = 0.022 mm and RMSE = 0.024 mm. Illuminance tests confirmed linear behavior under low-light conditions. Usability assessment reported a SUS score of 82.5/100, indicating operational acceptability of the system.

Index Terms—Artificial vision, digital image processing, ocular motility, pupillometry

Resumo—Apresenta-se um sistema não invasivo para avaliação pupilar e oculomotora voltado à consulta ambulatorial, baseado em processamento embarcado, captura binocular e estimulação luminosa controlada. A arquitetura integra uma unidade de processamento, um controlador dedicado para matriz endereçável e um subsistema de verificação de iluminância. O pipeline digital permite segmentação pupilar, estimativa do centro e geração de séries temporais quantificáveis para ambos os olhos. O ajuste espacial apresentou um fator médio de conversão de 0.190 mm/px com desvio padrão de 0.002 mm/px, e erros MAE = 0.022 mm e RMSE = 0.024 mm. A testes da iluminância confirmou comportamento linear em condições de penumbra. A avaliação de usabilidade apresentou SUS = 82.5/100, indicando aceitabilidade operacional do sistema.

Palavras chave—Motilidade ocular, processamento digital de imagens, pupilometria, visão artificial.

I. Introdução

A deficiência visual constitui uma carga sanitária relevante e concentra-se de forma desproporcional em países de baixa e média renda. O estudo Global Burden of Disease relatou que a magnitude global da perda visual requer estratégias de detecção e acompanhamento com capacidade de registro estruturado e comparável ao longo do tempo [1]. No contexto latino-americano, estudos populacionais baseados na metodologia RAAB descreveram uma demanda significativa por atendimento oftalmológico e necessidades de acesso à avaliação funcional em ambientes ambulatoriais comparáveis ao cenário boliviano [2].

A avaliação funcional do sistema visual inclui a análise da resposta pupilar e da motilidade ocular como parte do exame neuro-oftalmológico. Textos de referência descrevem que a avaliação clínica se apoia na observação da simetria pupilar, reflexo fotomotor e coordenação binocular, com aplicação de manobras de exploração visual para identificar alterações oculomotoras [3,4]. Do ponto de vista fisiológico, a dinâmica pupilar relaciona-se com o controle autonômico ocular e com respostas mensuráveis a estímulos luminosos, o que sustenta a necessidade de medições baseadas em critérios reprodutíveis [5]. No entanto, na prática ambulatorial, o registro permanece predominantemente qualitativo, o que reduz a padronização interobservador e limita a rastreabilidade entre episódios clínicos.

A literatura clínica descreveu o uso da pupilometria quantitativa para avaliação objetiva em cenários de maior complexidade e para detecção de mudanças sutis associadas a condições neurológicas ou farmacológicas [6,7], com considerações específicas para validação e qualidade metrológica das medições [8].

De forma complementar, revisões em video-oculografia e eye tracking relataram aplicações clínicas na análise de movimentos oculares, com métricas de acompanhamento e nistagmo, além de requisitos técnicos dependentes da arquitetura de captura e do processamento digital [9,10].

Contudo, os sistemas comerciais de pupilometria automatizada e eye tracking mantêm barreiras de adoção associadas ao custo de aquisição e a requisitos específicos de hardware e software [11,12]. Consequentemente, torna-se pertinente uma abordagem de engenharia biomédica voltada à integração de captura binocular, estimulação luminosa controlada e processamento digital de imagens em uma plataforma compacta para consulta ambulatorial.

Nesse contexto, o artigo apresenta um sistema não invasivo para avaliação pupilar e oculomotora baseado em uma unidade de processamento embarcado, controle dedicado de estímulos e aquisição com câmeras digitais, permitindo a geração de parâmetros quantificáveis para ambos os olhos.

O projeto incorporou critérios de segurança fotobiológica conforme a IEC 62471 [13], considerou princípios de usabilidade para dispositivos médicos segundo a IEC 62366-1 [14], e incluiu requisitos aplicáveis a instrumentos oftálmicos não invasivos descritos na ISO 15004-1 [15].

II. ARQUITETURA DO SISTEMA

Na Figura 1 apresenta-se o diagrama funcional do sistema, composto por quatro subsistemas:

1. Unidade de processamento embarcado
2. Controlador de estimulação luminosa
3. Subsistema de captura binocular
4. Subsistema de sensoriamento ambiental

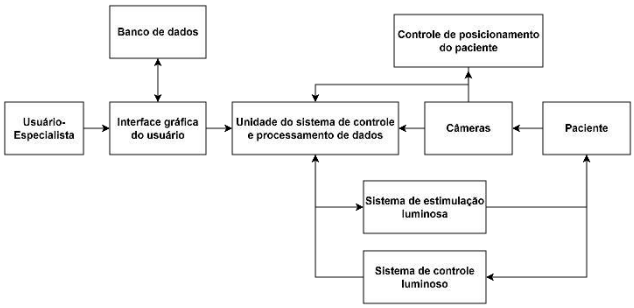


Fig. 1. Diagrama de blocos do sistema.

A unidade central de processamento coordena a aquisição de vídeo, o processamento digital de imagens, o armazenamento estruturado e a geração de relatórios. O controlador auxiliar gerencia a temporização e a geração de padrões de estimulação na matriz endereçável.

Na Tabela I são apresentados os principais componentes utilizados e a intercomunicação entre eles.

TABELA I
COMPONENTES PRINCIPAIS DO SISTEMA

Módulo	Função	Interface
Unidade de processamento (RPi 4)	Coordenação, aquisição de vídeo, processamento e banco de dados	USB, I2C, UART, HTTP
Microcontrolador (ESP32)	Controle de estímulos, temporização e confirmação	UART, GPIO
Matriz WS2812B	Estímulos visuais (cor/intensidade/padrão)	1 fio NRZ
Sensor BH1750	Verificação da iluminância ambiental	I2C

A geração de padrões na matriz LED é modelada pelo tempo de transmissão por quadro:

$$T_{frame}(N) = \frac{N*24}{f_{NRZ}} + T_{reset}, T_{reset} \geq 50\mu s \qquad (1)$$

Na equação (1), onde N é o número de LEDs, f_{NRZ} a frequência de transmissão e T_{reset} o pulso de latch requerido pelo protocolo do dispositivo.

O uso de buffer linear desacopla o endereçamento lógico da distribuição física, permitindo reprodutibilidade temporal na estimulação.

A. Subsistema óptico e de captura

A captura binocular é implementada com duas câmeras USB compatíveis com UVC, garantindo interoperabilidade com controladores padrão do sistema operacional. O regime operacional adotado é apresentado na Tabela II.

TABELA II
PARÂMETROS OPERACIONAIS DE CAPTURA

Parâmetro	Valor
Resolução	640×480 px
FPS	15 fps (OD) / 15 fps (OI)
Formato	MJPEG
Exposição	Manual fixa
Ganho / WB	Manual fixo
Foco	Manual

A seleção de MJPEG responde à estabilidade temporal e à menor complexidade de decodificação em sistemas embarcados. A resolução 640×480 representa um compromisso entre definição espacial e carga computacional.

B. Iluminação e sensoriamento ambiental

O sistema opera em penumbra controlada. A verificação da iluminância é realizada com sensor digital BH1750, cuja conversão nominal é expressa pela equação (2).

$$L_{lux} = \frac{raw}{1.2} \qquad (2)$$

O protocolo estabelece um intervalo de ajuste entre 0 e 20 lux, compatível com condições de avaliação pupilar.

C. Processamento digital de imagens

O pipeline organiza-se em duas vertentes:

c.1) Pupilometria

O processamento segue quatro etapas:

1. Aquisição e seleção da ROI
2. Pré-processamento (escala de cinza e filtragem)
3. Segmentação e ajuste geométrico
4. Conversão dimensional e geração de série temporal

A conversão de pixels para milímetros é obtida pela equação (3), utilizando k como fator de escala.

$$d_{mm}(t) = \bar{k} \cdot d_{px}(t) \qquad (3)$$

c.2) Oculomotricidade

O centro pupilar(x_p, y_p) é estimado em cada instante e os deslocamentos são calculados em relação a uma posição de referencia (x_{ref}, y_{ref}) pelas equações (4) e (5).

$$\Delta x(t) = x_p(t) - x_{ref} \qquad (4)$$

$$\Delta y(t) = y_p(t) - y_{ref} \qquad (5)$$

A equação (4) define o deslocamento horizontal $\Delta x(t)$ como a diferença entre a coordenada horizontal instantânea e a referência. De forma análoga, a equação (5) expressa o deslocamento vertical $\Delta y(t)$, permitindo quantificar variações no eixo vertical.

São geradas trajetórias 2D e curvas temporais para caracterização de sacadas, seguimento e fixação.

As Figuras 2, 3, 4 e 5 resumem o pipeline de visão artificial e processamento digital de imagens (PDI) aplicado à

segmentação pupilar. Na Fig. 2 realiza-se o pré-processamento com conversão para escala de cinza, equalização CLAHE e filtragem DoG e gaussiana para realçar bordas e reduzir ruído. A Fig. 3 mostra a definição da ROI e a detecção do glint. Na Fig. 4 aplica-se limiarização e operações morfológicas para refinar a máscara binária mediante análise de componentes conectados. Finalmente, na Fig. 5 extrai-se o contorno pupilar e ajusta-se uma circunferência para estimar centro e diâmetro, gerando métricas utilizadas na análise temporal.

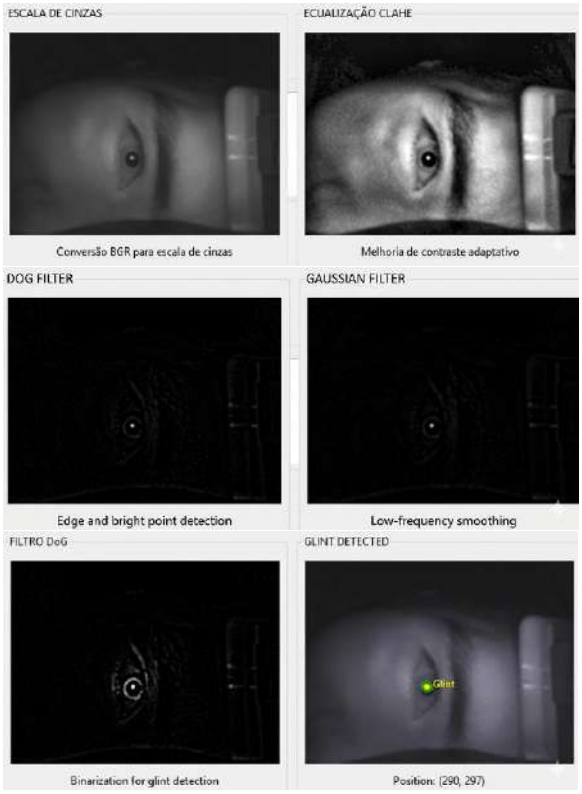


Fig. 2. Pipeline de pré-processamento de imagem: conversão para escala de cinza, equalização CLAHE e filtragem.

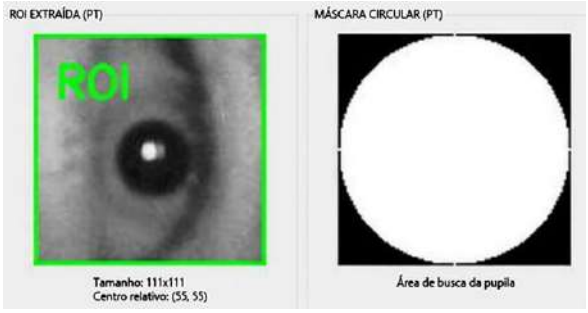


Fig. 3. Definição da região de interesse (ROI) e detecção do glint.

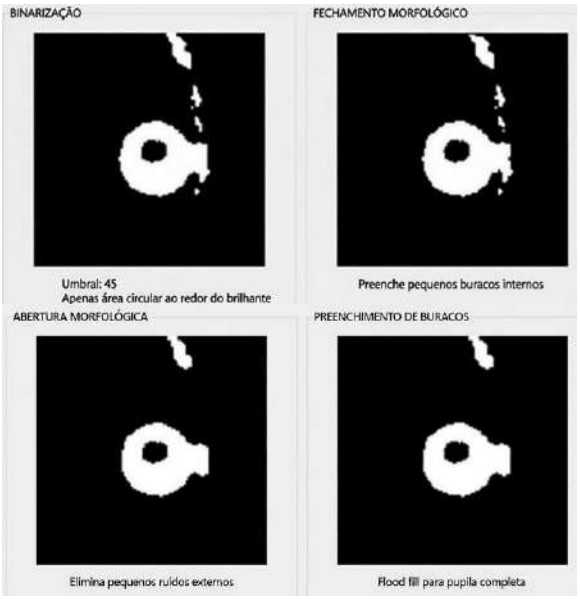


Fig. 4. Segmentação pupilar mediante limiarização e operações morfológicas.

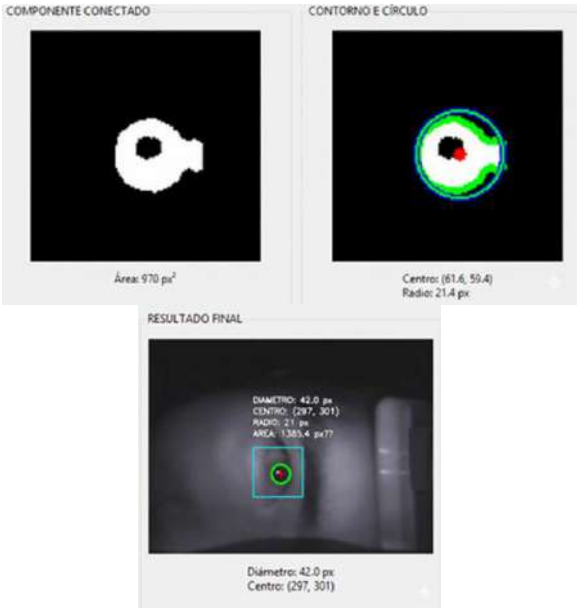


Fig. 5. Extração do contorno pupilar e ajuste circular.

D. Gestão de dados e rastreabilidade

A persistência é implementada por meio de um banco de dados relacional embarcado. A estrutura inclui:

- Identificação do paciente.
- Parâmetros da sessão.
- Condições de iluminância.
- Séries temporais.
- Evidência gráfica.

O modelo relacional permite a reconstrução íntegra de cada avaliação por meio de um identificador de relatório.

E. Validação e métricas

e.1) Validação da iluminância

Na equação (6) registra-se um conjunto de pares (L_{ref}, L_{sys}), onde L_{ref} corresponde ao luxímetro de referência e L_{sys} ao sensor.

$$L_{sys} = aL_{ref} + b \tag{6}$$

E calculam-se as métricas de erro:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |e_i| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{8}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} |\frac{e_i}{L_{ref,i}}| \tag{9}$$

As equações (7), (8) e (9) sustentam a análise quantitativa do sistema: permitem avaliar a linearização do sensor de iluminância em relação ao luxímetro de referência por meio de métricas de erro (MAE, RMSE e MAPE), assim como validar a conversão dimensional de pixels para milímetros empregada na estimativa do diâmetro pupilar.

e.2) Ajuste espacial (mm/px)

O fator médio de conversão obtido foi:

$$\bar{k} = 0.190 mm/px$$

com:

- $\sigma_k = 0.002 mm/px$
- MAE = 0.022 mm
- RMSE = 0.024 mm

Esses valores sustentam a precisão dimensional do sistema para medições pupilares.

III. Resultados

A. Validação de iluminância

O comportamento do sensor apresentou correlação linear no intervalo avaliado, como se observa na Figura 6, onde a dispersão L_{sys} *vs* L_{ref} evidencia proximidade em relação à referência e ausência de desvios não lineares relevantes no intervalo operacional.

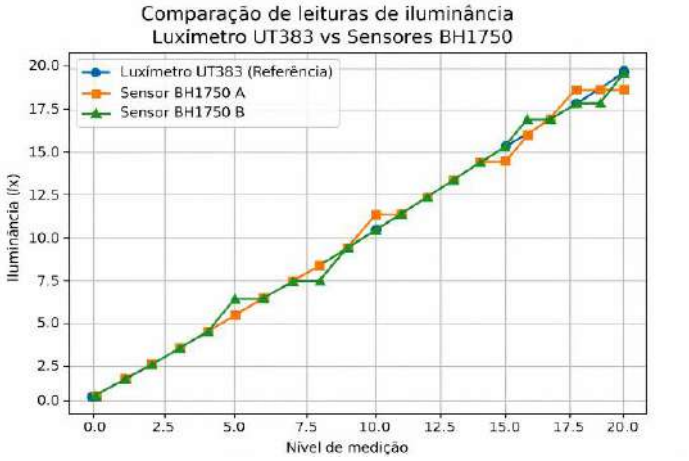


Fig. 6. Dispersão L_{sys} vs L_{ref}.

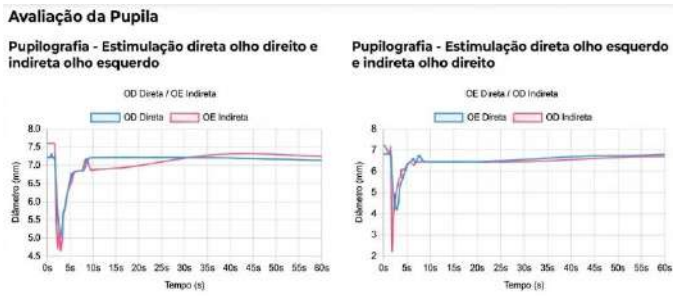
Os resíduos não evidenciaram comportamento não linear relevante no intervalo operacional.

B. Validação dimensional

O comportamento do sensor apresentou correlação linear no intervalo avaliado. O ajuste permitiu estabelecer equivalência funcional em condições de penumbra. O erro absoluto médio de 0.022 mm representa um desvio relativo inferior a 0.5 % para diâmetros pupilares clínicos habituais. Essa margem é considerada compatível com avaliação funcional ambulatorial.

A estabilidade de $\bar{k}$ sustenta o uso de uma constante única de conversão para cada configuração geométrica fixa.

C. Saídas funcionais



Métricas quantitativas da resposta pupilar

Parâmetro	Olho Direito (OD)	Olho Esquerdo (OE)	Diferença	Valor de Referência
Diâmetro basal (mm)	7.31	6.93	0.38	5.0 - 9.5 mm
Diâmetro mínimo (mm)	5.32	4.18	1.14	2.5 - 4.5 mm
% Constrição	27.2	39.7	12.5	> 25 %
Latência (ms)	200	150	50	200 - 500 ms
Tempo até o mínimo (ms)	840	800	40	600 - 1500 ms
Vel. de pico de contração (mm/s)	2.37	3.44	1.07	> 2.5 mm/s
Vel. de pico de dilatação (mm/s)	7.60	15.20	7.60	> 1.5 mm/s
Tempo de 75% de recuperação (ms)	1600	1000	600	1000 - 2500 ms

Fig. 7. Curvas temporais e métricas quantitativas da resposta pupilar.

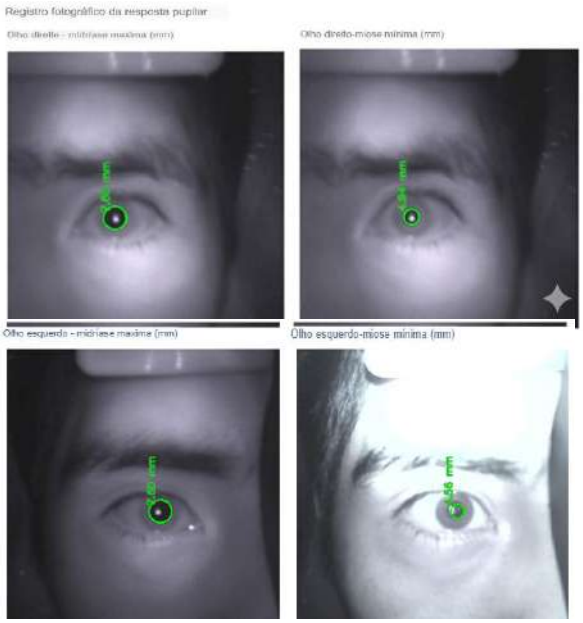
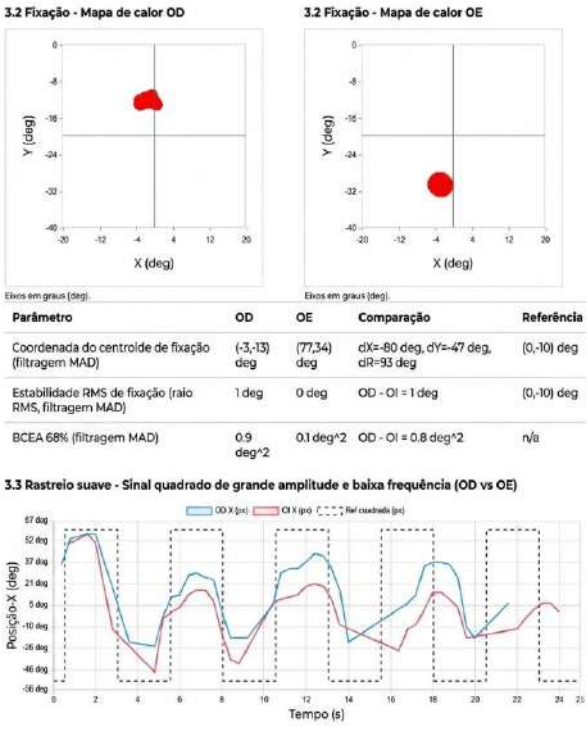


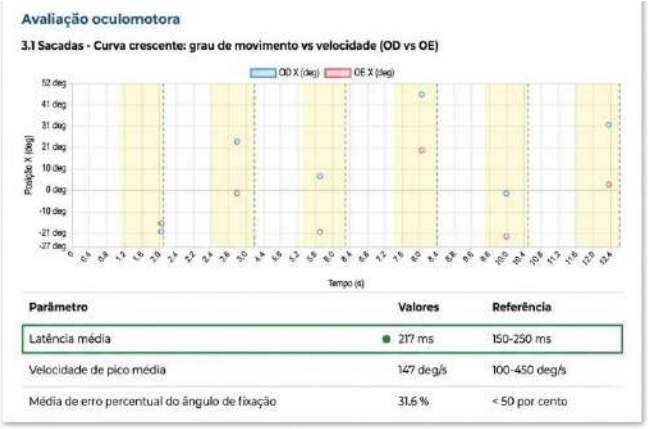
Fig. 8. Evidência visual da segmentação pupilar em midríase e miose máximas.

As Figuras 7 e 8 apresentam as saídas da avaliação de pupilometria. A Figura 7 reúne as curvas temporais da resposta pupilar sob estimulação direta e indireta e os parâmetros quantificáveis derivados do sinal, como diâmetro basal, diâmetro mínimo, latência e velocidades pupilares. A Figura 8 mostra os registros fotográficos das fases de midríase máxima e miose mínima para ambos os olhos obtidos pela segmentação pupilar.



Parâmetro	OD	OE	Comparação	Referência
Coordenada do centroide de fixação (filtragem MAD)	(-3,-13) deg	(77,34) deg	dX=-80 deg, dY=-47 deg, dR=93 deg	(0,-10) deg
Estabilidade RMS de fixação (raio RMS, filtragem MAD)	1 deg	0 deg	OD - OI = 1 deg	(0,-10) deg
BCEA 68% (filtragem MAD)	0.9 deg^2	0.1 deg^2	OD - OI = 0.8 deg^2	n/a

Fig. 9. Mapas de calor de fixação e registros de seguimento suave.



Parâmetro	Valores	Referência
Latência média	217 ms	150-250 ms
Velocidade de pico média	147 deg/s	100-450 deg/s
Média de erro percentual do ângulo de fixação	31.6 %	< 50 por cento

Fig. 10. Análise de sacadas baseada na relação ângulo–velocidade do movimento ocular.

As Figuras 9 e 10 apresentam as saídas da avaliação de motricidade ocular derivadas do rastreamento do centro pupilar. A Figura 9 inclui mapas de calor de fixação com métricas de dispersão e registros de seguimento suave com padrão temporal. A Figura 10 mostra a análise de sacadas por meio da relação entre ângulo de deslocamento e velocidade angular, permitindo caracterizar estabilidade de fixação, controle de seguimento e dinâmica sacádica.

D. Usabilidade

A avaliação por meio da System Usability Scale apresentou:

$$SUS = 82.5/100$$

Esse valor situa-se em faixas interpretativas associadas a alta aceitabilidade, conforme estudos de categorização do SUS.

E. Limitações

Não foi realizada comparação direta com pupilômetro clínico certificado nem com sistema de eye tracking de referência. A validação externa é definida como trabalho futuro. A frequência e a resolução foram restringidas pela capacidade de processamento embarcado, priorizando a estabilidade em relação ao aumento da definição espacial.

IV. Conclusão

O sistema integrou estimulação luminosa controlada, captura binocular, processamento digital de imagem e armazenamento estruturado em uma plataforma embarcada orientada à consulta ambulatorial. A validação metrológica confirmou precisão dimensional com fator médio de conversão de 0.190 mm/px, MAE de 0.022 mm e RMSE de 0.024 mm, bem como comportamento linear do sensor de iluminância no intervalo operacional de penumbra.

As saídas funcionais permitiram gerar pupilogramas bilaterais e trajetórias oculomotoras 2D com séries temporais reprodutíveis. Foram observadas diferenças quantificáveis em padrões de fixação, variações interoculares e comportamento diferencial na resposta pupilar direta e

consensual sob estimulação controlada. A arquitetura possibilita a análise de fotossensibilidade e comparação bilateral dentro de um ambiente de medição estruturado.

A principal contribuição reside na geração de registros quantificáveis e rastreáveis para avaliação funcional ocular ambulatorial por meio de uma arquitetura embarcada de baixo requisito computacional. Como trabalho futuro estabelece-se a validação comparativa dispositivo–dispositivo frente a sistemas clínicos certificados, bem como a ampliação do estudo em populações com alterações neuro-oftalmológicas documentadas.

Referências

[1] GBD 2020 Vision Loss Expert Group, "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The Right to Sight," *Lancet Global Health*, vol. 9, no. 2, pp. e144–e160, 2021.

[2] H. Limburg, F. Barria von-Bischhoffshausen, P. Gomez, J. C. Silva, and A. Foster, "Review of recent Rapid Assessment of Avoidable Blindness (RAAB) surveys in Latin America," *Ophthalmic Epidemiology*, vol. 21, no. 1, pp. 1–7, 2014.

[3] P. L. Kaufman and A. Alm, *Adler's Physiology of the Eye*, 12th ed. Amsterdam, The Netherlands: Elsevier, 2021.

[4] R. J. Leigh and D. S. Zee, *The Neurology of Eye Movements*, 5th ed. Oxford, U.K.: Oxford University Press, 2015.

[5] D. H. McDougal and P. D. Gamlin, "Autonomic control of the eye," *Comprehensive Physiology*, vol. 5, no. 1, pp. 439–473, 2015.

[6] C.-H. Hsu and L.-T. Kuo, "Application of pupillometry in neurocritical patients," *Journal of Personalized Medicine*, vol. 13, no. 7, p. 1100, 2023, doi: 10.3390/jpm13071100.

[7] W. R. Taylor, J. W. Chen, B. S. Meltzer, and J. R. Madsen, "Quantitative pupillometry in clinical practice," *Current Neurology and Neuroscience Report*s, vol. 21, no. 1, pp. 1–10, 2021.

[8] J. J. McAnany and Y. Lim, "Clinical considerations in the validation of pupillometry measurements," *Vision Research*, vol. 188, pp. 37–46, 2021.

[9] A. J. Larrazabal et al., "Video-oculography eye tracking towards clinical applications: A review," *Computers in Biology and Medicine*, vol. 108, pp. 57–66, 2019, doi: 10.1016/j.compbiomed.2019.03.025.

[10] G. C. Musat et al., "Ocular movement examination in peripheral vestibular disorders: A narrative review," *Medicina*, vol. 60, no. 10, p. 1665, 2024, doi: 10.3390/medicina60101665.

[11] VisuMed Equipment, "Neuroptics NPi-200 pupillometer," 2023. [Online]. Available: https://www.neuroptics.com

[12] B. Farnsworth, "Comprehensive guide to eye tracker prices: 15+ options reviewed for every budget," iMotions, Dec. 19, 2024. [Online]. Available: https://imotions.com

[13] IEC, IEC 62471:2006 – *Photobiological Safety of Lamps and Lamp Systems*. Geneva, Switzerland: International Electrotechnical Commission, 2006.

[14] IEC, IEC 62366-1:2015 – Medical Devices—*Application of Usability Engineering to Medical Devices*. Geneva, Switzerland: International Electrotechnical Commission, 2015.

[15] ISO, ISO 15004-1:2007 – *Ophthalmic Instruments—Fundamental Requirements and Test Methods*. Geneva, Switzerland: International Organization for Standardization, 2007.

Autores



Brayan Armando Zeballos Rios é egresso do curso de Engenharia Biomédica da Universidad Privada del Valle (Univalle), sede Cochabamba, Bolívia. É membro da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB). Seus interesses concentram-se no projeto e modelagem tridimensional de estruturas e dispositivos biomédicos, com aplicação em projetos de desenvolvimento tecnológico e em ambientes profissionais vinculados a equipamentos médicos. Além disso, dedica-se à análise de dados e imagens biomédicas por meio de técnicas de processamento digital, segmentação e extração de características a partir de sinais e informações visuais estruturadas.



Maria Rene Ledezma Flores Engenheira Biomédica pela Univalle - Universidad del Valle (2018), Mestrado em Engenharia Elétrica com ênfase em Biomecânica pela Udelar - Universidad de la República (2023). Atualmente, é professora em tempo integral de Engenharia Biomédica, trabalhando principalmente em projetos de pesquisa em biomecânica e ergonomia.



Eynar Calle Viles é mestre em Telecomunicações, com especialização em Redes de Comunicações e Telemática, e possui pós-graduação em Internet das Coisas (IoT) pelo Instituto Nacional de Telecomunicações (INATEL), Brasil.

Com mais de 10 anos de experiência, é especialista em projeto e implementação de redes, FTTH, IoT e sistemas embarcados. Desenvolveu diversas inovações tecnológicas, incluindo propostas de patentes, como o "Robô Móvel para Detecção de Metais" e o "Sistema de Localização para o Satélite Tupac Katari".

Além disso, possui várias publicações e participações em revistas e congressos internacionais. Suas principais áreas de interesse incluem antenas e propagação, segurança da informação e IoT, contribuindo ativamente para o desenvolvimento de soluções tecnológicas inovadoras.

Atualmente, atua como coordenador de pesquisa na Faculdade de Informática e Eletrônica da Universidad Privada del Valle (Univalle) e é presidente da IEEE Computer Society Bolivia.

Estimativa de Pressão Arterial e Glicemia por Meio de Oxímetro de Pulso e Inteligência Artificial

Jhonny Martin Camacho Gonzales, Mauricio Marcelo Peredo Claros, Eynar Calle Viles
Universidad Privada del Valle – Univalle Cochabamba - Bolivia
cgj0029756@est.univalle.edu, peredo.mauri@gmail.com, ecallevi@univalle.edu
ORCID: 0009-0002-1700-5927, 0000-0002-5884-3321, 0000-0003-1477-2813

Abstract— This paper proposes the development of CardioPress, a non-invasive auxiliary medical device that estimates blood pressure and blood glucose using photoplethysmographic (PPG) signals and artificial intelligence algorithms. Given the high prevalence of hypertension and diabetes in Bolivia, as well as the limitations of conventional monitoring methods, a technological solution is proposed to improve treatment adherence for cardiovascular and metabolic diseases. The system utilizes optical sensors to measure heart rate and oxygen saturation; these data are processed by an ESP32 microcontroller and analyzed by artificial neural networks trained to estimate clinical parameters. The neural network model was optimized to achieve a more robust and adaptive structure for patients with diverse physiological profiles. The prototype includes a web graphical interface, a relational database, and 3D-printed housing designed in Fusion 360. Experimental tests demonstrated high precision in blood pressure estimation ($R^2 > 0.88$) and a moderate correlation in glucose estimation ($R^2 \approx 0.68$), validating the feasibility of the non-invasive approach.

Index Terms—Artificial Intelligence (AI), Blood Glucose, Blood Pressure (BP), Photoplethysmography (PPG).

Resumo— Este artigo propõe o desenvolvimento do CardioPress, um dispositivo médico auxiliar não invasivo que estima a pressão arterial e a glicose sanguínea por meio de sinais fotopletismográficos (PPG) e algoritmos de inteligência artificial. Diante da alta prevalência de hipertensão e diabetes na Bolívia, e das limitações dos métodos convencionais de monitoramento, propõe-se uma solução tecnológica que melhore a adesão ao tratamento de doenças cardiovasculares e metabólicas. O sistema utiliza sensores ópticos para medir a frequência cardíaca e a saturação de oxigênio, cujos dados são processados por um microcontrolador ESP32 e analisados por redes neurais artificiais treinadas para estimar os parâmetros clínicos. O modelo de rede neural foi otimizado para apresentar uma estrutura mais robusta e adaptativa a pacientes com perfis fisiológicos distintos. O protótipo inclui uma interface gráfica web, um banco de dados relacional e uma carcaça 3D projetada no Fusion 360. Os testes experimentais demonstraram alta precisão na estimativa da pressão arterial ($R^2 > 0.88$) e uma correlação moderada na estimativa de glicose ($R^2 \approx 0.68$), validando a viabilidade da abordagem não invasiva.

Palavras-chave— Fotopletismografia (PPG), Glicose Sanguínea, Inteligência Artificial (IA), Pressão Arterial (PA).

I. INTRODUÇÃO

O débito cardíaco e as doenças cardiovasculares relacionadas, como a hipertensão arterial, as arritmias cardíacas e a diabetes mellitus, representam um crescente problema de saúde pública tanto em nível global quanto nacional. Segundo dados da OMS, na Bolívia, a prevalência da hipertensão afeta mais de 23% da população e a diabetes atinge 12,8%, estando estreitamente relacionada ao risco de desenvolvimento de complicações cardiovasculares e metabólicas. Para manejar essas condições de forma eficaz, é necessário o monitoramento contínuo de parâmetros como a pressão arterial e a glicemia [1].

Os métodos convencionais para monitorar parâmetros fisiológicos, como a pressão arterial e a glicose no sangue, apresentam limitações quando se exige vigilância contínua. Essas barreiras incluem a possibilidade de gerar desapego ou rejeição por parte de alguns pacientes devido ao desconforto dos procedimentos, como a punção capilar ou o uso de braçadeiras infláveis. Além disso, a falta de monitoramento representa um risco significativo para quem sofre de doenças como hipertensão e diabetes mellitus. A ausência de controle regular desses parâmetros pode desencadear emergências como crises hipertensivas ou hipoglicemias graves [2].

Para enfrentar essa problemática, propõe-se o desenvolvimento de um dispositivo não invasivo que emprega tecnologias como a fotopletismografia (PPG) e a inteligência artificial (IA). Este dispositivo foi projetado para estimar e monitorar a pressão arterial sem braçadeiras (cuffless) e a glicose no sangue sem a necessidade de punções ou tiras reativas. Utiliza um sensor óptico que mede a frequência cardíaca e a saturação de oxigênio, sendo os dados processados por um microcontrolador. Em seguida, são analisados por redes neurais artificiais para estimar os valores clínicos, melhorando a adesão ao tratamento e a gestão de doenças de base [3].

II. DESIGN E DESENVOLVIMENTO DO HARDWARE

O fluxograma do CardioPress, apresentado na Figura 1, inicia-se com a captura de sinais fotopletismográficos (PPG) por meio de um sensor de grau médico Medtronic Nellcor DS-100A, o qual é conectado a um microcontrolador ESP32 encarregado da digitalização e do pré-processamento em tempo real através de um filtro de resposta ao impulso finita (FIR) para garantir a estabilidade do sinal. Deste processo, extraem-se a frequência cardíaca (FC) e a saturação periférica de oxigênio (SpO_2), parâmetros que são exibidos localmente em uma tela sensível ao toque (TFT) de 2,8 polegadas com comunicação SPI, permitindo ao usuário gerenciar os modos de medição de forma intuitiva. Simultaneamente, o dispositivo serializa os dados em formato JSON e os transmite via protocolo HTTP para um servidor web, onde um modelo de Rede Neural Artificial (RNA) estima a pressão arterial sistólica, diastólica e

os níveis de glicose no sangue. Finalmente, os resultados calculados retornam ao microcontrolador para visualização imediata e são armazenados em um banco de dados MySQL, ficando disponíveis para consulta remota por meio de uma interface gráfica de usuário (GUI) que garante a rastreabilidade histórica do paciente.

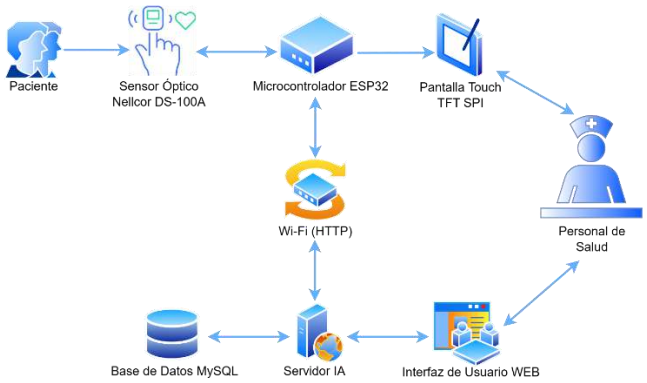


Fig. 1. Diagrama de blocos do projeto.

O sistema de aquisição, apresentado na Figura 2, utiliza um domínio de alimentação de 5 V para a etapa analógica e controle digital. A etapa de condicionamento divide-se em duas fases baseadas no amplificador operacional LM358: uma primeira fase de filtragem de altas frequências e uma segunda fase configurada como amplificador não inversor. O ganho teórico (G) desta etapa é fundamental para ajustar o sinal do fotodetector à faixa dinâmica do ADC do ESP32, e é definido pela relação de resistências, conforme mostrado na equação (1):

$$G = 1 + \frac{R_f}{R_{in}} \tag{1}$$

Onde:

- R_f é a resistência de realimentação, estabelecida em 220 $k\Omega$.
- R_{in} é a resistência de entrada, estabelecida em 2.7 $k\Omega$.

Substituindo os valores do projeto na Equação (1), obtém-se:

$$G = 1 + \frac{220\ 000}{2\ 700}$$

$$\mathrm{G} = 82$$

Esse ganho de 82 permite elevar a amplitude do sinal pulsátil a níveis ótimos (0–5 V), evitando a saturação graças a um trimmer de ajuste de offset que centraliza o sinal. Finalmente, foi incorporada uma rede de compensação com um capacitor de 22 pF no laço de realimentação para estabilizar o ganho em altas frequências e reduzir o ruído transitório, assegurando uma saída de baixa impedância compatível com o microcontrolador.

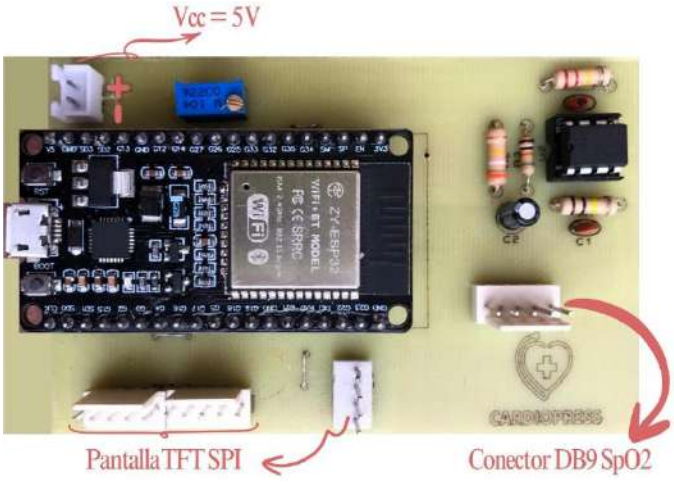


Fig. 2. Circuito Implementado.

Na Figura 2, é apresentado o resultado da placa de circuito impresso (PCI) do dispositivo, onde se observa uma distribuição focada no isolamento das seções analógica e digital para evitar ruídos na obtenção do sinal do sensor de oximetria de pulso, mantendo o conector Molex próximo; por outro lado, também estão presentes os conectores Molex para a tela TFT SPI de 2,8" com driver ST7789.

A. Cálculo de Consumo

Considerando a saída do módulo TP4056+ de 5 V a 1,4 A, alimentou-se o sensor de oximetria de pulso e a tela TFT sensível ao toque de 2,8 polegadas.

Os valores de Corrente Nominal I_N e Corrente Máxima I_{MAX}, são detalhados nas especificações de cada componente utilizado no dispositivo para calcular a Corrente Total do Circuito I_{TC} e determinar o valor de consumo ideal para a autonomia da bateria.

Tela sensível ao toque TFT ST7789 de 2,8 polegadas:

- $I1_N = 60\ mA$
- $I1_{MAX} = 120\ mA$

Sensor de oximetria de pulso Nellcor DS100A [4]:

- $I2_N = 30\ mA$
- $I2_{MAX} = 40\ mA$

Placa de Desenvolvimento: $I3_{MAX} = 350\ mA$.

A Corrente Total do Circuito I_{TC}, foi calculada por meio do somatório de todos os componentes do dispositivo, conforme mostrado na equação (2):

$$I_{TC} = I1_{MAX} + I2_{MAX} + I3_{MAX} \tag{2}$$

$$I_{TC} = (120 + 40 + 350)\ mA$$

$$I_{TC} = 510\ mA$$

Segundo AUTARQUÍA [5], no projeto de sistemas eletrônicos, é recomendável adicionar uma margem de segurança de 20% a 30% sobre o consumo estimado, especialmente em aplicações onde a carga pode variar dinamicamente ou onde foram utilizados componentes com tolerâncias amplas. Esta margem permite evitar sobrecargas, melhorar a eficiência térmica e prolongar a vida útil do sistema.

Portanto, o consumo total de corrente do circuito I_C acrescido de 30% foi calculado com base na equação (3):

$$I_C = (30\% * I_{TC}) + I_{TC} \quad (3)$$

$$I_C = 663\ mA$$

A autonomia ideal do sistema foi calculada com base nos dados da Tabela I:

TABELA I
DADOS PARA O CÁLCULO DE AUTONOMIA

COMPONENTE	VALOR
Capacidade da Bateria 18650 Cafini de 4,2 V	1200 mAh
Consumo Total +30% Margen	663 mA
Tensão de Saída do Módulo TP4056+	5 V

Substituindo os valores na fórmula básica da equação (4) para o cálculo de autonomia ideal [5]:

$$Autonomía_{IDEAL} = \frac{Capacidade\ da\ Bateria\ (mAh)}{Consumo\ Total\ (mA)} \quad (4)$$

$$Autonomía_{IDEAL} = \frac{1200}{663} \cong 1.81\ horas$$

Visto que o módulo utilizou um conversor elevador de tensão (boost converter), estimou-se uma eficiência de 85%, baseando-se no valor típico de eficiência para módulos de carga linear. Por conseguinte, isto reduziu a energia útil disponível ao considerar um fator de descarga real de aproximadamente 90% para a bateria [5].

Dessa forma, a autonomia real do dispositivo foi calculada a partir da equação (5):

$$Autonomía_{REAL} = Autonomía_{IDEAL} * 85\% * 90\% \quad (5)$$

$$Autonomía_{REAL} = 1.81 * 0.85 * 0.9$$

$$Autonomía_{REAL} = 1.4\ horas$$

Portanto, determinou-se que o sistema poderia operar de forma ideal por um período aproximado de 1,4 horas com uma bateria de 1200 mAh, sob um consumo contínuo do circuito de 663 mA, considerando a eficiência e a perda por descarga.

B. Desenvolvimento da Rede Neural:

Para a estimativa da pressão arterial (sistólica e diastólica) e da glicose sanguínea a partir dos sinais adquiridos, implementou-se uma Rede Neural Artificial (RNA) desenvolvida com as bibliotecas TensorFlow e Keras. O modelo utiliza como variáveis de entrada a frequência cardíaca (FC) e a saturação de oxigênio (SpO_2) obtidas pelo hardware, aplicando técnicas de normalização e padronização para otimizar o processamento. O treinamento foi realizado sobre um conjunto de dados de 410 registros experimentais, empregando algoritmos de otimização e esquemas de validação cruzada para assegurar a capacidade de generalização do sistema. A arquitetura foi ajustada para maximizar a correlação entre os parâmetros biopotenciais e as métricas clínicas de referência, permitindo que o dispositivo funcione como uma ferramenta auxiliar robusta para o monitoramento não invasivo de variáveis metabólicas e cardiovasculares.

TABELA II
RESUMO ESTATÍSTICO DO DATASET

PARÂMETRO	MÍN.	MÁX.	MÉDIA	UNIDADE
Frequência cardíaca (FC)	50	104	76.88	bat/min
Saturação de oxigênio (SpO_2)	85	100	92.56	%
Pressão sistólica (PS)	90	166	124.24	mmHg
Pressão diastólica (PD)	60	109	84.08	mmHg
Glicemia (BG)	71	154	97.21	mg/dL

Na Tabela II, apresentam-se os valores mínimos, máximos e as médias. Esta caracterização permitiu evidenciar a variabilidade fisiológica da amostra e a amplitude dos intervalos de treinamento do modelo neural. Os dados foram normalizados antes do processo de treinamento a fim de otimizar a convergência do modelo.

C. Projeto da Carcaça 3D

Nesta seção, realizou-se o projeto 3D da tampa e da base do dispositivo.

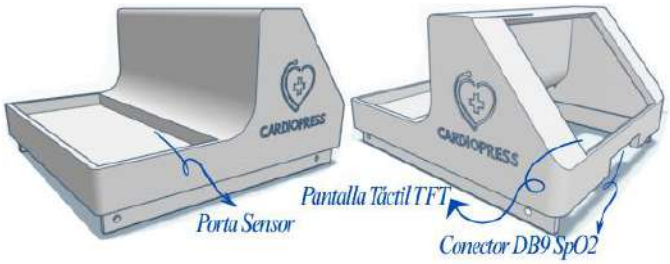


Fig. 3. Design 3D da tampa.

Na Figura 3, apresenta-se o design superior do CardioPress, desenvolvido mediante modelagem tridimensional sob critérios de ergonomia e proteção de componentes, estruturando-se em duas seções principais impressas em PLA. A tampa superior atua como interface operacional, integrando de forma estratégica a tela sensível ao toque (TFT) para a visualização de parâmetros, uma porta DB9 lateral que reduz a fadiga mecânica do sensor Nellcor DS100A e uma cavidade específica (Porta Sensor) projetada para o resguardo do transdutor biopotencial.

O design da base do dispositivo, mostrado na Figura 4, constitui o suporte estrutural principal, onde são alojados de maneira ordenada a placa eletrônica, o sistema de gestão de carga e a bateria de Li-ion. Este projeto inclui aberturas para ventilação passiva e bordas suavizadas que favorecem a portabilidade, alcançando um equilíbrio entre a funcionalidade

biomédica e a robustez necessária para o monitoramento ambulatorial.

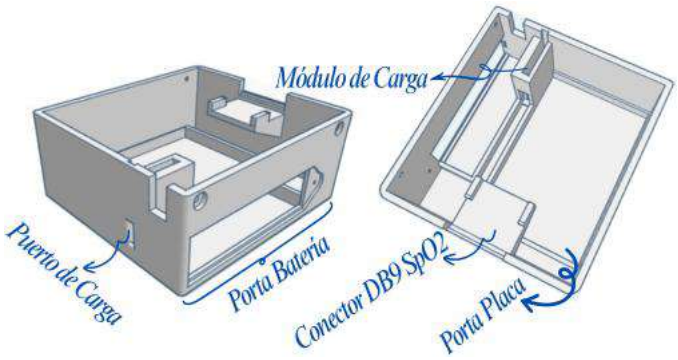


Fig. 4. Design 3D da base.

O interior da base do CardioPress foi estruturado mediante cinco zonas funcionais otimizadas para a estabilidade mecânica e a gestão térmica do sistema. O projeto inclui um alojamento central com guias de posicionamento para a PCB principal, assegurando uma fixação rígida que previne deslocamentos durante o transporte. Adjacente a esta seção, integrou-se um compartimento para uma célula de íon-lítio tipo 18650, cuja localização estratégica contribui para o equilíbrio do centro de massa do dispositivo e favorece a dissipação térmica passiva. Finalmente, na extremidade lateral, dispôs-se uma cavidade dedicada para a porta de carga e o módulo TP4056+ com conversor elevador de tensão integrado; esta configuração permite o acesso externo para a alimentação via micro-USB sem comprometer a integridade do encapsulamento, isolando termicamente o módulo de potência dos sensores analógicos sensíveis.

III. RESULTADOS

Os resultados obtidos durante o desenvolvimento, implementação e validação do dispositivo biomédico CardioPress correspondem ao monitoramento não invasivo da pressão arterial e da glicose sanguínea por meio de oximetria de pulso e inteligência artificial. Ao longo do projeto, foram realizados testes experimentais, processos de calibração de sensores, treinamentos de redes neurais artificiais e melhorias contínuas no design do sistema, tanto no componente de hardware quanto no de software.

A. Testes de Precisão do Sensor de Oximetria de Pulso

A fase inicial de validação concentrou-se na caracterização do sensor de oximetria para garantir a integridade das variáveis de entrada do sistema. Dado que a frequência cardíaca (FC) e a saturação de oxigênio (SpO_2) constituem a fonte primária de dados para o modelo de rede neural.

Na Figura 5, apresentam-se os testes de precisão realizados com o simulador de Oximetria de Pulso (SpO_2) CONTEC modelo MS100 [6].



Fig. 5. Teste experimental de medição de Frequência Cardíaca e Saturação de Oxigênio com simulador certificado.

Este equipamento permitiu gerar sinais fisiológicos controlados e certificados, os quais foram contrastados com as leituras obtidas pelo hardware desenvolvido para verificar a concordância e a linearidade na etapa de transdução. Os resultados do processo de ajuste asseguraram uma alta precisão na captura de biopotenciais, garantindo que os parâmetros transmitidos ao servidor para processamento via inteligência artificial fossem um reflexo fiel das condições fisiológicas simuladas, mitigando, assim, erros de propagação na estimativa clínica.

B. Ensaio de Validação do Canal de FC e SpO_2 do Sensor

Para verificar a precisão do sensor óptico Nellcor DS100A, realizaram-se testes controlados utilizando um simulador de oximetria de pulso CONTEC MS100 [1]. O objetivo do ensaio foi comparar a leitura de frequência cardíaca (FC) registrada pelo sensor conectado ao sistema CardioPress com os valores de referência gerados pelo simulador.

A Figura 6 apresenta o gráfico de dispersão que mostra a correlação entre os valores de frequência cardíaca (FC) obtidos pelo sensor Nellcor DS100A e os gerados pelo simulador de SpO_2 CONTEC MS100. No eixo horizontal, representam-se os valores do simulador, enquanto no eixo vertical são plotados os valores medidos pelo sensor.

Cada ponto vermelho corresponde a uma medição individual, e a linha pontilhada cinza representa o ajuste linear por regressão, evidenciando a relação entre ambas as variáveis. O gráfico indica os seguintes valores estatísticos fundamentais:

- Coeficiente de correlação (r): 0.994
- Coeficiente de determinação (R^2): 0.975

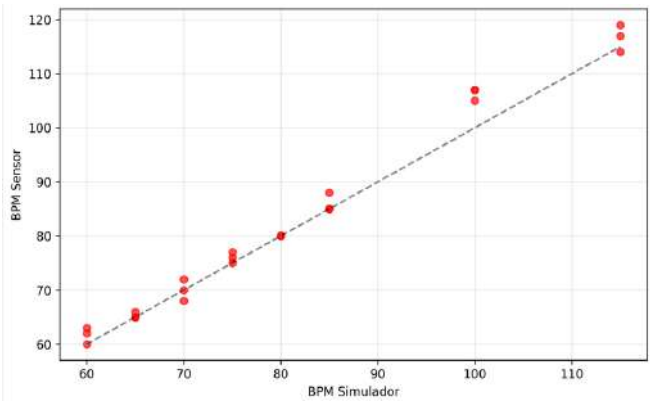


Fig. 6. Correlação entre leituras de Frequência Cardíaca (FC) do sensor Nellcor DS100A e o simulador CONTEC MS100.

A Figura 7 apresenta o gráfico de dispersão que representa a correlação entre os valores de saturação de oxigênio (SpO_2) obtidos pelo sensor Nellcor DS100A e os valores gerados pelo simulador CONTEC MS100. O eixo horizontal indica os valores simulados, enquanto o eixo vertical mostra os valores medidos pelo sensor. Os pontos azuis representam as amostras individuais, e uma linha de regressão tracejada indica o ajuste linear entre ambos os conjuntos de dados.

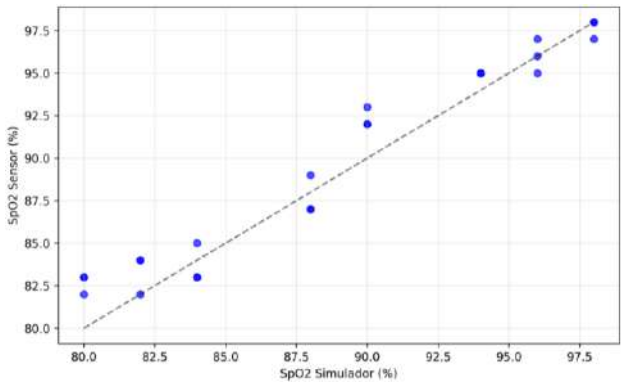


Fig. 7. Correlação de valores simulados e medidos de oxigênio (SpO_2) entre o sensor Nellcor DS100A vs. simulador CONTEC MS100.

O gráfico reporta um coeficiente de correlação de Pearson r = 0.977 e um coeficiente de determinação R^2 = 0.937, o que evidencia uma correlação positiva muito forte e uma alta precisão do sensor na medição de SpO_2 sob condições simuladas.

C. Testes Experimentais em Ambiente Real e Cotidiano

O processo de coleta de dados baseou-se no monitoramento diário de sinais vitais (pressão arterial, glicemia, frequência cardíaca e saturação de oxigênio) em pacientes com perfis fisiológicos detalhados, estabelecendo um regime de quatro medições diárias para capturar variações circadianas. Para otimizar a precisão do dispositivo CardioPress, o sistema integra o perfil patológico individual nas etapas de desnormalização e padronização dos dados, permitindo um ajuste dinâmico da saída da rede neural de acordo com a faixa fisiológica específica de cada usuário.



Fig. 8. Paciente com medições convencionais e CardioPress.

As medições mostradas na Figura 8 correspondem a um paciente masculino de 53 anos com diagnóstico clínico de hipertensão arterial e diabetes tipo 2. Devido a uma adesão farmacológica irregular, este indivíduo apresenta variabilidades hemodinâmicas significativas, o que proporcionou um cenário crítico para avaliar a capacidade de resposta do sistema diante de flutuações biológicas não reguladas por medicação. Mediante o contraste entre a instrumentação médica convencional e as estimativas do CardioPress, verificou-se uma resposta ideal da rede neural após o processo de retreinamento e personalização de dados. Os resultados demonstraram que o algoritmo possui a robustez necessária para processar sinais complexos e alinhar as estimativas de pressão arterial e glicemia com os valores de referência, validando a eficácia do modelo de inteligência artificial para adaptar-se a perfis metabólicos específicos, mesmo em condições de instabilidade terapêutica.

TABELA III
TABELA DE ERRO PERCENTUAL DO PACIENTE

TIPO	ERRO PA (%)	ERRO Glu (%)	ERRO FC (%)	ERRO SpO_2 (%)
Pós-Jantar	5.89	11.29	5.26	2.17
Pós-Almoço	4.78	2.86	1.45	2.17
Pós-Jejum (curto)	5.82	6.29	1.39	3.19
Em Jejum	7.95	3.47	1.49	1.08
Erro Médio	6.11	5.98	2.40	2.15

A Tabela III mostra o desempenho do sistema para o sujeito de estudo avaliado mediante a análise de erros percentuais frente a padrões de referência em quatro estados metabólicos, cujos resultados sintetizam-se em um erro médio geral notavelmente baixo.

IV. CONCLUSÃO

O desenvolvimento do CardioPress valida a viabilidade técnica do emprego de sinais fotopletismográficos e aprendizado profundo (deep learning) para a estimativa não invasiva de parâmetros clínicos, consolidando uma ferramenta auxiliar com alta aceitação por parte de especialistas e

pacientes. Através da integração de um sensor Nellcor DS100A e um microcontrolador ESP32, alcançou-se uma autonomia de 1,4 horas e uma precisão notável em ambientes controlados, com erros de 2,01% em frequência cardíaca e 1,50% em SpO_2. O modelo de rede neural em ambientes reais depende do uso de históricos médicos independentes para mitigar a variabilidade interindividual.

Finalmente, a implementação de uma arquitetura em nuvem via HTTP/Wi-Fi e uma interface TFT sensível ao toque permitiu desacoplar a carga computacional e garantir a rastreabilidade dos dados, consolidando um sistema portátil e funcional que, sob condições de carga superiores a 50%, oferece uma alternativa robusta para o monitoramento preventivo de doenças cardiovasculares e metabólicas.

REFERENCIAS

[1] OMS, «Obesidad y sobrepeso», 2023. [En línea]. Disponible en: https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight

[2] L. E. Litwak, I. Querzoli, C. Musso, A. Dain, S. Houssay, y J. E. Costa Gil, «Monitoreo continuo de glucosa: Utilidad e indicaciones», *Medicina (Buenos Aires)*, vol. 79, n.º 1, pp. 44-52, feb. 2019, [En línea]. Disponible en: https://www.scielo.org.ar/scielo.php?script=sci_abstract&pid=S0025-76802019000100007&lng=es&nrm=iso&tlng=es

[3] J. M. C. Gonzales, E. C. Viles, y M. P. Claros, «Estimation of Blood Pressure and Glucose through Photoplethysmography and Artificial Intelligence», en *2024 3rd International Congress of Biomedical Engineering and Bioengineering (CIIBBI)*, 2024, pp. 1-4. doi: 10.1109/CIIBBI63846.2024.10784558.

[4] Nellcor Puritan Bennett Inc., «DS-100A Durasensor®», 2005.

[5] A. Y. AUTARQUÍA, «SELECCIÓN Y CÁLCULO DE LAS PROTECCIONES SEGÚN EL REBT», junio de 2021. [En línea]. Disponible en: https://autarquiapersonal.com/2021/06/19/seleccion-de-las-protecciones-electricas-segun-el-rebt/

[6] XingMedical, «SIMULADOR SPO2 CONTEC MS100», 2025. [En línea]. Disponible en: https://xingmedical.com/simulador-spo2-contec-ms100

AUTORES



Jhonny Martin Camacho Gonzales é pesquisador do Departamento de Pesquisa da Universidad Privada del Valle (UNIVALLE), sede Cochabamba, Bolívia. Atualmente, está cursando o último período da graduação em Engenharia Biomédica na mesma instituição, onde tem se destacado por seu foco na integração de sistemas eletrônicos e aprendizado profundo (deep learning) aplicados ao diagnóstico clínico. Seu trabalho de pesquisa centra-se no desenvolvimento de dispositivos médicos não invasivos, captura de sinais fotopletismográficos e processamento por meio de redes neurais artificiais. Liderou o design e a implementação do protótipo CardioPress, integrando hardware de alta precisão e arquiteturas em nuvem para monitoramento ambulatorial. Suas principais áreas de interesse incluem o projeto de sistemas embarcados, instrumentação médica avançada e telemedicina.



Mauricio Marcelo Peredo Claros é mestre em tecnologias da informação e docente universitário na área de Engenharia Biomédica. Seu trabalho centra-se na integração de tecnologias emergentes, particularmente na adoção de inteligência artificial em processos educativos e de pesquisa. Conta com diversas publicações científicas em revistas internacionais relacionadas à tecnologia educativa, inteligência artificial aplicada, telemedicina e desenvolvimento de soluções biomédicas. Participa ativamente em processos de pesquisa aplicada e inovação tecnológica, promovendo o uso de ferramentas digitais, sensores biomédicos e tecnologias de inteligência artificial na formação universitária. Suas principais linhas de interesse incluem inteligência artificial, inovação educativa, instrumentação biomédica e tecnologias para a melhoria da qualidade educativa.



Eynar Calle Viles é mestre em Telecomunicações, com especialização em Redes de Comunicações e Telemática, e possui pós-graduação em Internet das Coisas (IoT) pelo Instituto Nacional de Telecomunicações (INATEL), Brasil. Com mais de 10 anos de experiência, é especialista em projeto e implementação de redes, FTTH, IoT e sistemas embarcados. Desenvolveu diversas inovações tecnológicas, incluindo propostas de patentes, como o “Robô Móvel para Detecção de Metais” e o “Sistema de Localização para o Satélite Tupac Katari”. Além disso, possui várias publicações e participações em revistas e congressos internacionais. Suas principais áreas de interesse incluem antenas e propagação, segurança da informação e IoT, contribuindo ativamente para o desenvolvimento de soluções tecnológicas inovadoras. Atualmente, atua como coordenador de pesquisa na Faculdade de Informática e Eletrônica da Universidad Privada del Valle (Univalle) e é presidente da IEEE Computer Society Bolivia.

Criação de Projeto para Inclusão de Robótica em Colégios do Ensino Médio

Alexandre Baratella Lugli, Eugênio Moreira Jonas da Silva,
Egídio Raimundo Neto, Luiz Fernando Domiciano Ribeiro dos Santos
Maria Luiza Muquem Alves, Marco Antônio Shima, Wanderson Eleutério Saldanha
Instituto Nacional de Telecomunicações - Inatel
baratella@inatel.br, eugenio@inatel.br, egidioneto@inatel.br
luizf.santos@inatel.br, maria.alves@gel.inatel.br, marco.shima@gea.inatel.br, wsaldanha@inatel.br

***Abstract*—This study addresses the inclusion of robotics in high schools, an initiative developed by the National Institute of Telecommunications (Inatel) in collaboration with the Start Inatel team. The project stands out for promoting inclusion, sustainability, and accessibility, encouraging the development of skills such as mechanical construction, programming, and teamwork through the creation of Cupim robots. The initiative highlights the importance of robotics in the education of high school students, fostering its implementation in schools.**

***Index Terms*—Cupim, High School, Robotics, STEM Learning.**

***Resumo*—Este estudo aborda a inclusão da robótica no ensino médio, uma iniciativa desenvolvida pelo Instituto Nacional de Telecomunicações (Inatel) juntamente com a equipe Start Inatel. O projeto destaca-se por promover inclusão, sustentabilidade e acessibilidade, incentivando o desenvolvimento de habilidades como construção mecânica, programação e trabalho em equipe por meio da criação de robôs da categoria Cupim. A iniciativa reforça a importância da robótica na formação dos alunos do ensino médio, estimulando sua implementação nas escolas.**

***Palavras chave*—Aprendizagem em STEM, Cupim, Ensino médio, Robótica.**

A inserção da robótica no ensino médio tem se consolidado como uma abordagem inovadora para promover o desenvolvimento de competências técnicas, cognitivas e sociais nos estudantes. Mais do que ensinar conceitos relacionados à programação, eletrônica e construção mecânica, a robótica educacional proporciona um ambiente de aprendizagem prática no qual os alunos podem aplicar conhecimentos teóricos em desafios reais. Esse processo contribui para o fortalecimento do raciocínio lógico, da criatividade, da capacidade [1] analítica e da resolução de problemas, competências consideradas essenciais para a formação de profissionais preparados para os desafios tecnológicos e científicos contemporâneos. Além dos ganhos técnicos, a robótica educacional também favorece a utilização de metodologias ativas de ensino, nas quais o estudante assume papel central no processo de aprendizagem. Nesse contexto, o aluno deixa de ser apenas receptor de conteúdo e passa a atuar como protagonista na construção do conhecimento, por meio da experimentação, testes, tomada de decisões e aprimoramento de soluções. Essa abordagem torna o ensino mais dinâmico e significativo, ampliando o engajamento dos estudantes e contribuindo para o desenvolvimento da autonomia, do pensamento crítico e do trabalho colaborativo. Outro aspecto relevante da robótica aplicada ao ambiente escolar [2] é seu caráter interdisciplinar. Projetos robóticos possibilitam integrar conhecimentos de áreas como matemática, física, computação e engenharia,[3] permitindo que os conteúdos sejam trabalhados de forma prática e contextualizada. Além disso, atividades dessa natureza estimulam habilidades socioemocionais importantes, como cooperação, comunicação, liderança e gestão de desafios, características amplamente valorizadas tanto no ambiente acadêmico quanto profissional. Apesar desses benefícios, ainda existem obstáculos para a implementação estruturada da robótica em muitas escolas de ensino médio. Entre as principais dificuldades estão a limitação de recursos, a falta de iniciativas contínuas voltadas à prática competitiva e a escassez de projetos que aproximem os estudantes de experiências mais desafiadoras e motivadoras. Embora existam diversas ações voltadas ao ensino de robótica, muitas delas se concentram em atividades introdutórias, sem explorar plenamente o potencial das competições como ferramenta educacional para aprofundamento técnico e desenvolvimento estratégico. Nesse cenário, competições de robótica surgem como uma solução promissora para ampliar o interesse dos alunos e fortalecer o aprendizado prático. Ao envolver desafios reais, restrições de projeto e necessidade de tomada de decisão sob competição, essas iniciativas criam um ambiente que estimula inovação, criatividade e aplicação de conhecimentos em situações concretas. Além disso, competições proporcionam experiências que aproximam os estudantes de contextos semelhantes aos encontrados em projetos de engenharia e desenvolvimento tecnológico. Dentro desse contexto, destaca-se a categoria Cupim como uma proposta diferenciada e acessível para o ambiente educacional. Essa modalidade de competição utiliza predominantemente madeira na construção dos robôs, impondo desafios específicos relacionados à resistência estrutural, durabilidade e desempenho mecânico [4]. Essas limitações tornam o processo de desenvolvimento mais criativo e exigem dos participantes soluções inovadoras para otimização do projeto. Dessa forma, a categoria não apenas promove aprendizado técnico, mas também incentiva experimentação e pensamento estratégico. Um importante referencial nesse cenário é o Start Inatel, equipe do Instituto Nacional de Telecomunicações especializada no desenvolvimento e participação em competições da categoria Cupim. A experiência acumulada pela equipe oferece base

técnica e metodológica para adaptação desse modelo competitivo ao ambiente escolar, tornando possível sua utilização como ferramenta de incentivo à aprendizagem e à inovação entre estudantes do ensino médio. Diante disso, identifica-se como problema central deste trabalho a necessidade de criar estratégias que promovam maior inserção dos alunos do ensino médio em atividades práticas e competitivas [5] de robótica, superando limitações presentes em abordagens tradicionais e ampliando oportunidades de desenvolvimento técnico e educacional. Como resposta a esse problema, este artigo propõe o desenvolvimento de competições voltadas para estudantes do ensino médio inspiradas na categoria Cupim, utilizando a experiência do Start Inatel como referência para estruturação da proposta. Assim, o objetivo deste artigo é propor e desenvolver competições educacionais de robótica para alunos do ensino médio, utilizando a categoria Cupim [6] como base para fomentar o interesse pela tecnologia, estimular o aprendizado interdisciplinar e contribuir para a formação de estudantes mais preparados para desafios futuros nas áreas de ciência, tecnologia e inovação.

EXEMPLO DE UM ROBÔ:

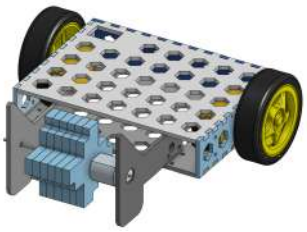


Figura 1. ROBÔ

I. Fundamentação Teórica

A. Regras e Limitações

Na categoria Cupim [6], os robôs devem pesar entre 150 gramas e 27 quilos, o que exige dos participantes estratégias na construção dos seus modelos. As batalhas têm duração máxima de 3 minutos, e o principal objetivo é nocautear o robô adversário ou, caso isso não seja possível, aplicar o maior número de golpes eficazes durante esse tempo. Além disso, os robôs precisam ser projetados de forma que consigam passar pelas portas da arena, o que impõe uma limitação de tamanho e exige precisão na construção para garantir que o robô consiga acessar a área de combate.

As arenas possuem dimensões específicas, o que restringe ainda mais as opções de design, forçando as equipes a adaptarem seus projetos para serem competitivos dentro desse espaço limitado. As lutas são desafiadoras, já que, além de um bom desempenho ofensivo, os robôs precisam ser resistentes o suficiente para suportar os impactos e ataques do oponente, além de serem rápidos o bastante para esquivar-se ou contra-atacar. Essas regras e limitações fazem com que os participantes desenvolvam estratégias e soluções criativas.

B. Desafios da categoria

Os robôs da categoria Cupim enfrentam uma série de desafios técnicos e estratégicos que testam os limites da criatividade e inovação dos participantes. Um dos principais desafios é a construção de um robô robusto o suficiente para suportar os impactos violentos durante as batalhas, sem perder a agilidade necessária para manobras rápidas e eficazes [4]. A escolha de materiais também é crucial, pois, embora a madeira seja leve, ela pode ser menos resistente em comparação com outros materiais, exigindo soluções inteligentes para garantir durabilidade sem comprometer o desempenho.

Outro desafio importante é a programação, que deve permitir que o robô reaja rapidamente às ações do adversário, seja para atacar, esquivar-se ou se defender. As equipes também precisam considerar a estratégia de combate, que envolve não só a capacidade de aplicar golpes, mas também a habilidade de detectar fraquezas no oponente. Esses desafios tornam a categoria Cupim uma categoria emocionante e complexa dentro das competições de robótica.

Os robôs da categoria Cupim são construídos com uma variedade de materiais, sendo que cada escolha possui um papel específico na funcionalidade, resistência e desempenho do robô durante as batalhas. A madeira também permite fácil personalização do design, podendo ser cortada, lixada e montada com ferramentas simples. Os principais materiais utilizados são:



Figura 2. Madeira MDF

Madeira: A madeira é o material principal na construção dos robôs Cupim, como o nome da categoria sugere. Ela é escolhida por ser leve, de baixo custo e fácil de manusear, sendo usada principalmente na estrutura do robô, como base e chassi. Entre os tipos utilizados destacam-se o MDF (Medium Density Fiberboard), popular devido à sua rigidez e facilidade de corte, e o compensado, que é leve e relativamente resistente.



Figura 3. Motor DC 130

Motores e servos: Os motores elétricos DC são responsáveis pela locomoção do robô, acionando as rodas ou outros

sistemas de movimentação. Já os servos são utilizados para controlar mecanismos de ataque, como braços ou lâminas.

Suporte para baterias: Os suportes para baterias são utilizados para manter as baterias firmemente fixadas, garantindo conexão elétrica segura e estável entre as baterias e o circuito.



Figura 4. Baterias

Baterias: As baterias fornecem a energia necessária para o funcionamento do robô. Baterias de lítio, como as LiPo, são amplamente utilizadas devido à alta densidade de energia, oferecendo autonomia e potência em formato compacto e leve.



Figura 5. Regulador de Tensão

Regulador de tensão: O regulador é utilizado para reduzir a tensão da bateria, que pode variar de 6,4 V a 8,4 V, para 5 V, valor necessário para alimentação da ESP32. Sua principal função é manter a tensão de saída constante mesmo com variações na alimentação.



Figura 6. Esp32

ESP32: O ESP32 é o microcontrolador central do robô, responsável pelo controle e comunicação do sistema. Possui Wi-Fi e Bluetooth integrados, opera em nível lógico de 3,3 V e pode ser alimentado pelo pino VIN com 5 V regulados. Para evitar interferências eletromagnéticas e falhas de comunicação, a antena deve permanecer afastada de fios de potência e motores.



Figura 7. Fios e conectores

Fios e conectores: São responsáveis pela ligação entre os componentes eletrônicos, como motores, baterias e placa de controle. A escolha dos fios depende da corrente exigida pelo sistema.



Figura 8. Rodas

Rodas: As rodas podem ser fabricadas em plástico, borracha ou metal, dependendo das necessidades de tração e durabilidade. Frequentemente são revestidas com borracha para melhorar a aderência ao solo.

Colas e parafusos: São utilizados para fixação estrutural dos componentes, garantindo que não se soltem durante o combate.

Elementos de ataque: Dependendo da estratégia adotada, o robô pode possuir mecanismos de ataque, somente usando o MDF como material, projetados para empurrar ou danificar o oponente.

Circuitos e placas eletrônicas: São responsáveis pelo gerenciamento dos sinais dos sensores, acionamento dos motores e integração entre os componentes do robô.



Figura 9. Chave Liga/desliga

Chave liga/desliga: Controla o fornecimento de energia ao sistema. Na posição desligada o circuito é interrompido; na posição ligada, a corrente elétrica é fornecida aos componentes.



Figura 10. LED

LED indicador: O LED atua como indicador visual do estado do robô, sinalizando quando o sistema está energizado. Por ser polarizado, requer ligação correta entre Vcc e GND.

Aplicação prática:

Na aplicação prática, esses materiais foram integrados para compor um robô funcional da categoria Cupim. A madeira foi utilizada na estrutura principal devido à sua leveza e facilidade de usinagem. Componentes plásticos e metálicos foram aplicados em regiões sujeitas a impacto e em mecanismos estruturais específicos.

O sistema de movimentação foi implementado com motores DC acoplados às rodas, enquanto a alimentação elétrica foi realizada por baterias LiPo conectadas a um regulador de tensão para fornecer 5 V à ESP32. O microcontrolador foi utilizado para realizar o controle do robô via comunicação sem fio.

Fios, conectores, chave liga/desliga e LED indicador foram empregados para completar a integração elétrica do sistema, garantindo funcionamento seguro e monitoramento do estado do robô.

Dessa forma, a combinação desses materiais e componentes possibilitou o desenvolvimento de um robô funcional, resistente e adequado às exigências da categoria.

II. Estudos

A. Resultados

No ano de 2024, diversas competições foram organizadas para incentivar os alunos do ensino médio a vivenciarem experiências práticas na categoria Cupim. A iniciativa teve como objetivo estimular o interesse dos jovens pela tecnologia, robótica e trabalho em equipe. Para isso, o Inatel visitou diferentes escolas, promovendo a importância do projeto e engajando estudantes e professores na criação de soluções criativas e inovadoras.

Cada escola participante foi convidada a formar grupos compostos por três alunos, que se dedicaram ao desenvolvimento e montagem de robôs. Esse processo incluiu desde o planejamento e construção até testes e ajustes necessários para o bom desempenho nas competições. Após a etapa de montagem, os grupos competiram entre si nas escolas aonde o time vencedor de cada escola foi selecionado para representar a instituição na grande final aonde acontece no Inatel.

O projeto foi incluído em 11 escolas, envolvendo diretamente 246 alunos distribuídos da seguinte forma:
-78 alunos do 1.º ano do ensino médio,
-83 alunos do 2.º ano do ensino médio,
-73 alunos do 3.º ano do ensino médio,
-12 participantes cuja série não foi informada.

A etapa final foi realizada no próprio Inatel, contando com a participação de 306 inscritos, incluindo alunos e professores das escolas participantes. Esse evento final foi um momento de grande integração, com demonstrações de criatividade, espírito de equipe e talento por parte dos competidores.

Além de promover a inclusão da robótica, o projeto também incentivou a troca de experiências entre os estudantes de diferentes escolas. A categoria Cupim continua a desempenhar um papel fundamental na formação de jovens interessados em tecnologia, abrindo portas para futuras carreiras e mostrando que o trabalho em equipe é essencial para carreira.

III. Conclusão

A construção do robô Cupim trouxe uma ótica incrível de aprendizado prático, ao permitir uma exploração mais aprofundada de conceitos fundamentais em robótica e eletrônica. Durante a atividade educativa, houve a disponibilização de recursos extra para ajudar os alunos a entenderem os temas discutidos de forma mais clara e objetiva. A montagem da placa foi feita antecipadamente, com o propósito de simplificar o entendimento e permitir que os estudantes pudessem replicar o processo com maior facilidade. Isso foi feito com o objetivo principal de facilitar a aprendizagem e garantir a participação ativa de todos no projeto.

Ao longo do progresso do projeto, encontramos dificuldades, e novas ideias surgiram, destacando a importância da colaboração em equipe e da análise crítica para criar solucionamentos eficazes.

Para futuros trabalhos, sugerimos a elaboração de vídeos mais detalhados e avançados, para apresentar uma visão técnica organizada das etapas de desenvolvimento dos robôs aos estudantes, visando melhorias no desempenho para competir com sucesso no final da competição.

Esses esforços não só vão aumentar a concorrência, mas também vão ajudar a educar alunos mais bem preparados para lidar com desafios complicados de forma crítica e criativa.

Referências

[1] L. Brando, R. M. Dutra, and E. C. Furtado, “Proposta de uma metodologia de robótica educacional como um instrumento redutor de evasão em escolas públicas,” in *Congresso sobre Tecnologias na Educação (Ctrl+e)*. SBC, 2021, pp. 21–30.

[2] ABENGE, “Artigo do congresso brasileiro de educação em engenharia,” in *Congresso Brasileiro de Educação em Engenharia (COBENGE)*. Associação Brasileira de Educação em Engenharia (ABENGE), 2026. [Online]. Available: https://www.abenge.org.br/cobenge/legado/arquivos/8/sessoestec/art1869.pdf

[3] Universidade Estadual Paulista (UNESP), “Publicação disponível no repositório institucional unesp,” 2026, acesso em: 28 abr. 2026. [Online]. Available: https://repositorio.unesp.br/entities/publication/e3495a4e-efa3-4979-a7fa-3f7ae0520271

[4] Autores do artigo, “Artigo publicado nos anais do enalic 2025,” in *Encontro Nacional das Licenciaturas (ENALIC)*. Realize Editora, 2025. [Online]. Available: https://www.editorarealize.com.br/editora/anais/enalic/2025/69bd550aa7c28_20032026110914.pdf

[5] startinatel. Materiais cupim. [Online]. Available: https://github.com/nrc-cupim/cupim

[6] RoboCore®. Regras combate cupim. [Online]. Available: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Combate+Cupim.pdf

AGRADECIMENTOS À EQUIPE START INATEL

Os autores gostariam de expressar sua gratidão a todos os membros da equipe Start Inatel, cujas contribuições foram essenciais para a realização deste artigo. O comprometimento e a dedicação foram fundamentais para aprofundar a pesquisa, possibilitando uma abordagem mais ampla e enriquecedora do tema. Sem o apoio da equipe, não seria possível alcançar os resultados apresentados neste trabalho.

AUTORES



Marco Antônio Teixeira Shima é graduando do 2º período do curso de Engenharia de Controle e Automação no Instituto Nacional de Telecomunicações (INATEL). É integrante da equipe Robotbulls, atuando nas categorias Seguidor de Linha e Perseguidor de Linha. Atualmente, é bolsista de Iniciação Científica, contribuindo para projetos na Start Inatel.



Maria Luiza Muquem Alves é graduanda do 2º período de Engenharia Elétrica no Instituto Nacional de Telecomunicações (INATEL). Membro da equipe Robotbulls, destaca-se na categoria Mini-Sumô e atua como bolsista de Iniciação Científica em projetos da Start Inatel.

Estudo de Robôs Inteligentes Aplicados ao Futebol de Robôs: Very Small Size Soccer (VSSS)

Alexandre B. Lugli, Egídio R. Neto, Eugênio M. J. da Silva, Enzo R. R. Lemos,
Gustavo Mundim Penaforte, Kaynan H. Dias de Melo, Luiz Fernando D. R. Santos, Nathan A. Costa,
Wanderson E. Saldanha
Instituto Nacional de Telecomunicações – Inatel, Brasil
baratella@inatel.br, egidio.neto@inatel.br, eugenio@inatel.br, wsaldanha@inatel.br,
enzo.rodrigues@gea.inatel.br, gustavo.penaforte@get.inatel.br, kaynan.melo@gec.inatel.br,
luizf.santos@inatel.br, nathan.costa@ges.inatel.br

***Abstract*—This paper presents a study on intelligent robots applied to the IEEE Very Small Size Soccer (VSSS) category. The work analyzes the integration of artificial intelligence techniques with robotic systems to improve decision-making, coordination, and autonomous strategies in dynamic game environments. Emphasis is given to perception, communication, and control methods that allow teams of robots to operate cooperatively with greater efficiency. The study aims to consolidate relevant approaches and share insights that can guide future research and contribute to the advancement of robotic soccer.**

***Index Terms*—Artificial Intelligence, Robotics, Multi-agent Systems, Very Small Size Soccer.**

***Resumo*—Este artigo apresenta um estudo sobre robôs inteligentes aplicados à categoria IEEE Very Small Size Soccer (VSSS). O trabalho analisa a integração de técnicas de inteligência artificial com sistemas robóticos para aprimorar a tomada de decisão, a coordenação e as estratégias autônomas em ambientes dinâmicos de jogo. Dá-se ênfase a métodos de percepção, comunicação e controle que permitem que equipes de robôs operem de forma cooperativa com maior eficiência. O estudo visa consolidar abordagens relevantes e compartilhar conhecimentos que possam orientar pesquisas futuras e contribuir para o avanço do futebol de robôs.**

***Index Terms*—Robótica, Inteligência Artificial, VSSS, Futebol de Robôs**

I. Introdução

O estudo conduzido pela equipe **RobotBulls** do Instituto Nacional de Telecomunicações (INATEL), ilustrado na Figura 1, concentra-se no desenvolvimento de robôs inteligentes voltados à competição IEEE *Very Small Size Soccer* (VSSS) [1]. Esta modalidade de futebol de robôs exige sistemas totalmente autônomos, capazes de operar em tempo real ao tomar decisões complexas a partir da percepção do ambiente. Para isso, os agentes devem identificar com precisão a posição da bola, dos adversários e de seus pares por meio de sistemas avançados de visão computacional, processando informações de maneira rápida para executar estratégias de jogo coordenadas.

O trabalho da equipe RobotBulls demonstra como a robótica inteligente é aplicada em cenários dinâmicos e competitivos, nos quais o planejamento estratégico, o controle preciso e a coordenação multiagente são determinantes para o desempenho [2]. Além de representar um desafio tecnológico, a participação na categoria VSSS oferece uma plataforma de pesquisa e ensino, permitindo a implementação de técnicas avançadas de inteligência artificial, algoritmos de tomada de decisão e interação entre múltiplos agentes.

A relevância deste estudo transcende a aplicação prática da robótica autônoma em esportes, consolidando-se como uma contribuição ao desenvolvimento acadêmico e científico [3], ao promover a integração efetiva entre fundamentação teórica, experimentação prática e inovação tecnológica.



Figura 1. Logo da equipe RobotBulls.

II. Fundamentação Teórica

A categoria IEEE Very Small Size Soccer (VSSS) consolida-se como uma importante plataforma de pesquisa em robótica móvel, visão computacional e sistemas multiagentes [1], [4]. Para compreender a aplicação das técnicas de controle propostas neste trabalho, é necessário detalhar o ambiente regulamentado da competição, bem como as limitações físicas e os desafios dinâmicos inerentes à modalidade.

A. Regras e Caracterização do Ambiente

A competição ocorre em um campo retangular de superfície preta fosca, com dimensões padronizadas (tipicamente 150 cm $\times$ 130 cm), delimitado por paredes brancas.Conforme ilustrado na Figura 2, as dimensões do campo seguem rigorosamente as especificações da competição [5]. O jogo é disputado por duas equipes de três robôs cada, com o objetivo de marcar gols utilizando uma bola de golfe laranja, cor escolhida para facilitar a segmentação por visão computacional.

Os robôs devem obedecer a estritas limitações dimensionais, devendo caber integralmente em um cubo de $7,5$ cm de aresta.

Além disso, regras específicas impedem que os robôs "segurem" a bola ou obstruam o jogo deliberadamente; o controle da posse deve ser realizado apenas através do posicionamento e da cinemática do robô. A identificação dos agentes é feita por meio de etiquetas coloridas (patchs) no topo de cada robô, permitindo que o sistema de visão determine a posição (x, y, θ) de cada agente.

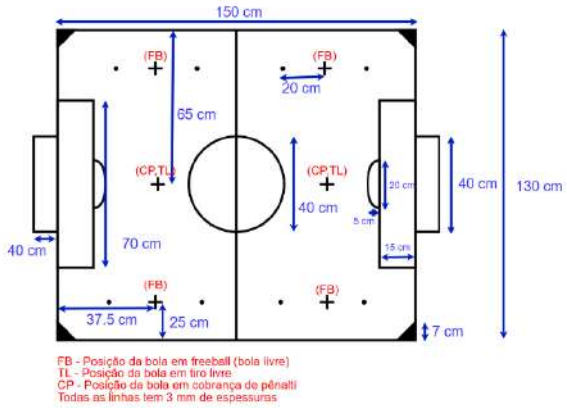


Figura 2. Dimensões oficiais do campo da categoria VSSS.

B. Arquitetura do Sistema e Latência

A arquitetura padrão do VSSS é centralizada, conforme apresentado na Figura 3. Uma câmera posicionada sobre o campo captura as imagens, que são enviadas a um computador externo. Este computador processa a visão, executa a estratégia de jogo e envia os comandos de velocidade linear e angular via radiofrequência para os robôs [6].

Essa arquitetura impõe uma limitação crítica ao projeto de controle: a latência. O tempo decorrido entre a captura da imagem, o processamento, o envio do pacote de rádio e a atuação dos motores cria um atraso (*delay*) no sistema em malha fechada [7]. Para controladores de alto ganho, esse atraso pode introduzir instabilidade, exigindo um projeto robusto de controle PD sintonizado para evitar oscilações indesejadas.



Figura 3. Arquitetura geral do sistema VSSS adotado pela equipe RobotBulls.

C. Desafios Dinâmicos e Não-Linearidades

A maioria dos robôs da categoria, incluindo os da equipe RobotBulls, utiliza acionamento diferencial (dois motores independentes). Embora mecanicamente robusto, esse sistema apresenta uma dinâmica não-holonômica — o robô possui restrições de movimento, podendo girar e deslocar-se longitudinalmente, mas não lateralmente de forma instantânea [2].

O ambiente físico introduz não-linearidades severas que dificultam a modelagem ideal:

- **Derrapagem (*Slippage*):** Devido às altas acelerações exigidas para interceptar a bola, as rodas frequentemente perdem tração.
- **Atrito Variável:** O atrito cinético entre as rodas e o tapete, e o atrito seco interno das caixas de redução, alteram a resposta do sistema.
- **Zona Morta:** Em baixas velocidades, necessárias para ajustes finos de ângulo, os motores podem não responder linearmente ao sinal PWM enviado.

Nesse cenário, a aplicação de um controle em malha aberta torna-se inviável. A implementação de controladores PD é mandatória para compensar os erros de estado estacionário e melhorar a resposta transitória, garantindo que a estratégia de alto nível seja executada com fidelidade física pelos atuadores [7].

III. Materiais e Métodos

O desenvolvimento da plataforma robótica da equipe integra mecânica de precisão, eletrônica embarcada de alto desempenho e ferramentas avançadas de simulação [6]. Esta seção detalha a arquitetura dos agentes e as tecnologias empregadas na modelagem e controle do sistema.

A. Arquitetura de Hardware e Eletrônica

Os robôs foram projetados para atender às restrições dimensionais da categoria VSSS (cubo de $7,5$ cm). A eletrônica centralizada foi desenvolvida no software *EasyEDA*, condensando os sistemas de potência e processamento em uma única placa de circuito impresso.

O processamento central é realizado pelo microcontrolador **STM32F303RE**, baseado na arquitetura Arm® Cortex®-M4 de 32 bits operando a 72 MHz, escolhido pela sua capacidade de processamento de ponto flutuante e temporizadores dedicados para controle PWM. A atuação é feita por motores **Maxon DC-MAX 22S** de 6V, acoplados a encoders **ENX10 EASY** de 1024 pulsos, garantindo alta resolução para a malha de controle [7]. O acionamento dos motores utiliza drivers ponte-H **DRV-8231** com regulação de corrente integrada.

A comunicação com o computador da estratégia é feita via módulos de radiofrequência **NRF24L01** (2.4 GHz) utilizando protocolo SPI, o que assegura baixa latência no envio dos setpoints de velocidade.

B. Projeto Mecânico e Manufatura

A estrutura mecânica foi modelada no software *SolidWorks*, priorizando o rebaixamento do centro de gravidade para aumentar a estabilidade. O chassi é manufaturado através de impressão 3D em material **PETG**, selecionado por sua resistência a impactos. Para maximizar a tração e mitigar a derrapagem — um desafio crítico no controle dinâmico — as rodas são revestidas com poliuretano líquido (*Poly 74-20*) e o chassi incorpora uma base de aço de 1 mm para aumentar a força normal.

C. Ferramentas de Simulação e Validação

Para validar as estratégias de jogo e a lógica de movimentação antes da implementação física, utilizou-se o simulador **FIRASim**. Esta ferramenta utiliza a biblioteca ODE (*Open Dynamics Engine*) para simular a física de corpos rígidos, permitindo testes seguros de colisão e trajetória [2].

IV. Inteligência Artificial

A Inteligência Artificial (IA) desempenha um papel central na arquitetura do sistema desenvolvido pela equipe RobotBulls, atuando desde a percepção até a execução das ações de campo [3]. O sistema integra visão computacional, tomada de decisão baseada em máquina de estados e navegação avançada utilizando Unit Vector Field (UVF), além de controle em múltiplas camadas. O fluxo geral da IA empregada pode ser observado na Figura 4.

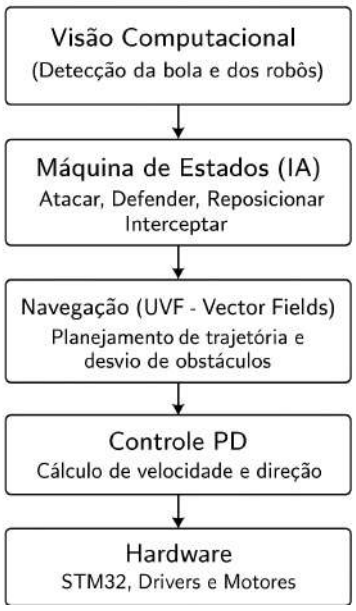


Figura 4. Fluxo geral da Inteligência Artificial utilizada no sistema VSSS, incluindo percepção, tomada de decisão e controle.

A. Tomada de Decisão Estratégica com Máquina de Estados

A equipe implementa uma máquina de estados que atua como um "técnico virtual", definindo o comportamento e a função de cada robô durante a partida [3]. A posição da bola é utilizada como referência principal para determinar a estratégia (ofensiva ou defensiva) e o papel de cada agente. O sistema analisa em tempo real o comportamento do time adversário – incluindo posicionamento e número de atacantes – para antecipar jogadas e evitar surpresas, permitindo uma defesa mais eficaz e uma tomada de decisão assertiva.

B. Navegação Avançada com Unit Vector Field (UVF)

Para o deslocamento preciso e estratégico, é utilizado o conceito de Unit Vector Field (UVF) para calcular e traçar as trajetórias dos robôs [6]. Esse método é combinado com um sistema de controle PD em duas camadas:

- **Controle por software:** Ajusta a movimentação com base na diferença angular do robô.
- **Controle por firmware:** Monitora constantemente a leitura dos encoders para um controle preciso da velocidade e rotação dos motores.

Essa integração permite movimentos fluidos, desvio eficiente de obstáculos e alta precisão na navegação.

C. Visão Computacional e Processamento de Imagens

O sistema de visão computacional emprega técnicas de IA para identificar, localizar e processar as imagens capturadas pelas câmeras [4]. O processo inclui:

- Transformação linear 2D para converter a imagem capturada e remover distorções da lente.
- Filtragem de ruídos e conversão do espaço de cores RGB para HSV, facilitando a segmentação e identificação dos robôs e da bola.
- Um sistema de calibração dinâmica de cores, por meio de uma interface dedicada, que ajusta as cores em função da iluminação ambiente, garantindo robustez e evitando falhas de identificação durante a partida.

D. Simulação como Ferramenta de Desenvolvimento da IA

A equipe utiliza o simulador FIRASim para testar e aprimorar o desempenho da IA em um ambiente virtual que replica as condições reais do jogo [3]. Isso permite desenvolver e avaliar diferentes estratégias antes da implementação no hardware físico, realizar testes extensivos do código de controle de forma isolada, reduzindo custos e riscos, além de acelerar o ciclo de desenvolvimento e refino das táticas da equipe.

E. Coordenação e Reposicionamento Autônomo

A IA também gerencia a coordenação entre os robôs e operações específicas como o *reposicionamento automático*, exigido pelas regras da competição após faltas, gols ou pênaltis [5]. O sistema altera o estado do robô para um modo autônomo de recolocação, utilizando dados de posicionamento fornecidos pela visão computacional. Uma interface de reposicionamento manual também está disponível para ajustes e testes.

F. Perspectivas Futuras

A base integrada de IA, visão, navegação e controle estabelecida pelo projeto abre caminho para futuros avanços, como:

- Aplicação de técnicas de aprendizado de máquina para otimização ainda mais fina das trajetórias e decisões táticas [8].
- Implementação de algoritmos de previsão de movimento para antecipar ações dos adversários.
- Exploração de métodos de IA para a geração automática de estratégias inovadoras.

V. Testes e Resultados

A validação do sistema proposto foi realizada em três frentes complementares: (1) testes experimentais de bancada para sintonia do controle PD; (2) avaliação qualitativa e comparativa do sistema de IA em ambiente simulado; e (3) comprovação empírica do desempenho integrado durante o calendário competitivo de 2025 [6]. Esta abordagem permitiu isolar o impacto individual de cada componente (controle, navegação, decisão) antes de validar o sistema completo em cenários reais.

A. Resultados Experimentais e Sintonia PD

A aplicação do método do Lugar Geométrico das Raízes (LGR) permitiu definir os parâmetros do controlador PD visando um critério de desempenho rigoroso: tempo de acomodação (t_s) de 40 ms e um máximo pico (sobressinal) limitado a 0,1 [7].

Utilizando a aquisição de dados via encoders, validou-se que a estrutura PD oferece a resposta dinâmica ideal para a categoria VSSS [6]. A ação **Proporcional** (K_p) garantiu a agressividade necessária para que o robô atingisse rapidamente a velocidade de setpoint, fundamental para interceptações. Simultaneamente, a ação **Derivativa** (K_d) atuou de forma preditiva, amortecendo as oscilações e impedindo o *overshoot* (ultrapassagem do alvo) durante frenagens bruscas ou mudanças de direção.

Essa configuração mostrou-se superior ao controle puramente proporcional, permitindo ganhos mais elevados sem comprometer a estabilidade. O erro de regime permanente foi reduzido significativamente, resultando em uma movimentação precisa e responsiva, essencial para o sistema de navegação por Vetores Unitários (UVF).

B. Avaliação Qualitativa do Sistema de IA

Para avaliar o impacto da inteligência artificial no desempenho do time, foram realizados testes comparativos no simulador FIRASim, contrastando a configuração completa (IA + UVF + PD) contra uma linha de base reativa simples (controle proporcional sem tomada de decisão estratégica).

1) Melhoria na Interceptação e Navegação: Observou-se uma melhoria visível na taxa de interceptação da bola quando utilizando o sistema UVF aliado à máquina de estados [6]. Enquanto a abordagem reativa apresentava dificuldades em cenários de disputa direta, o sistema com IA demonstrou maior assertividade ao calcular trajetórias ótimas e antecipar movimentos adversários. O erro de trajetória, avaliado qualitativamente através da inspeção visual das rotas percorridas, mostrou-se consideravelmente menor com a implementação do UVF, resultando em movimentos mais suaves e precisos.

2) Eficácia da Tomada de Decisão: O sistema de máquina de estados mostrou-se eficaz na seleção apropriada de comportamentos durante as partidas simuladas [3]. As transições entre estados (ex.: DEFESA → ATAQUE) ocorriam de forma coerente com o contexto do jogo, respondendo adequadamente à movimentação da bola e à formação adversária. Notou-se também uma redução significativa em colisões entre robôs da mesma equipe, graças ao estado de MEIO-CAMPO/TRANSIÇÃO que coordena o reposicionamento estratégico.

3) Integração e Comportamento Emergente: A combinação da IA com o controle PD permitiu a emergência de comportamentos táticos complexos, como:

- **Cobertura defensiva inteligente:** Os robôs se redistribuíam automaticamente para cobrir espaços vazios quando um companheiro se deslocava para o ataque.
- **Transições ofensivas fluidas:** A passagem de um sistema defensivo organizado para um ataque rápido ocorria de forma coordenada, com múltiplos robôs assumindo papéis complementares.
- **Adaptação contextual:** Em situações de superioridade numérica (ex.: 3v2), o sistema priorizava movimentos de cerco e pressão, enquanto em desvantagem adotava formação mais conservadora.

C. Desempenho em Competições (Temporada 2025)



Figura 5. Robôs da equipe RobotBulls durante partida oficial. A estabilidade dinâmica permitida pelo PD sintonizado e a tomada de decisão rápida da IA foram fatores decisivos nas conquistas de 2025.

A robustez do sistema integrado (IA + UVF + PD) foi comprovada durante as competições oficiais de 2025. Em situações reais de jogo, onde fatores como latência de comunicação e atrito variável estão sempre presentes, a combinação do controle PD estabilizado com a tomada de decisão estratégica da IA mostrou-se decisiva para o desempenho competitivo.

Como resultado desta abordagem integrada, a equipe RobotBulls obteve uma série histórica de conquistas no cenário nacional:

- **Campeão da IRONCUP 2025:** A precisão do controle PD foi fundamental nas disputas de bola, enquanto o posicionamento inteligente dos robôs (gerenciado pela IA) resultou em uma das defesas mais sólidas da competição.
- **Campeão do RSM Challenge 2025:** A estratégia ofensiva, caracterizada por movimentos rápidos e trajetórias otimizadas, demonstrou eficácia superior contra defesas organizadas.
- **Vice-Campeão (Runner-Up) da RCX - CPBR16 2025:** O desempenho consistente em todas as fases da competi-

ção confirmou a competitividade do sistema em alto nível técnico.

D. Qualificação Internacional e Observações de Campo

O desempenho consistente ao longo da temporada garantiu à equipe a qualificação para a **LARC (Latin American Robotics Competition)**, sediada em Monterrey, no México [5]. A análise das partidas oficiais permitiu observar que a máquina de estados demonstrou alta confiabilidade na seleção de ações defensivas, o sistema de reposicionamento automático operou com precisão cumprindo as regras sem infrações, e a integração entre os módulos de visão, IA e controle mostrou-se robusta mesmo em momentos de alta dinâmica do jogo [6].

Esta trajetória competitiva bem-sucedida consolida a eficácia da arquitetura proposta. A capacidade dos robôs de operarem de forma coordenada, com estabilidade de movimento e decisão contextual, confirma que a combinação sinérgica de técnicas de IA com controle clássico representa uma abordagem promissora para a categoria VSSS [1].

VI. Conclusão

Este trabalho apresentou o desenvolvimento e a implementação de robôs inteligentes aplicados à categoria IEEE Very Small Size Soccer (VSSS), destacando a integração de técnicas de inteligência artificial com sistemas robóticos para controle, percepção e coordenação multiagente [1]. A análise da arquitetura de hardware, eletrônica embarcada e estratégias de controle PD demonstrou que é possível alcançar elevada precisão na movimentação e estabilidade dinâmica, mesmo diante de desafios como latência, atrito variável e derrapagem [7].

Os testes experimentais e o desempenho em competições oficiais de 2025 evidenciam a eficácia do método proposto [6], confirmando que a sintonia adequada dos controladores e a modelagem precisa do sistema físico são fundamentais para a execução de estratégias autônomas. Além disso, os resultados obtidos reforçam a relevância do VSSS como plataforma de pesquisa em robótica móvel e sistemas multiagentes [4], oferecendo insights aplicáveis a outros contextos de robótica autônoma e cooperação entre agentes.

Como perspectiva futura, destaca-se a exploração de técnicas avançadas de inteligência artificial, incluindo aprendizado de máquina [8] e planejamento preditivo, visando otimizar ainda mais a tomada de decisão dos robôs em tempo real e ampliar a capacidade de adaptação a diferentes cenários de jogo.

Referências

[1] H. Kitano et al., "RoboCup: A challenge problem for AI and robotics," in *Proc. RoboCup*, 1997.
[2] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.
[3] A. B. Lugli et al., "Sistema de Controle para Robôs aplicados no IEEE Very Small Size Soccer (VSSS) Baseado em Inteligência Artificial," in *Anais do Congresso de Iniciação Científica do Inatel (INCITEL)*, Santa Rita do Sapucaí, Brasil, 2024.
[4] B. Siciliano and O. Khatib, *Springer Handbook of Robotics*. Springer, 2016.
[5] LARC, "LARC 2025 Very Small Size Soccer (VSSS) Rules," 2025. [Online]. Disponível em: https://static1.squarespace.com/static/68811cea22a48812262c6b93/t/689f9746e33f0a416adfc808/1755289414689/LARC-VSSS-2025-PR.pdf. [Acesso em: 09/12/2025].
[6] A. B. Lugli et al., "Estudo do Sistema de Controle PID para Robôs aplicados no IEEE Very Small Size Soccer (VSSS)," in *Anais do Congresso de Iniciação Científica do Inatel (INCITEL)*, Santa Rita do Sapucaí, Brasil, 2025.
[7] K. Ogata, *Modern Control Engineering*, 5th ed. Prentice Hall, 2010.
[8] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press, 2018.

Autores



Enzo Rodrigues Ribeiro Lemos é graduando no curso de Engenharia de Controle e Automação pelo INATEL (Instituto Nacional de Telecomunicações). Foi voluntário na Cas@Viva, na reciclagem de lixo eletrônico. Ele também é membro ativo da equipe de robótica Robotbulls, participando da categoria VSSS (Very Small Size Soccer).



Gustavo Mundin Penaforte é aluno de graduação do curso de Engenharia de Telecomunicações do INATEL (Instituto Nacional de Telecomunicações). E-mail: gustavo.penaforte@get.inatel.br. Nascido em Divinópolis (MG), atual membro da Robotbulls. Tem interesse nas áreas de Tecnologia.



Kaynan Henrique Dias de Melo é graduando no curso de Engenharia de Computação pelo INATEL (Instituto Nacional de Telecomunicações) e membro ativo da equipe de robótica Robotbulls, representando na categoria Very Small Size Soccer (VSSS) e competindo pela categoria Mini Sumo.



Nathan Arruola da Costa é graduando do curso de Engenharia de Software pelo INATEL (Instituto Nacional de Telecomunicações). Atuou como voluntário na Cas@Viva, contribuindo para o processo de reciclagem de lixo eletrônico. Além disso, é membro ativo da equipe de robótica Robotbulls, participando na categoria VSSS (Very Small Size Soccer).

Estudo de Robôs Inteligentes Aplicados ao Combate de Sumô LEGO®

Alexandre Baratella Lugli, Eugênio Moreira Jonas da Silva,
Egídio Raimundo Neto, Luís Fernando Domiciano Ribeiro dos Santos,
Wanderson Eleutério Saldanha, Luiz Gustavo de Oliveira Costa, Rodrigo G. S. Santos

Instituto Nacional de Telecomunicações - Inatel
baratella@inatel.br, eugenio@inatel.br, egidioneto@inatel.br, luizf.santos@inatel.br,
wsaldanha@inatel.br, luiz.g@gec.inatel.br, rodrigo.geraldo@gea.inatel.br

***Resumo*—O Instituto NacionaL de Telecomunicações (INATEL), reconhecido por sua excelência em ensino, pesquisa e desenvolvimento tecnológico, incentiva a capacitação de novos profissionais por meio de projetos que integram conhecimento teórico e aplicação prática. Entre as iniciativas conduzidas por estudantes e docentes do laboratório de Engenharia Elétrica e Engenharia de Controle e Automação, conhecido como "Robotbulls", destacam-se os robôs da modalidade Sumô LEGO®.**

Este artigo tem como propósito apresentar uma Análise de Técnicas de Programação para Robôs da Categoria Sumô LEGO®, com ênfase nos controladores NXT e EV3. O estudo discute métodos de programação baseados em sensores, algoritmos de controle e estratégias de adaptação comportamental, examinando as particularidades e capacidades de cada plataforma.

Além de buscar melhorias no desempenho dos robôs em competições, este trabalho também visa fortalecer a formação tecnológica dos estudantes do INATEL, reforçando o papel da instituição como referência acadêmica no campo da robótica.

***Palavras chave*—LEGO®, Programação, Sumô, Controladores, NXT, EV3.**

***Abstract*—The National Institute of Telecommunications (INATEL), renowned for its excellence in education, research, and technological innovation, fosters the development of future professionals through initiatives that integrate theoretical knowledge with practical application. Among the projects carried out by students and professors in the Electrical Engineering and Control and Automation Engineering laboratory, known as "Robotbulls," the LEGO® Sumo robot category stands out.**

This article aims to present an Analysis of Programming Techniques for LEGO® Sumo Robots, with emphasis on the NXT and EV3 controllers. The study explores programming approaches based on sensors, control algorithms, and behavioral adaptation strategies, examining the potential and specific features of each platform.

Beyond seeking to enhance robot performance in competitions, this work also contributes to the technological training of INATEL students, reinforcing the institution's role as an academic reference in the field of robotics.

***Index Terms*—LEGO®, Programming, Sumo, Controllers, NXT, EV3.**

I. Introdução

Atualmente, a LEGO® é reconhecida não apenas por seus tradicionais blocos de montagem, mas também por sua relevância no campo da robótica, especialmente em ambientes competitivos.

Nos últimos anos, diversos estudos têm ressaltado o papel da robótica educacional associada à competição como instrumento de desenvolvimento técnico e cognitivo. Pesquisas recentes, como a de Maurer (2022), mostram que plataformas como o LEGO Spike Prime têm sido cada vez mais utilizadas em torneios devido à sua versatilidade e ao suporte a linguagens modernas, como MicroPython[1]. Além disso, iniciativas internacionais, como a RoboCup Junior[2] e o All Japan Robot-Sumo Tournament[3], reforçam a importância global dessa área, evidenciando que os kits LEGO continuam sendo fundamentais para o desenvolvimento de estratégias de inteligência robótica em cenários restritos e altamente competitivos.

Dentro desse contexto, a categoria Sumô LEGO® se destaca por exigir que os robôs sejam construídos exclusivamente com peças LEGO®, o que estimula habilidades relacionadas à programação, engenharia e resolução de problemas. Este artigo analisa as principais técnicas de programação aplicadas aos robôs da categoria Sumô LEGO®, com foco nos controladores NXT e EV3, buscando compreender de que forma essas plataformas contribuem para melhorar o desempenho dos robôs em disputas.

O robô da categoria LEGO® depende, além de sua estrutura física, de um componente essencial: seus controladores. Entre os modelos disponíveis, destacam-se o NXT e o EV3, cada um oferecendo diferentes capacidades e recursos. O NXT foi um dos primeiros sistemas capazes de viabilizar robôs autônomos, trazendo conectividade, motores e sensores. O EV3, por sua vez, representa uma evolução dessa tecnologia, oferecendo maior poder de processamento, melhor conectividade e uma variedade ampliada de sensores.

A modalidade consiste em fazer com que um robô tente empurrar o oponente para fora de um Dojo — uma arena circular feita em MDF e seguindo dimensões padronizadas — de forma semelhante ao sumô tradicional. A diferença é que, nessa categoria, os competidores são robôs montados exclusivamente com peças LEGO®.

II. Fundamentação Teórica

A. Regras

A categoria Sumô LEGO® segue um conjunto de normas definidas pela RoboCore® [4]. Cada duelo é disputado entre duas equipes, formadas por até dois integrantes presentes na arena, sendo um responsável pela operação do robô (piloto) e outro como assistente. Apenas o piloto pode manipular o robô

durante cada round. Uma partida é composta por três rounds e, caso ocorra empate, um quarto round é realizado para definir o vencedor.

A categoria estabelece limites rigorosos relacionados à construção dos robôs, incluindo tipo de material, peso e dimensões, que são avaliados antes de cada combate. Um gabarito é utilizado para verificar se o robô se mantém dentro das dimensões máximas permitidas (15,2 cm de largura e comprimento), enquanto uma balança de precisão confirma se o peso não ultrapassa 1 kg. Essas limitações visam equilibrar o nível da competição, fazendo com que a estratégia de montagem e a programação desempenhem papéis decisivos no desempenho de cada robô.

O Dojô [1], onde ocorrem as lutas, é confeccionado em madeira laminada de alta pressão, possui formato circular e diâmetro de 77 cm. Sua borda é marcada por uma faixa branca de 2,5 cm de largura. No centro da arena, há duas linhas marrons chamadas Shikiri, cada uma com 1 cm de largura e 10 cm de comprimento, separadas por 10 cm de distância.



Fig. 1. Dojô.

Antes do início da luta, as equipes posicionam seus robôs de acordo com suas estratégias após uma contagem regressiva de três segundos. Esse posicionamento deve ocorrer atrás das linhas Shikiri; se isso não for respeitado, a equipe recebe uma advertência. Acumular duas advertências resulta na desclassificação da equipe no round em andamento. Após o posicionamento válido e a confirmação do juiz, uma nova contagem regressiva marca o início da partida. O piloto ativa o robô e uma espera obrigatória de cinco segundos deve ser respeitada antes de qualquer movimento; o descumprimento também gera advertência.

A luta termina quando um dos robôs é empurrado para fora do Dojô. O robô que permanecer recebe um ponto. A equipe que atingir dois pontos primeiro dentro de até três rounds vence a partida. Em caso de igualdade, o juiz realiza um round extra, no qual ele próprio posiciona os robôs, definindo suas posições iniciais. Após isso, o combate segue normalmente, e o vencedor desse round leva a vitória final.

B. Programação

Devido às restrições físicas impostas pela categoria, a programação se torna um elemento essencial para o bom desempenho do robô. É por meio dela que as equipes definem o comportamento da máquina, considerando tanto suas próprias ações quanto possíveis movimentos do adversário. Cada equipe cria a lógica que considera mais eficaz para seus robôs, levando em conta fatores como o estilo de combate, o tipo de controlador utilizado (NXT ou EV3) e os sensores disponíveis, buscando otimizar o tempo de resposta e a eficiência durante o confronto.

C. Controladores

Os controladores são componentes fundamentais para o funcionamento dos robôs, atuando como a unidade responsável por processar informações e coordenar ações. Eles recebem dados dos sensores, interpretam essas informações e enviam comandos aos motores e outros elementos. Em competições como o Sumô LEGO®, os controladores mais utilizados são o NXT e o EV3, ambos conhecidos pela facilidade de programação e flexibilidade na integração de sensores e motores.

Controladores Disponíveis

* LEGO® NXT [5]

O NXT faz parte da linha LEGO Mindstorms e é amplamente adotado em projetos educacionais e competitivos. Lançado em 2006, contém um microcontrolador ARM7 de 32 bits operando a 48 MHz, três portas para motores e quatro para sensores. Possui 64 KB de RAM, 256 KB de memória flash, e conectividade via USB e Bluetooth.

Sua programação pode ser realizada tanto por meio de interfaces visuais (como o NXT-G) quanto por linguagens textuais, como NXC (Not eXactly C). Pela combinação de simplicidade e capacidade, o NXT continua sendo uma opção popular entre iniciantes e equipes experientes.



Fig. 2. Controlador NXT.

* LEGO® EV3 [6]

Sucessor direto do NXT, o EV3 foi lançado em 2013 com melhorias significativas. Ele utiliza um processador ARM9 de 300 MHz, oferecendo maior desempenho para processamento de dados e tomadas de decisão mais complexas. Conta com 64 MB de RAM, 16 MB de memória flash, Wi-Fi, USB e Bluetooth, e permite conexão simultânea de 4 motores e 4 sensores.

Sua programação pode ser feita pela interface visual do LEGO Mindstorms EV3 ou por linguagens como Python, C e Java, tornando-o uma plataforma versátil e adequada tanto para iniciantes quanto para programadores avançados.



Fig. 3. Controlador EV3.

* LEGO® Spike Prime [7]

O Spike Prime é uma plataforma moderna voltada principalmente ao ensino de robótica e programação. Seu hub possui processador de 100 MHz, 320 KB de RAM, 1 MB de memória flash e armazenamento adicional de 32 MB, além de conectividade Bluetooth e uma matriz de LEDs 5×5. Ele suporta programação em MicroPython, o que o torna bastante flexível.

Apesar disso, sua presença na categoria Sumô LEGO® ainda é limitada. Isso ocorre porque o Spike Prime é relativamente novo, possui custo mais elevado e é menos tradicional em competições desse tipo quando comparado ao NXT e EV3, que já possuem histórico consolidado e maior aceitação entre as equipes.



Fig. 4. Controlador Spike Prime.

III. Metodologia

A. Sensores

Os sensores desempenham um papel fundamental na robótica, especialmente em competições como o Sumô LEGO, onde a precisão e a rápida adaptação às condições do ambiente são determinantes para o desempenho do robô. Nos controladores LEGO, como o NXT e o EV3, é possível integrar diversos sensores que ampliam a capacidade de percepção e interação do robô com seu entorno.

- Sensor Ultrassônico: Responsável por medir a distância entre o robô e objetos ao seu redor, sendo essencial para identificar obstáculos ou adversários e auxiliar no posicionamento. Sua boa precisão e rápida capacidade de resposta fazem dele um componente indispensável para maximizar o desempenho em competições, permitindo tomadas de decisão imediatas.
- Sensor de Toque: Detecta o contato direto do robô com um objeto, sendo um recurso importante para estratégias específicas durante o combate.

Quando utilizados em conjunto com uma programação bem estruturada, esses sensores ampliam a capacidade do robô de reagir ao ambiente, garantindo a agilidade e eficiência necessárias para enfrentar os desafios da arena de combate.

B. Estilos de programação

b.1) Programação com Blocos

O LEGO Mindstorms[8] é uma plataforma de robótica educacional criada para facilitar o aprendizado de programação e conceitos fundamentais de engenharia. Uma de suas principais características é o uso de uma interface de programação baseada em blocos, disponível em softwares como o LEGO Mindstorms EV3 Software ou o SPIKE App, dependendo do kit utilizado.

Essa abordagem visual permite desenvolver programas por meio da organização de blocos que representam ações específicas, como controle de motores, leitura de sensores, utilização de loops e estruturas condicionais. Por dispensar o conhecimento prévio em linguagens textuais, essa interface promove um ambiente acessível, interativo e intuitivo, especialmente para iniciantes.

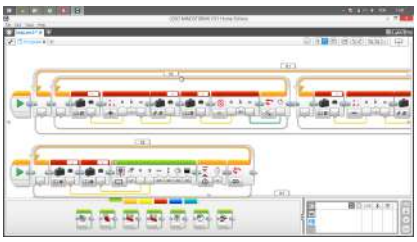


Fig. 5. Exemplo de programação em blocos no software LEGO Mindstorms EV3.

b.2) Linguagens de programação

Os controladores NXT e EV3, da linha LEGO Mindstorms, oferecem compatibilidade com diversas linguagens de programação, abrangendo desde opções introdutórias até ferramentas avançadas. O NXT pode ser programado utilizando NXT-G, uma interface visual voltada para iniciantes; NXC (Not eXactly C), baseada em C, que proporciona maior controle; LeJOS, que permite programar em Java; e RobotC, amplamente empregado em competições devido à sua flexibilidade e precisão.

O EV3, por sua vez, apresenta maior capacidade de processamento e suporta linguagens como EV3-G, sua interface visual nativa; Python, oficialmente integrado para programação de alto nível; RobotC; Java via LeJOS; e até C/C++ e MATLAB/Simulink, ampliando as possibilidades de controle, análise e simulação. Assim, ambos os controladores são bastante versáteis, embora o EV3 seja mais indicado para projetos complexos devido ao seu desempenho superior.

Entre as opções disponíveis, o RobotC se destaca como uma das linguagens mais utilizadas em competições de robótica. Sua eficiência, flexibilidade e excelente controle sobre sensores e motores o tornam uma escolha popular. Baseado em C, permite o desenvolvimento de algoritmos avançados e multitarefas, essenciais para estratégias dinâmicas em competições como o Sumô LEGO. Além disso, seu ambiente de desenvolvimento conta com ferramentas de depuração e simulação, facilitando a criação e otimização de programas, garantindo respostas

rápidas e precisas — características indispensáveis em disputas de alta performance.

C. Biblioteca

Na linguagem visual EV3-G, utilizada no LEGO Mindstorms EV3, não é possível criar bibliotecas no sentido tradicional, como ocorre em linguagens textuais. Contudo, o ambiente permite a criação de blocos personalizados, que funcionam como módulos reutilizáveis que agrupam funções ou sequências de comandos. Esses blocos ajudam a tornar o código mais organizado e modular, ainda que com limitações em comparação às bibliotecas tradicionais.

Já no RobotC, embora não haja suporte direto para bibliotecas no formato clássico (.h ou .c), é possível estruturar o código por meio de funções e procedimentos. Essas funções permitem encapsular trechos reutilizáveis, resultando em uma modularização parcial. Entretanto, o código geralmente precisa permanecer em um único arquivo, o que restringe a modularização completa oferecida em linguagens como C e C++. Assim, apesar de permitir alguma organização interna, o RobotC não fornece um sistema nativo robusto para criação e gerenciamento de bibliotecas externas, reduzindo a flexibilidade da reutilização de código.

IV. Estudo

A. Detectando o adversário

A detecção precisa da posição do adversário é essencial durante a luta, pois permite que o robô execute ataques estratégicos e se movimente de forma inteligente pelo dojô. Para isso, utiliza-se uma combinação entre sensores ultrassônicos, sensores de toque e uma programação bem estruturada, garantindo navegação controlada e eficiente.

A maior parte dos robôs emprega dois sensores ultrassônicos, cuja leitura conjunta permite ajustar com precisão o controle dos motores, resultando em movimentações estratégicas mais eficazes durante a competição. A seguir, apresenta-se um exemplo de abordagem utilizada para detectar o adversário e controlar os motores de maneira eficiente.



Fig. 6. Exemplo de código em RobotC para o robô adversário.

B. Jogadas

Uma das estratégias mais comuns no Sumô LEGO é o empurrão direto, em que o robô avança em linha reta com o objetivo de empurrar o oponente para fora do dojô. Essa técnica é típica de robôs do tipo aríete, projetados para aplicar grande força de impacto — muitas vezes com a intenção de tombar o adversário. Robôs mais robustos utilizam essa abordagem para exercer pressão contínua, dificultando que o oponente consiga se reposicionar ou reagir. Os algoritmos desse tipo de robô priorizam o deslocamento retilíneo, garantindo maior eficiência no ataque frontal.

Em contraste, robôs com outras configurações — como modelos rápidos ou com design em "asa" — tendem a adotar táticas mais dinâmicas. Esses robôs frequentemente utilizam movimentos curvos ou circulares para evitar o confronto frontal direto, buscando atacar pelas laterais. Esse tipo de estratégia permite explorar ângulos favoráveis e aproveitar brechas deixadas pelo adversário, aumentando as chances de um empurrão bem-sucedido.

Pilotagem:

O sistema proposto inclui um módulo de pilotagem configurável, permitindo ao operador selecionar, antes do início da partida, tanto o tipo de curva quanto sua direção. Essa interface de escolha utiliza os botões do NXT e oferece três opções de curvatura: fechada, média e aberta, cada uma associada a relações específicas de velocidade entre os motores e a tempos definidos de execução.

A direção da curva inicial pode ser ajustada independentemente, permitindo escolher entre direita, esquerda ou simplesmente optar por um deslocamento retilíneo. Essa manobra é aplicada exclusivamente no início da partida, com o objetivo de reposicionar o robô de maneira controlada e estratégica. Assim, o operador pode configurar o posicionamento inicial conforme o contexto da competição, sem necessidade de reprogramar o sistema.

Costas:

A estratégia chamada "costas" consiste em um recuo inicial rápido, seguido de um avanço frontal. Quando selecionada, o robô realiza uma marcha reversa de um segundo em potência máxima e sem feedback sensorial. Em seguida, avança em linha reta por 0,75 segundo, também sem controle por sensores.

Essa movimentação predefinida altera deliberadamente a posição inicial do robô no dojô, ajustando seu alinhamento e a distância em relação ao adversário. Após essa etapa, o sistema passa para o modo reativo, baseado em sensores ultrassônicos.

Diferentemente da pilotagem configurável, essa estratégia não utiliza parâmetros ajustáveis: sua execução é determinística, com tempos e velocidades fixas, priorizando simplicidade e imprevisibilidade no início do combate.

Desvia:

A estratégia "desvia" utiliza uma curva inicial assimétrica configurada previamente pelo operador. Após um breve recuo automático para criar espaço de manobra, o robô executa uma curva cuja intensidade e duração dependem dos parâmetros de pilotagem — que definem tanto o diferencial de velocidade entre os motores quanto o tempo total da manobra. A direção da curva (direita ou esquerda) também é escolhida antes da luta.

Ao finalizar essa etapa inicial, o robô entra em modo reativo, utilizando sensores ultrassônicos frontais para localizar o adversário. Se ambos detectam proximidade, o robô avança rapidamente; se apenas um sensor detecta, o robô gira para alinhar-se; caso nenhum identifique o inimigo, um movimento exploratório alternado é executado.

Diferentemente das trajetórias lineares fixas, o "desvia" combina deslocamento angular planejado com rastreamento autônomo posterior, tornando a estratégia mais adaptável e

menos previsível, sem acrescentar grande complexidade ao sistema.

V. Comparativo

Os controladores NXT, EV3 e Spike apresentam características distintas, adequadas para diferentes contextos de uso. O NXT é uma opção ideal para iniciantes, oferecendo simplicidade e facilidade de uso, embora limitado em processamento e capacidade de expansão. O EV3, mais avançado, possui maior poder de processamento, mais memória e suporte a diversas linguagens de programação, o que o torna a escolha mais adequada para competições e projetos complexos.

Já o Spike, apesar de oferecer recursos modernos e boa flexibilidade, ainda não possui ampla adoção, o que reduz seu uso em competições como o Sumô LEGO. Assim, a decisão entre esses controladores depende da experiência da equipe, da complexidade pretendida e das necessidades de conectividade.

De modo geral, o EV3 costuma ser a opção mais robusta para desafios exigentes, o NXT atende bem a quem está começando, e o Spike, embora promissor, ainda enfrenta limitações devido à sua adoção reduzida.

VI. Conclusão

A programação exerce um papel decisivo no desempenho de um robô de Sumô LEGO®, constituindo o principal fator que diferencia uma equipe vencedora das demais. Por meio de algoritmos bem estruturados, o robô consegue interpretar informações dos sensores, tomar decisões rápidas e executar estratégias eficazes dentro do dojô. Embora uma montagem resistente e bem planejada seja fundamental, é a programação que possibilita movimentos precisos, detecção eficiente do oponente e adaptação às condições dinâmicas da luta. Dessa forma, uma lógica programada de maneira inteligente não apenas melhora o rendimento do robô, como também evidencia a capacidade estratégica e inovadora da equipe, fortalecendo sua vantagem competitiva.

A. Resultados Finais

A equipe Robotbulls tem se destacado nas competições de robótica no Brasil graças ao constante aperfeiçoamento da programação de seus robôs. O uso eficiente dos controladores NXT e EV3 possibilita um controle preciso dos motores e sensores, garantindo movimentações rápidas, consistentes e orientadas à estratégia. A programação avançada é o elemento-chave que mantém a equipe entre as melhores, permitindo respostas ágeis a diferentes cenários enfrentados no dojô.

Ao longo de sua trajetória, a equipe Robotbulls acumulou diversas conquistas, destacando-se em competições como:

- 1º lugar na Iron Cup 2020;
- 1º e 3º lugar na Iron Cup 2022;
- 2º e 3º lugar na Iron Cup 2023;
- 1º, 2º e 3º lugar na Iron Cup 2024;
- 3º lugar na RCX 2024;
- 3º lugar na Copa Turing 2018.

Referências

[1] Aaron Maurer. *Design Innovative Robots with LEGO SPIKE Prime by Aaron Maurer*. 2022. URL: https://www.oreilly.com/library/view/design-innovative-robots/9781801811576/.

[2] RobocupJunior. *Soccer, Rescue, and OnStage Rules 2023*. 2023. URL: https://junior.robocup.org/.

[3] FUJISOFT Inc. *All Japan Robot-Sumo Tournament*. 2024. URL: https://www.fsi.co.jp/robot/.

[4] RoboCore®. *Regras SUMÔ*. URL: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Sumo.pdf (acesso em 27/11/2024).

[5] *LEGO Mindstorms NXT*. 2024. URL: https://en.wikipedia.org/wiki/Lego_Mindstorms_NXT.

[6] *LEGO Mindstorms EV3*. 2024. URL: https://education.lego.com/en-us/products/lego-mindstorms-education-ev3-core-set-/5003400.

[7] *LEGO Education Spike Prime*. 2024. URL: https://spike.legoeducation.com/prime/.

[8] FLLCasts. *EV3 Phi. Starting EV3-G software. What are projects and programs?* 2024. URL: https://www.fllcasts.com/tutorials/364-ev3-phi-starting-ev3-g-software-what-are-projects-and-programs.

Autores



Luiz Gustavo de Oliveira Costa é graduando período do curso de Engenharia de Computac Instituto Nacional de Telecomunicaç ões –Além disso, atua como bolsista de iniciaç equipe de robótica Robotbulls, integrante Sumô LEGO®, contribuindo diretamente desenvolvimento de robôs.



Rodrigo Geraldo Silva Santos é graduando período do curso de Engenharia de Controle Automação no Instituto Nacional de Telecomunicac INATEL. Além disso, atua como bolsista de científica na equipe de rob ótica Robotbulls, categoria Sumô LEGO®, contribuindo diretamente no desenvolvimento de robôs.

Estudo para Desenvolvimento de um Robô de Combate para Categoria Antweight - 454g (1lb)

Alexandre B. Lugli; Egídio R. Neto; Eugenio M. J. da Silva; Felipe S. Barbosa; Téo S. Marques; Wanderson E. Saldanha Luiz Fernando D. R. Santos; Wagner L. C. Anastácio
Instituto Nacional de Telecomunicações - Inatel
baratella@inatel.br; egidioneto@inatel.br; eugenio@inatel.br; felipe.s@ges.inatel.br; teo.marques@ges.inatel.br; wsaldanha@inatel.br luizf.santos@inatel.br; wagner.anastacio@gea.inatel.br

***Abstract*— This paper presents the development of an Antweight category combat robot (454 g), named JigglyBullff, developed by the RobotBulls team from Inatel. The work describes the process from the theoretical foundation and category rules to the material selection, electronics, and manufacturing of the robot.**

***Index Terms*—Combat Robotics, Antweight, Mechanical Design, Control Engineering.**

Resumo— Este artigo apresenta o desenvolvimento de um robô de combate da categoria Antweight (454 g), denominado JigglyBullff, desenvolvido pela equipe RobotBulls do Inatel. O trabalho descreve desde a fundamentação teórica e regras da categoria até a seleção de materiais, eletrônica e manufatura do robô.

Palavras-chave— Robótica de Combate, Antweight, Projeto Mecânico, Engenharia de Controle.



Fig. 1. Equipe Robotbulls na RSM 2025

I. Introdução

A robótica de combate vem ganhando cada vez mais destaque no Brasil e no exterior, tanto como atividade competitiva quanto como ambiente de desenvolvimento tecnológico. Nas competições, robôs projetados e construídos por equipes especializadas se enfrentam dentro de arenas protegidas, seguindo um conjunto de regras que garantem segurança, equilíbrio técnico e condições justas de disputa. Esse contexto incentiva o estudo prático de engenharia, reunindo mecânica, eletrônica, controle e fabricação em um único projeto.

A RobotBulls é a equipe de Robótica e Inteligência Artificial do Instituto Nacional de Telecomunicações (Inatel), localizada em Santa Rita do Sapucaí, Minas Gerais. A equipe conta com laboratório próprio para prototipagem e testes, além do suporte de professores orientadores e da participação regular de estudantes estrangeiros em programas de intercâmbio. Essa combinação de infraestrutura e diversidade tem impulsionado a evolução técnica e científica dos projetos desenvolvidos.

Atualmente, a equipe atua em 16 categorias diferentes, indo de modalidades tradicionais como Mini Sumo, Sumo 3 kg, Seguidor de Linha e Futebol de Robôs, até categorias de combate como *Fairyweight*, *Antweight* e *Beetleweight*. Entre seus projetos mais reconhecidos está o robô *Hobbyweight* Bull'dan (12 lb / 5,44 kg), referência mundial na categoria. O Bull'dan é campeão latino-americano no *RCX* (*RoboCore Experience*) e alcançou o Top 4 no campeonato internacional da *NHRL* (*National Havoc Robot League*), resultado que garantiu vaga direta no mundial e consolidou ainda mais o nome da equipe no cenário competitivo.

No Brasil, as competições de maior destaque são o RCX (RoboCore Experience) e a IronCup (Inatel Robotics National Cup), esta última uma das maiores competições de robótica do país. Nessas arenas, os robôs de combate se enfrentam utilizando armas ativas ou passivas, testando soluções de tração, resistência de materiais, controle eletrônico, estratégias de luta e a capacidade de absorver impacto. Além do caráter competitivo, o ambiente se torna uma plataforma importante de aprendizado, permitindo que novas soluções sejam experimentadas e refinadas diretamente em situações reais de combate.

Este artigo apresenta o desenvolvimento completo de um robô de combate para a categoria *Antweight*, com peso máximo de 454 g (1 lb), nomeado JigglyBullff. O robô utiliza tração 2WD tangenciada e uma arma ativa do tipo *eggbeater*, sendo totalmente controlado por rádio controle. As regras da RoboCore permitem liberdade no desenho e nas dimensões, desde que o peso máximo seja respeitado, o que abre espaço para diferentes abordagens mecânicas e eletrônicas, mas exige escolhas cuidadosas para equilibrar desempenho, durabilidade e massa total.

A organização deste trabalho segue a seguinte estrutura: a Seção II apresenta a fundamentação teórica e as principais regras da categoria Antweight; a Seção III detalha os materiais adotados e a aplicação prática no desenvolvimento do JigglyBullff, incluindo estrutura, locomoção, sistema de arma,

eletrônica embarcada e controle; a Seção IV descreve os testes realizados e o desempenho em competições; e a Seção V discute as conclusões, aprendizados e sugestões para aprimoramentos futuros.



Fig. 2. Equipe de combate Robotbulls na NHRL 2025

II. Fundamentação Teórica

Esta seção aborda o contexto da categoria escolhida, bem como as normas que regem o desenvolvimento do projeto.

A. Categoria Antweight e contexto das competições

A categoria Antweight, utilizada nas competições organizadas pela RoboCore, inclui robôs de combate com peso máximo de 454 g (1 lb). Por serem robôs muito compactos, essa categoria costuma exigir projetos bem otimizados, já que toda a estrutura, bateria, sistema de tração, eletrônica e possíveis armas precisam caber dentro desse limite de peso.

As lutas seguem o padrão da organização, com duração de até dois minutos, e o resultado é decidido por nocaute ou pela avaliação dos juízes com base em dano, agressividade e controle [1].

B. Regras principais da categoria

As regras oficiais definem que o robô deve respeitar rigorosamente o peso máximo da classe, e também possuir um peso mínimo equivalente a 40% desse valor [2]. A categoria permite a participação de "*multibots*", desde que o peso total da equipe não ultrapasse os 454 g. Não há um limite de dimensões pré-definido; a única exigência é que o robô consiga passar pela porta da arena, o que dá liberdade ao projeto, mas ainda assim obriga o time a trabalhar com soluções compactas e bem distribuídas.

No que diz respeito à mobilidade, o robô deve apresentar deslocamento controlado utilizando rodas, esteiras ou pernas. São proibidas soluções que resultem em voo ou sustentação aerostática. Diferente de algumas ligas internacionais [3], a RoboCore não oferece bônus de peso para robôs caminhantes, então todos precisam cumprir o mesmo limite independentemente do tipo de locomoção escolhido.

Outra exigência importante está relacionada ao sistema de rádio-controle. Cada robô deve utilizar um rádio próprio, evitando interferência com outros competidores do evento. Além disso, robôs inscritos na categoria de combate não podem participar simultaneamente de outras modalidades na mesma competição, garantindo que o foco e a configuração estejam adequados ao combate.

C. Inspeção técnica e requisitos de segurança

Antes de entrar na arena, todos os robôs passam por uma inspeção técnica. Nessa etapa, são verificadas condições como o peso final do robô, integridade estrutural, proteção da bateria, fixação das peças, funcionamento do fail-safe e segurança geral do conjunto. Como as regras são atualizadas periodicamente, as equipes precisam revisar a versão vigente a cada edição da competição.

D. Desafios da categoria Antweight

A limitação de peso é um dos fatores que mais influencia o desenvolvimento de robôs Antweight. É necessário equilibrar resistência mecânica, autonomia, velocidade e simplicidade do projeto. Materiais leves e estruturas otimizadas são fundamentais, assim como a seleção cuidadosa de motores, eletrônica e bateria. Quando o robô utiliza uma arma ativa, esse equilíbrio se torna ainda mais crítico, já que o sistema precisa ser leve, eficiente e resistente aos impactos repetidos que acontecem durante as lutas.

Além do peso, outro ponto desafiador é a confiabilidade. Robôs tão pequenos estão sujeitos a vibração intensa, picos de corrente, deformações e danos causados pelos impactos. Por isso, as equipes costumam priorizar soluções robustas e fáceis de reparar entre as lutas, garantindo que o robô consiga completar todas as rodadas do evento.

Assim, a fundamentação teórica da categoria *Antweight* envolve compreender as regras, limitações e requisitos de segurança, mas também lidar com as restrições práticas do projeto e com os desafios característicos das competições.

E. Sistema eletrônico

O sistema eletrônico de um robô de combate fundamenta-se nos princípios da eletrônica de potência, circuitos elétricos e modelagem de máquinas elétricas. O correto dimensionamento da alimentação e dos atuadores é essencial para garantir desempenho, eficiência e segurança operacional durante as competições.

O controle de velocidade é geralmente realizado por meio de modulação por largura de pulso (PWM), técnica que ajusta a tensão média aplicada ao motor por meio da variação do ciclo de trabalho do sinal digital. Esse método apresenta alta eficiência energética, reduzindo perdas térmicas quando comparado a controles lineares.

Adicionalmente, a confiabilidade do sistema eletrônico depende da implementação de mecanismos de proteção, como capacitores de desacoplamento, sistemas de fail-safe, fixação adequada dos componentes e proteção física da bateria contra impactos. Em categorias como a *Antweight*, onde vibrações e choques são frequentes, esses cuidados tornam-se ainda mais críticos.

F. Microcontrolador

O microcontrolador atua como unidade central do sistema embarcado, sendo responsável por interpretar os sinais do rádio-controle e gerar comandos aos atuadores. Seu funcionamento baseia-se em sistemas digitais discretos no tempo, processando sinais do tipo PWM ou protocolos equivalentes.

O desempenho do sistema depende da frequência de processamento, da eficiência do código implementado e da baixa latência na leitura e geração de sinais, fatores essenciais para garantir resposta rápida durante o combate.

III. Materiais e Aplicação Prática

A. Desenvolvimento do robô JigglyBullff

O robô desenvolvido para este projeto é um robô de combate da categoria Antweight configurado como 2WD tangenciado com arma do tipo *eggbeater*. Desde o início do desenvolvimento, a equipe buscou uma solução que combinasse resistência, eficiência e facilidade de manutenção dentro do limite de 454 gramas da categoria.



Fig. 3. Vista geral do robô JigglyBullff

B. Estrutura e materiais utilizados

A estrutura principal do robô foi projetada para ser leve e rígida. As laterais foram feitas em alumínio 7075 que é uma liga muito utilizada na robótica de combate por suportar impactos sem adicionar peso excessivo ao projeto. A travessa central foi impressa em ABS e funciona como suporte interno para a eletrônica e parte da arma.

As torres frontais e as proteções traseiras e laterais foram impressas em TPU 98A. Essas peças são consideradas parte de sacrifício porque absorvem impactos diretos e protegem elementos mais importantes do robô. O TPU oferece flexibilidade suficiente para deformar sem quebrar e pode ser substituído rapidamente quando danificado.

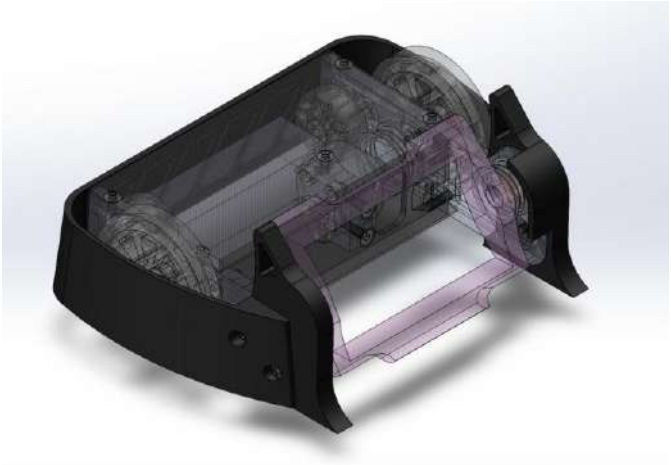


Fig. 4. Detalhe das proteções impressas em TPU.

As tampas superior e inferior são feitas em fibra de carbono de 1 mm. Esse material foi escolhido por ser muito rígido e, ao mesmo tempo, leve, o que ajuda a manter a integridade estrutural do robô sem ultrapassar o limite de peso. Toda a fixação do conjunto é feita com parafusos de aço liga 10.9 que aumentam a resistência contra impactos e vibrações.

C. Sistema de locomoção

O sistema de tração do Jiggly utiliza um conjunto tangenciado. Nesse sistema uma luva instalada diretamente no eixo do motor fica em contato com a roda e esse atrito funciona como uma redução mecânica. O sistema é eficiente porque elimina engrenagens e polias economizando espaço e peso.

As rodas foram produzidas pela própria equipe com resina Vytaflex 60A que fornece boa aderência na arena. O núcleo interno das rodas foi impresso em TPU 98A para aumentar a durabilidade. Para mover o robô foram escolhidos motores *brushless* iFlight Xing 1806 de 1600KV que entregam torque suficiente mesmo com o sistema tangenciado.

D. Sistema de arma

A arma do Jiggly é um *eggbeater* construído em aço 4340 temperado a 48 HRC. Essa escolha garante que a peça seja resistente aos impactos gerados durante as lutas. O acionamento é feito por uma correia TB2 assíncrona que impede que a energia do impacto seja transferida de volta para o motor. Dessa forma o motor não trava e não sofre danos nos rolamentos quando a arma atinge outro robô. O motor utilizado para girar a arma é um T-Motor F40 Pro IV de 1950KV que oferece alta rotação e boa resposta.

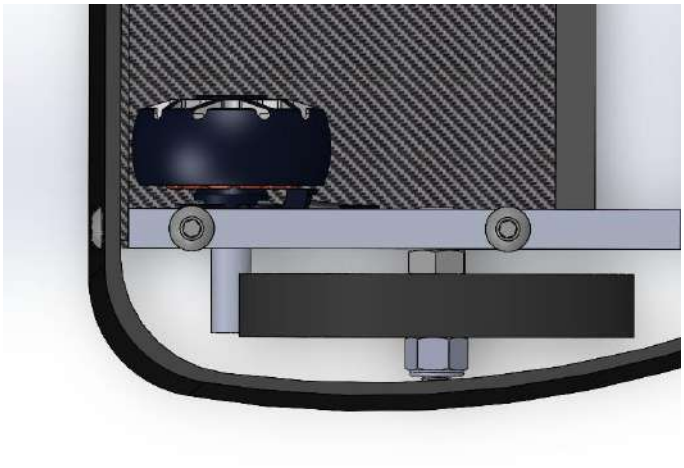


Fig. 5. Sistema de tração tangenciado.

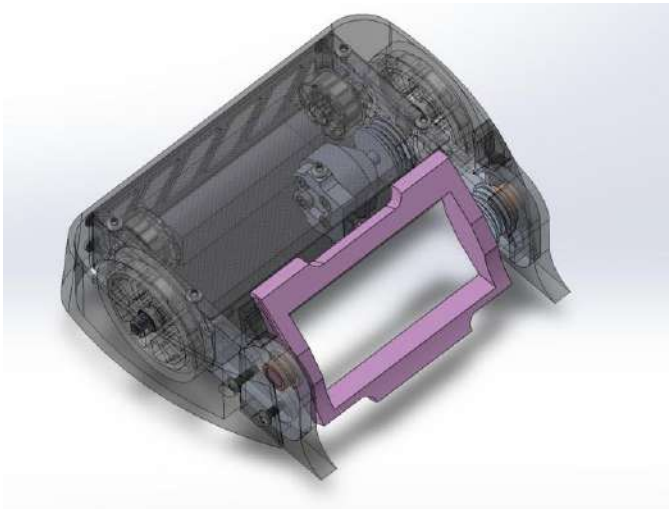


Fig. 6. Detalhe do sistema de arma (eggbeater).

E. Eletrônica embarcada

A alimentação do robô é feita por uma bateria LiPo 3S de 450 mAh. Os motores de tração usam ESCs de 20A BlHeli_S e a arma utiliza um ESC de 30A. Por exigência do regulamento o robô possui uma chave geral FingerTech Mini Switch e um LED de alto brilho para indicar quando o sistema está ativado. O receptor utilizado é um FS2A que trabalha com o protocolo AFHDS 2A. Para alimentar esse conjunto foi utilizado um BEC simples de 5 V que garante tensão estável para os sistemas de controle.

F. Controle remoto e pilotagem

O controle do robô é feito com um transmissor FlySky FS-i6 configurado com mixagem de canais para a locomoção. O eixo vertical do *stick* direito controla o movimento para frente e para trás enquanto o eixo horizontal do *stick* esquerdo controla o giro. Esse modo facilita a pilotagem e deixa o robô mais responsivo. A arma é controlada por um potenciômetro analógico, o que permite ajustar a velocidade da arma com precisão.

G. Ferramentas de desenvolvimento

Todo o projeto mecânico do Jiggly foi desenvolvido no software SolidWorks que permitiu modelar cada parte do robô e ajustar o encaixe dos componentes antes da fabricação. O uso do CAD também facilitou o processo de testes e revisão das peças ao longo do desenvolvimento.

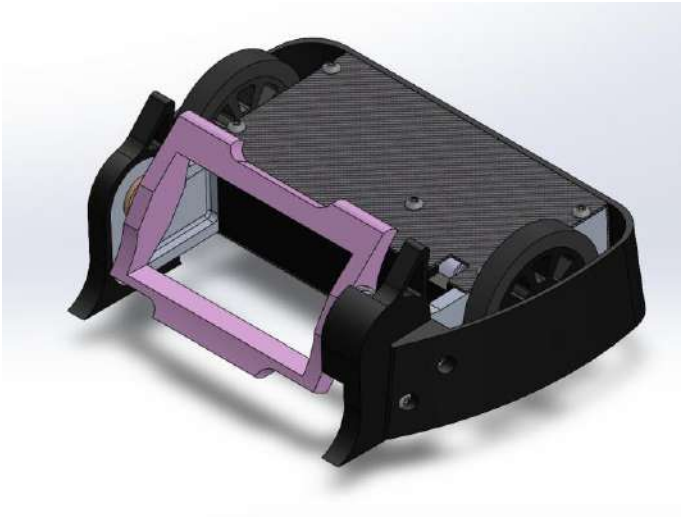


Fig. 7. Renderização do modelo CAD no SolidWorks.

IV. Testes e Resultados

Durante o desenvolvimento do JigglyBullff, diversas etapas de teste foram realizadas para validar o funcionamento do robô antes de qualquer participação em competição. Esses testes incluem a verificação da locomoção, avaliando resposta dos motores, aderência das rodas e comportamento do sistema tangenciado; a validação do sistema de arma, inicialmente em rotações baixas para observar vibração e balanceamento e depois em rotações mais altas para confirmar o funcionamento do motor, do ESC e da transmissão por correia; e a checagem dos sistemas de segurança, especialmente o failsafe, que garante a interrupção imediata do robô caso o sinal de rádio seja perdido. Esse conjunto de verificações faz parte do processo padrão da equipe e permite identificar falhas antes que elas se tornem um problema durante as lutas.

O JigglyBullff participou de duas competições ao longo de 2025: a *IronCup* 2025 e o *RSM Challenge* 2025. Na *IronCup*, o robô demonstrou um desempenho consistente, encerrando sua participação com duas vitórias e uma derrota. Nas primeiras lutas, tanto a tração quanto a arma se comportaram como esperado, e o robô conseguiu aplicar golpes eficientes. No entanto, um problema inesperado ocorreu na arma principal. A barra *eggbeater*, fabricada em aço 4340, apresentou uma têmpera insuficiente e acabou entortando após um impacto mais forte na última luta. Como a equipe não tinha uma arma reserva disponível, o robô precisou encerrar sua participação

mesmo ainda tendo condições gerais de funcionamento. Apesar do ocorrido, o desempenho geral do JigglyBullff na competição foi considerado satisfatório e mostrou que a base do projeto estava bem encaminhada.

No *RSM Challenge* 2025, o desempenho foi mais limitado. O robô finalizou o evento com uma vitória e duas derrotas, resultado influenciado por falhas que já haviam sido percebidas durante os testes pré-luta. A locomoção apresentou comportamento inconsistente, com perda parcial de tração em determinados momentos, e o sistema de arma não manteve a rotação ideal durante as lutas. Esses problemas combinados reduziram a capacidade ofensiva e dificultaram a execução de estratégias que dependeriam de manobras rápidas e ataques precisos. Como consequência, o JigglyBullff não conseguiu repetir o desempenho obtido na *IronCup*.

Mesmo com essas dificuldades, a participação nas duas competições foi essencial para identificar melhorias fundamentais no projeto. Os eventos mostraram a necessidade de aprimorar o tratamento térmico da arma, reforçar o sistema de tração e revisar o balanceamento da transmissão da arma. Essas conclusões orientaram os ajustes na fase seguinte de desenvolvimento e contribuíram para consolidar a evolução técnica da equipe dentro da categoria Antweight.

V. Conclusão

O desenvolvimento do JigglyBullff permitiu avaliar na prática quais soluções adotadas no projeto funcionaram bem e quais precisam ser aprimoradas para que o robô se torne mais competitivo dentro da categoria Antweight. As participações na *IronCup* 2025 e no *RSM Challenge* 2025 mostraram que a arquitetura geral do robô é promissora, especialmente em relação ao desempenho inicial da tração e à eficiência da arma quando o conjunto está operando corretamente. Além disso, o sistema eletrônico embarcado, incluindo ESCs, receptor, bateria e circuito de alimentação, demonstrou ser confiável durante todas as lutas, sem apresentar falhas ou apagões inesperados. Esse comportamento consistente reforça que a configuração adotada para a parte elétrica está bem dimensionada para a categoria.

Apesar dos pontos positivos, os testes e as competições evidenciaram algumas fragilidades estruturais e mecânicas que precisam ser tratadas nas próximas versões. Um dos problemas identificados foi o uso de bucha de bronze no sistema de rotação da arma, que não apresentou a durabilidade necessária e contribuiu para desgaste e instabilidade durante as lutas. Essa peça será substituída por um rolamento esférico, garantindo menor atrito e maior vida útil para o conjunto. A proteção em TPU 98A, utilizada como peça de sacrifício, também se mostrou insuficiente por ser fina demais para suportar impactos mais fortes. A solução prevista é substituir essa proteção por rodas mais espessas, que além de cumprirem função estrutural, absorvem melhor pancadas diretas.

Outro ponto crítico está na travessa de ABS responsável pela união das laterais do robô. O material não resistiu adequadamente às deformações geradas pelos impactos e será substituído por uma travessa em alumínio 7075, que oferece resistência mecânica muito superior sem comprometer a meta de peso do projeto. O eixo das rodas, atualmente baseado em parafusos M4, também demonstrou fragilidade, entortando após pancadas mais severas. Para resolver esse problema, será fabricado um eixo dedicado em aço 4340, garantindo robustez e maior confiabilidade na locomoção.

Com base nesses resultados, os próximos passos do projeto envolvem reforçar os elementos estruturais, revisar o sistema de arma, implementar eixos e travessas mais resistentes e realizar uma bateria de testes mais extensa antes das competições. A equipe também planeja produzir armas reservas com tratamento térmico adequado para evitar falhas como a que encerrou a participação do robô na IronCup. Essas melhorias permitirão que o JigglyBullff avance para as próximas temporadas de forma mais madura, resistente e competitiva, consolidando o aprendizado obtido ao longo do processo.

Referências

[1] RoboCore, “Critério de julgamento - combate,” https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Crit\%C3\%A9rio_Julgamento+-+Combate.pdf, acesso em: 2024.

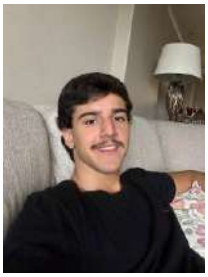
[2] ——, “Regras - combate,” https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Combate.pdf, acesso em: 2024.

[3] NHRL, “Nhrl open rules - 2025: Locomotion based weight bonuses,” https://wiki.nhrl.io/wiki/index.php?title=NHRL_Open_Rules_-_2025#Locomotion_Based_Weight_Bonuses, acesso em: 2024.

Autores



Felipe Silveira Barbosa é graduando em Engenharia de Software pelo Inatel – Instituto Nacional de Telecomunicações, e é bolsista de Iniciação Científica na equipe de robótica RobotBulls, atuando no estudo e desenvolvimento de robôs de combate.



Téo Sacramento Marques é graduando em Engenharia de Software pelo Inatel – Instituto Nacional de Telecomunicações, e é bolsista de Iniciação Científica na equipe RobotBulls, contribuindo para o estudo e desenvolvimento de robôs de combate.



Wagner Luís Caetano Anastácio é graduando em Engenharia de Controle e Automação pelo Inatel - Instituto Nacional de Telecomunicações e é bolsista de Iniciação Científica na equipe de robótica RobotBulls, atuandono estudo e desenvolvimento de robôs de combate.

Estudo de Robôs Inteligentes Aplicados ao Futebol de Robôs Virtual 2D

Júlia R. Garcia, Renata B. Cabral, Débora de Farias Negri
Alexandre B. Lugli, Eugênio M. J. da Silva, Egídio R. Neto,
Luiz Fernando D. R. dos Santos e Wanderson E. Saldanha

julia.garcia@ges.inatel.br, renata.cabral@ges.inatel.br, baratella@inatel.br, eugenio@inatel.br,
egidio.neto@inatel.br, luizf.santos@inatel.br, wsaldanha@inatel.br

Instituto Nacional de Telecomunicações – Inatel

***Abstract* —** *This paper presents the development and organization of a digital soccer team in the RoboCup 2D simulation league. The work focuses on the strategies adopted, the computational methods employed and the artificial intelligence techniques implemented throughout the project. In addition, the study discusses the evolution of the team during the development process and analyzes the main results obtained in simulated matches.*

***Index Terms* —** *artificial intelligence, digital soccer, robotics, simulation*

Resumo — Este artigo apresenta o desenvolvimento e a estruturação de uma equipe de futebol digital na categoria 2D da *RoboCup*. O trabalho descreve as estratégias adotadas, os métodos computacionais utilizados e as técnicas de inteligência artificial aplicadas durante o projeto. Também é discutida a evolução da equipe ao longo do processo de desenvolvimento, bem como os principais resultados obtidos em ambiente de simulação.

Palavras-Chave — inteligência artificial, futebol digital, robótica, simulação

I. INTRODUÇÃO

O avanço da inteligência artificial e da robótica autônoma tem impulsionado o desenvolvimento de sistemas capazes de tomar decisões complexas, atuar de forma independente e cooperar em ambientes dinâmicos. Ao longo das últimas décadas, pesquisas nessa área deixaram de se restringir ao campo teórico e passaram a ser aplicadas em diferentes contextos, como automação industrial, veículos autônomos, sistemas multiagentes e plataformas de simulação voltadas à pesquisa, percepção, planejamento, cooperação e competição [1].

Entre esses ambientes, o futebol robótico consolidou-se como um cenário relevante para o estudo e a experimentação de técnicas de inteligência artificial. A utilização do futebol como plataforma de pesquisa permite explorar problemas típicos de sistemas autônomos, como percepção limitada do ambiente, planejamento estratégico, adaptação a situações dinâmicas e cooperação entre múltiplos agentes.

Dentro desse contexto, a liga de simulação 2D da *RoboCup* destaca-se por oferecer um ambiente computacional acessível e eficiente, no qual agentes de *software* competem em partidas de futebol virtual [1]. Cada equipe é composta por agentes autônomos que precisam interpretar informações parciais do ambiente e tomar decisões em tempo real.

Este trabalho apresenta o desenvolvimento de uma equipe de futebol digital para a liga de simulação 2D. São descritas as etapas de implementação dos agentes, as estratégias utilizadas para movimentação e tomada de decisão e os principais comportamentos autônomos desenvolvidos. O estudo aborda aspectos fundamentais do jogo, como posicionamento tático, controle da bola, ações ofensivas e defensivas e coordenação entre jogadores.

Além da implementação da equipe, o projeto também evidencia o potencial educacional dos ambientes de simulação. Plataformas desse tipo permitem que estudantes e pesquisadores desenvolvam habilidades em programação, modelagem de sistemas inteligentes, tomada de decisão e engenharia de *software*. Ao mesmo tempo, oferecem um ambiente controlado para testes, análise de resultados e aprimoramento de algoritmos.

II. FUNDAMENTAÇÃO TEÓRICA

A liga de simulação 2D da *RoboCup* fornece um ambiente computacional projetado para o estudo de sistemas multiagentes autônomos. Nesse cenário, duas equipes compostas por onze agentes de *software* disputam uma partida de futebol em um campo bidimensional [2]. Toda a dinâmica do jogo é administrada pelo *SoccerServer*, responsável por executar a simulação física, aplicar as regras da partida e garantir condições iguais para todas as equipes.

Esse ambiente controlado permite que pesquisadores comparem diferentes estratégias e algoritmos de inteligência artificial em condições padronizadas [3]. Como resultado, a plataforma tornou-se amplamente utilizada em pesquisas relacionadas à cooperação entre agentes, tomada de decisão em tempo real e comportamento emergente em sistemas distribuídos.

A. Regras do jogo

Durante as partidas, um conjunto de regras inspirado no futebol tradicional é aplicado para garantir que os agentes sigam comportamentos compatíveis com o esporte. Grande parte dessas regras é administrada automaticamente pelo servidor da simulação, enquanto algumas situações específicas podem exigir supervisão humana.

B. Arbitragem automática

O controle principal da partida é realizado pelo árbitro automático implementado no *SoccerServer* [4]. Antes do início de cada tempo e após a ocorrência de um gol, o jogo é interrompido por alguns segundos para que todos os jogadores retornem a posições válidas em campo. Caso algum agente permaneça em posição irregular, o próprio servidor realiza o reposicionamento.

Quando um gol é marcado, o sistema registra o evento, atualiza o placar e reposiciona a bola no centro do campo para reinício da partida. Situações como laterais, escanteios, tiros de meta e reposições também são gerenciadas automaticamente pelo servidor.

O sistema ainda verifica impedimentos, controla a distância mínima de adversários durante cobranças e impede ações consideradas irregulares, como múltiplos chutes dentro do mesmo ciclo de simulação [5].

Cada partida é dividida em dois tempos de aproximadamente cinco minutos de simulação. Em caso de empate, pode ocorrer prorrogação com aplicação da regra do *goal* de ouro.

C. Supervisão humana

Apesar do elevado grau de automação, algumas situações podem exigir interpretação humana. Nesses casos, um árbitro pode intervir para aplicar penalidades em situações como obstrução deliberada de adversários, bloqueio excessivo da área do gol, envio abusivo de comandos ou atraso proposital da partida.

Essa supervisão garante que as equipes não explorem limitações do simulador e que o jogo permaneça equilibrado do ponto de vista estratégico.

D. Limitações do ambiente de simulação

Embora o simulador reproduza muitas características do futebol real, o ambiente apresenta simplificações necessárias para manter a eficiência computacional. O campo é representado em duas dimensões e diversos efeitos físicos, como gravidade, resistência do ar e colisões tridimensionais, são abstraídos.

Outro aspecto importante é que os agentes possuem acesso apenas a informações parciais do ambiente. Cada jogador recebe percepções limitadas e sujeitas a ruído, semelhantes ao que ocorre com sensores em robôs reais. Dessa forma, os agentes precisam estimar o estado do jogo a partir de dados incompletos, o que torna a tomada de decisão mais desafiadora.

E. Desafios da categoria 2D

Os principais desafios da liga de simulação 2D estão relacionados à coordenação de múltiplos agentes, à necessidade de tomada de decisão em tempo real e ao fato de que os jogadores operam com conhecimento incompleto do ambiente.

Cada agente dispõe de aproximadamente 100 milissegundos para interpretar suas percepções e selecionar uma ação. Esse limite exige algoritmos eficientes capazes de equilibrar decisões individuais, como deslocamento ou marcação de adversários, com estratégias coletivas envolvendo posicionamento tático, passes e cobertura defensiva.

Como a comunicação entre jogadores é limitada, as equipes geralmente adotam abordagens descentralizadas, nas quais cada agente atua de forma autônoma, mas alinhada ao comportamento coletivo do time.

A Figura 1 apresenta o fluxo geral de percepção e decisão dos agentes no ambiente do jogo.

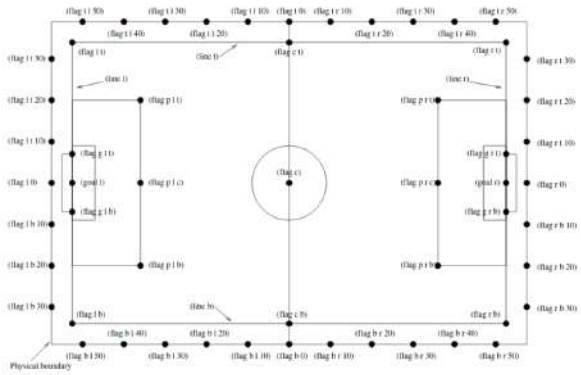


Figure 1: Fluxo geral de percepção e decisão dos agentes no ambiente do jogo.

III. MATERIAIS E APLICAÇÃO PRÁTICA

O desenvolvimento da equipe foi realizado em um ambiente computacional baseado no sistema operacional [6]*Linux*, utilizando a distribuição *Ubuntu* 22.04 LTS em arquitetura de 64 *bits*. A escolha desse sistema se deve à sua estabilidade e à ampla compatibilidade com as bibliotecas utilizadas pelo *RoboCup Soccer Simulator*.

Para facilitar a configuração do ambiente de desenvolvimento, foi utilizada uma máquina virtual executada no *VirtualBox*. Essa abordagem permitiu manter o ambiente isolado e simplificar a instalação de dependências.

A implementação dos agentes foi realizada em linguagem C++, amplamente utilizada em aplicações que exigem alto desempenho computacional. A linguagem oferece suporte à programação orientada a objetos e permite desenvolver sistemas modulares, característica importante para a implementação de agentes autônomos.

No projeto, cada jogador foi modelado como um objeto independente, contendo atributos relacionados ao seu estado e métodos responsáveis por executar ações como movimentação, passe e chute. Essa abordagem facilitou a organização do código e permitiu a implementação gradual de novos comportamentos.

O desenvolvimento foi realizado utilizando o ambiente *NetBeans IDE*, que forneceu ferramentas de organização do projeto, compilação automática e depuração do código.

O núcleo da simulação foi o *RoboCup Soccer Simulator*, composto pelos módulos *rcssserver* e *rcssmonitor*. O primeiro é responsável por executar a simulação e controlar as regras do jogo, enquanto o segundo fornece a interface gráfica para visualização das partidas.

A comunicação entre os agentes e o servidor foi implementada por meio da biblioteca *librcsc*, que disponibiliza funções para recepção de percepções do ambiente e execução das ações dos jogadores [4].

Como base inicial do projeto foi utilizado o *Helios-*

Base, um *framework* amplamente utilizado na comunidade da *RoboCup*. Essa estrutura fornece implementações básicas de percepção, movimentação e controle dos agentes, permitindo que novas estratégias sejam desenvolvidas sobre uma base já consolidada.

IV. TESTES E RESULTADOS

Os testes da equipe foram realizados no ambiente oficial do *RoboCup Soccer Simulator* [5]. Foram conduzidas onze partidas ao longo de dois dias de simulação, mantendo as mesmas condições utilizadas em competições da categoria.

Durante os experimentos foram analisados indicadores de desempenho como número de vitórias, empates e derrotas, quantidade de gols marcados e sofridos, organização tática da equipe e comportamento do goleiro em situações críticas.

Ao final das partidas, a equipe acumulou cinco pontos e marcou nove gols [7] [8]. O setor ofensivo apresentou desempenho satisfatório, com criação consistente de jogadas e movimentação coordenada entre os jogadores de meio-campo e ataque.

Os testes também revelaram fragilidades no sistema defensivo, principalmente no comportamento do goleiro. Em diversas situações foram observados atrasos de reação a chutes diretos, dificuldades de posicionamento em rebotes e limitações na cobertura de bolas rápidas.

Essas limitações contribuíram para um número elevado de gols sofridos, impactando o desempenho geral da equipe.

De modo geral, os resultados indicam que a arquitetura utilizada para movimentação coletiva e construção de jogadas ofensivas apresenta bom potencial. Por outro lado, o sistema defensivo ainda necessita de melhorias significativas, especialmente nos algoritmos responsáveis pela tomada de decisão do goleiro.

V. CONCLUSÃO

A análise das partidas simuladas permitiu identificar avanços importantes no desempenho da equipe, principalmente na organização tática, na movimentação coletiva e na criação de jogadas ofensivas. Os agentes demonstraram capacidade de coordenar ações e explorar espaços no campo, o que resultou em um número significativo de oportunidades de finalização.

Apesar desses progressos, o estudo também revelou limitações que ainda comprometem o desempenho global do time. O comportamento do goleiro destacou-se como o principal ponto crítico, apresentando respostas lentas e inconsistências de posicionamento em situações de pressão ofensiva.

Como continuidade deste trabalho, pretende-se desenvolver novos mecanismos de tomada de decisão para o goleiro, explorando técnicas de *machine learning* que permitam maior capacidade de antecipação e adaptação ao contexto da partida. Também serão investigadas melhorias no controle de bola e na coordenação entre os jogadores de linha.

Além disso, futuros desenvolvimentos incluem a implementação de estratégias táticas mais adaptativas, capazes de ajustar automaticamente o posicionamento da equipe de acordo com o estado do jogo. Espera-se que essas melhorias contribuam para aumentar a competitividade da equipe e consolidar sua evolução em futuras participações na *RoboCup Simulation 2D*.

References

[1] ROBOCUP. *A Brief History of RoboCup*. Disponível em: https://www.robocup.org/a_brief_history_of_robocup. Acesso em: 12 jan. 2025.

[2] RCSS DOCUMENTATION. *Rules of the Game – RoboCup Soccer Simulator*. Disponível em: https://rcsoccersim.readthedocs.io/en/latest/overview.html#the-rules-of-the-game. Acesso em: 12 jan. 2025.

[3] GIMENES, André B. *Aplicação de Redes Neurais e Lógica Fuzzy na Simulação do Jogo de Barganha com Ultimato*. Campinas: UNICAMP, 2009.

[4] ROBOCUP LARC/CBR. *Soccer Simulation League 2D – 2014*. Disponível em: http://www.dainf.ct.utfpr.edu.br/~fabro/larc2014/. Acesso em: 16 fev. 2016.

[5] ROBOCUP. *RoboCup Soccer Simulator*. Disponível em: https://ssim.robocup.org. Acesso em: 29 jul. 2024.

[6] UBUNTU. *Ubuntu 12.04.5 LTS (Precise Pangolin)*. Disponível em: http://releases.ubuntu.com/12.04/. Acesso em: 11 jan. 2016.

[7] INATEL. *Equipe de Robótica do Inatel Conquista Nove Títulos na IronCup 2022*. Disponível em: https://inatel.br/noticias/equipe-de-robotica-do-inatel-conquista-nove-trofeus-na-i Acesso em: 12 jan. 2025.

[8] ROBOCORE. *Resultados – RSM 2025, Team 544*. Disponível em: https://events.robocore.net/rsm-2025/teams/544/results. Acesso em: 12 jan. 2025.

AUTORES

Júlia Ribeiro Garcia

Cursa Engenharia de Software no Instituto Nacional de Telecomunicações (INATEL), onde também participa de projetos relacionados ao desenvolvimento de robôs 2D e à aplicação de conceitos de programação orientada a objetos.



Foto autora

Renata Bernardes Cabral

É graduanda de Engenharia de Software no Instituto Nacional de Telecomunicações (INATEL), onde também colabora em iniciativas voltadas ao desenvolvimento de robôs 2D e à implementação de técnicas de programação orientada a objetos.



Foto autora

Débora de Farias Negri

É graduanda de Engenharia de Software no Instituto Nacional de Telecomunicações (INATEL), onde também colabora em iniciativas voltadas ao desenvolvimento de robôs 2D e à implementação de técnicas de programação orientada a objetos.



Foto autora

Matemática para Engenharias – Uma adequação aos conceitos e habilidades, necessários

Ana Luísa Lyra Pavanelli, Felipe Ferreira de C. G. Pereira, Guilherme Muniz de O. Reis,
Lucas do Lago Figueiredo, Maria Clara B. Ferrer, Luiz Felipe Simões de Godoy
Instituto Nacional de Telecomunicações - Inatel
a.luisa@gea.inatel.br, f.ferreira@ges.inatel.br, guilherme.muniz@ges.inatel.br,
lucas.lago@geb.inatel.br, maria.ferrer@geb.inatel.br, luizfelipe@inatel.br

***Abstract*—This study presents an initiative aimed at improving the academic adaptation of incoming engineering students at Inatel, given the recurring history of failures in the first semester Mathematics course. To mitigate difficulties associated with prior basic education, a free Mathematics leveling course was offered to third year high school students from Santa Rita do Sapucaí MG, taught by engineering undergraduates from the institution. The research focused on students who participated in the course in 2024 and subsequently enrolled at Inatel the following year. Grades from the Mathematics – M01 course, collected at the end of the first semester of 2025, were analyzed to quantitatively assess the influence of the leveling course on academic performance.**

***Index Terms*—Academic Adaptation, Academic Performance, Basic Education, Engineering, Mathematics Leveling.**

***Resumo*—Este trabalho apresenta uma iniciativa voltada à melhoria da adaptação acadêmica dos ingressantes nos cursos de Engenharia do Instituto Nacional de Telecomunicações (Inatel), diante do histórico de reprovações na disciplina de Matemática – M01, do 1º período. Para mitigar dificuldades atribuídas à formação básica, foi oferecido gratuitamente um curso de nivelamento em Matemática aos estudantes do terceiro ano do Ensino Médio da cidade de Santa Rita do Sapucaí-MG, ministrado por graduandos de Engenharia, com o acompanhamento do professor da disciplina M01 que é o orientador desta IC. A pesquisa analisou especificamente os alunos que participaram do curso em 2024 e ingressaram no Inatel no ano seguinte. As notas da disciplina Matemática – M01, coletadas ao final do primeiro semestre de 2025, foram utilizadas para verificar, por meio de análise quantitativa, a influência do curso no desempenho acadêmico. Foram levantadas algumas questões de reflexão sobre o comportamento pedagógico para aprendizagem dos alunos, que serão apresentadas nas análises finais.**

***Palavras chave*—Adaptação Acadêmica, Educação Básica, Engenharia, Desempenho acadêmico, Nivelamento em Matemática.**

I. Introdução

Este trabalho teve como objetivo contribuir para a adaptação dos alunos ingressantes nos cursos de Engenharia do Instituto Nacional de Telecomunicações (Inatel). Considerando o histórico de reprovações na disciplina de Matemática, componente curricular do 1º período, e reconhecendo que uma das possíveis causas está relacionada à formação recebida no ensino básico como indicado em [1]. Assim, foi oferecido, de forma gratuita, um curso de nivelamento em Matemática destinado aos estudantes do terceiro ano do Ensino Médio dos colégios de Santa Rita do Sapucaí-MG. As aulas foram ministradas por alunos de graduação em Engenharia desta IES.

Embora o curso tenha sido aberto a todos os estudantes interessados, os sujeitos desta pesquisa são os alunos que, ao final de 2024, participaram do Processo Seletivo e ingressaram no Inatel. Ao término do primeiro semestre de 2025, foram coletadas as notas da disciplina de Matemática – M01 referentes aos alunos que participaram do projeto em 2024. O objetivo dessa coleta foi verificar, por meio de análise quantitativa, se a participação no curso de nivelamento contribuiu para um melhor desempenho acadêmico, medido por notas e taxas de aprovação.

Dessa forma, este estudo busca comparar o rendimento escolar dos participantes do projeto com o dos demais alunos ingressantes, a fim de avaliar a efetividade da iniciativa na redução das dificuldades em Matemática e, consequentemente, no processo de adaptação acadêmica dos novos estudantes.

II. Desenvolvimento

Acredita-se que os obstáculos de aprendizagem dos alunos na disciplina de Cálculo I estão relacionados a diversos fatores, mas, dentre eles, pode-se destacar a falta de domínio dos conhecimentos de Matemática Básica oriundos do Ensino Fundamental II e Médio, mais especificamente, Álgebra, Geometria, Funções e Trigonometria. De forma especial, os alunos que vivenciaram a pandemia do Covid 19, em qualquer etapa do ensino Fundamental II e/ou Ensino Médio, atestam maior carência de conhecimentos básicos, necessários para o desenvolvimento das habilidades em Matemática. Tais deficiências acabam gerando desgastes excessivos aos alunos ingressantes nos cursos de Engenharia, pois, além da preocupação com os novos conceitos apresentados no ensino superior, percebem que não estão suficientemente aptos para aplicar os objetos e ferramentas matemáticas, estudadas do ensino básico. Em consequência disto, verifica-se em todos semestres o abandono e/ou a reprovação na disciplina de Matemática – M01, e ainda, relacionado às dificuldades em Matemática, uma expressiva reprovação no semestre posterior na disciplina de Cálculo I. Segundo [2], são várias as causas que competem para o alto índice de reprovação em Cálculo I, e aponta como causa importante a má formação adquirida durante o ensino básico, em que grande parte dos alunos chega à universidade sem domínio de conceitos básicos, com pouca capacidade crítica, sem hábitos de estudar e bastante inseguros.

Além disso, a partir da indicação e discussão com o professor orientador desta IC, utilizou-se neste trabalho o conceito de Registros de Representação Matemática de Raymond Duval [3], para avaliar o tipo de cobrança exigida nas questões aplicadas na disciplina M01, a fim de avaliar a implicância

ou não das dificuldades próprias das questões utilizadas na graduação.

A. O curso de nivelamento

O curso de nivelamento foi estruturado com base na apostila da disciplina Matemática – M01, componente curricular do primeiro período dos cursos de Engenharia do Inatel, e desenvolvido sob a avaliação do professor orientador desta iniciação científica. As atividades foram realizadas entre abril e novembro de 2024 em duas turmas independentes, cada uma com carga horária semanal de 4 horas. Para atender estudantes do terceiro ano do Ensino Médio matriculados no período noturno, foi criada uma turma adicional aos sábados, no período matutino. Os conteúdos foram previamente organizados de forma a serem concluídos em um total de 120 horas, permitindo que os participantes conciliassem suas aulas regulares nos colégios com as atividades de nivelamento no Inatel.

No início do curso, a turma de quinta feira registrou 43 alunos inscritos, enquanto a turma de sábado contabilizou 42, totalizando 85 participantes. Entretanto, observou se redução progressiva na assiduidade ao longo do segundo semestre, especialmente após a realização do Exame Nacional do Ensino Médio (ENEM). Parte dos estudantes relatou a necessidade de priorizar outras disciplinas devido à proximidade do ENEM. Outro fator que pode ter influenciado à assiduidade, é que, por se tratar de um curso gratuito, o comprometimento com a frequência mostrou-se abaixo do esperado. No último mês de atividades, em novembro, a presença média oscilou entre 10 e 15 alunos por turma.

Para a análise de desempenho acadêmico no primeiro semestre da graduação, foram considerados apenas os estudantes que atingiram, no mínimo, 70% de presença ao longo do curso de nivelamento e que, posteriormente, ingressaram nesta Instituição de Ensino Superior. Esse critério buscou garantir maior confiabilidade na avaliação do impacto da participação no projeto sobre o rendimento na disciplina Matemática – M01.

B. Análise comparativa

Considerando apenas os estudantes que atingiram, no mínimo, 70% de frequência nas atividades ao longo do curso e que ingressaram no Inatel no primeiro semestre de 2025, foram selecionados 12 participantes para a análise. Como discutido anteriormente, observou-se uma redução progressiva na participação ao longo do desenvolvimento do projeto. Além disso, alguns estudantes que inicialmente aderiram às atividades optaram por migrar para outras instituições de ensino superior, após concluírem que suas áreas de interesse profissional não se alinhavam às opções oferecidas pelos cursos de Engenharia.

Para fins comparativos, analisou-se também o conjunto de estudantes matriculados na disciplina M01 no primeiro semestre de 2025, excluindo-se aqueles que trancaram ou desistiram da disciplina. Nesse período, a disciplina M01 – Matemática Aplicada à Engenharia – registrou um total de 637 alunos distribuídos em oito turmas. Entre eles, 12 participaram deste projeto de iniciação científica, mantendo frequência igual ou superior a 70% durante o ano de 2024. Os demais estudantes envolvidos no curso apresentaram assiduidade insuficiente e, por esse motivo, não foram incluídos na análise.

No período investigado, o índice de reprovação ao final do semestre foi de 26,4% (168 alunos reprovados) entre os estudantes que concluíram a disciplina, evidenciando as dificuldades enfrentadas pelos ingressantes nos cursos de Engenharia do Inatel. Entre os 12 estudantes selecionados como sujeitos desta pesquisa, não foram registrados casos de reprovação. (Tabela I)

	Total	Sujeitos da pesquisa
Matriculados	637	12
Reprovados	168	Não houve
Percentual	26,40%	0%

TABELA I
ANÁLISE COMPARATIVA DE RESULTADOS FINAIS. FONTE: AUTOR (2025)

Entre os 12 estudantes selecionados, sete concluíram a disciplina de Matemática M01 no primeiro semestre de 2025 com nota igual ou superior a 75 pontos, enquanto os outros cinco obtiveram notas entre 60 e 75 pontos. Embora se trate de uma amostra reduzida, 12 estudantes em um universo de 637 matriculados, vale destacar que todos os participantes do projeto em 2024 foram aprovados na disciplina. Esse resultado sugere uma possível relação positiva entre a participação no projeto e o desempenho acadêmico, ainda que não seja suficiente para inferências estatísticas mais robustas.

Em [2] reforça os dados estatísticos apresentados neste estudo sobre os índices de reprovação em Matemática no primeiro período dos cursos de Engenharia do Inatel. O autor destaca que as dificuldades enfrentadas pelos estudantes nas avaliações de Cálculo Diferencial e Integral I estão fortemente associadas a lacunas em conteúdos de Matemática Básica, conhecimentos que deveriam ter sido consolidados ao longo do Ensino Fundamental II e do Ensino Médio.

C. Análise dos resultados

Embora o número de participantes efetivos neste projeto tenha sido inferior ao previsto inicialmente, foi possível identificar aspectos relevantes para reflexão e análise. O primeiro diz respeito à baixa resiliência demonstrada por parte dos estudantes em manter uma rotina consistente de estudos, assiduidade às aulas e compromisso com o próprio processo de aprendizagem. Observou-se que os 12 alunos que concluíram o projeto com frequência superior a 70% apresentaram, no semestre subsequente, desempenho satisfatório na primeira disciplina de Matemática do curso de Engenharia. Assim, é plausível inferir que muitos dos estudantes que não participaram de forma efetiva das atividades em 2024, e que, consequentemente, não obtiveram resultados favoráveis na disciplina de Matemática, também poderiam ter alcançado desempenhos satisfatórios caso tivessem mantido um nível adequado de engajamento e participação. Contudo, esse cenário suscita um questionamento importante sobre a postura dos estudantes matriculados no primeiro período nos cursos de Engenharia: qual é o nível de engajamento desses alunos diante do novo ritmo de estudo, das dificuldades inerentes ao ensino superior e da necessidade de

resiliência frente aos insucessos? A reflexão que motivou essa investigação surgiu da seguinte inquietação: por que estudantes que acabaram de concluir o Ensino Médio apresentam tanta dificuldade em aplicar conceitos matemáticos que já deveriam dominar, considerando que a ementa da disciplina analisada, é formada estritamente por conteúdos trabalhados no Ensino Fundamental II e no Ensino Médio?

Também foi avaliado, ainda que de forma secundária, a maneira como os tópicos de Matemática Básica são abordados na disciplina M01. A análise das médias das avaliações aplicadas, no primeiro semestre de 2025, mostrou que a avaliação sobre funções foi aquela em que os alunos apresentaram o pior desempenho. Observou-se também que as questões envolvendo conversões entre representações gráficas em representações algébricas registraram os maiores índices de erro. Outro tipo de questão com elevado número de respostas incorretas ou deixadas em branco foram as questões contextualizadas, formuladas no formato de situação-problema. As questões abaixo foram aplicadas na primeira avaliação da disciplina de Matemática M01, no primeiro semestre de 2025, para uma das oito turmas, 47 alunos participaram desta prova (Figuras 1 e 2).

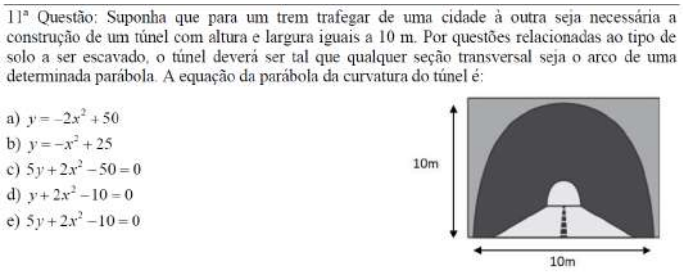


Fig. 1. 11ª Questão da 1ª Prova de M01 – 1ºSem 2025. Fonte: Autor (2025)

Dos 47 alunos que realizaram a prova, 23 acertaram a 11ª questão (Figura 1), 3 deixaram a resposta em branco e 21 escolheram a alternativa errada. Isso corresponde a um índice de 51% de insucesso. Trata-se de uma questão contextualizada, que utiliza uma representação figural para auxiliar na interpretação.

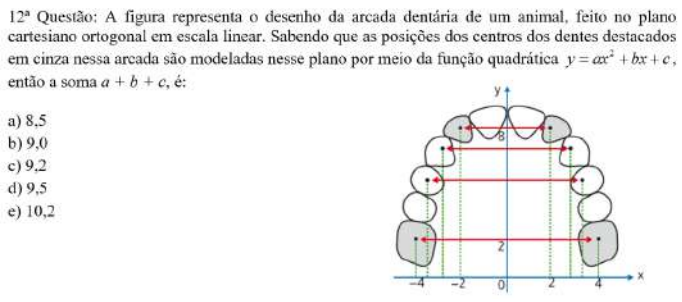


Fig. 2. 12ª Questão da 1ª Prova de M01 – 1ºSem 2025. Fonte: Autor (2025)

Na 12ª questão (Figura 2), o índice de insucesso foi ainda maior. Dos 47 alunos que fizeram a prova, apenas 16 acertaram a questão, enquanto 5 deixaram a resposta em branco e 26 marcaram a alternativa errada. Assim, 66% dos alunos não conseguiram resolver a questão.

Essa questão exige que o estudante interprete a figura apresentada para compreender o gráfico de uma função do 2º grau, realizando a conversão entre o registro figural e o algébrico para chegar à resposta correta.

É relevante destacar que, durante o curso de Matemática oferecido no âmbito desta pesquisa de Iniciação Científica, os participantes trabalharam exercícios semelhantes aos propostos na disciplina M01 do primeiro semestre da graduação. Esses exercícios incluíam tanto problemas contextualizados quanto tarefas que exigiam a conversão de representações gráficas para expressões algébricas. O desempenho satisfatório dos 12 alunos que participaram do curso, em 2024, em comparação com a expressiva reprovação dos alunos matriculados na disciplina, sugere que, ao longo do Ensino Fundamental e Médio, os materiais didáticos e/ou práticas pedagógicas podem não enfatizar, com o rigor necessário, esse tipo de raciocínio matemático. Esse diagnóstico é corroborado por Nasser, Sousa e Torraca (2015) [4], que investigaram dificuldades recorrentes no estudo de Cálculo e constataram que compreender o enunciado de um problema e convertê-lo adequadamente para uma expressão analítica é um desafio tanto para alunos do Ensino Médio quanto para ingressantes na graduação.

III. Conclusão

Os resultados obtidos nesta pesquisa evidenciam que o curso de nivelamento em Matemática, que foi oferecido aos alunos do terceiro ano do Ensino Médio de Santa Rita do Sapucaí–MG, apresentou potencial significativo para contribuir com a adaptação acadêmica dos ingressantes nos cursos de Engenharia do Inatel. Embora a adesão efetiva ao longo do projeto tenha sido inferior ao esperado, os 12 estudantes que mantiveram frequência igual ou superior a 70% demonstraram desempenho satisfatório na disciplina Matemática – M01 no primeiro semestre de 2025, não havendo registros de reprovação entre eles. Esse dado contrasta com o índice geral de reprovação da disciplina, que alcançou 26,4% no mesmo período, sugerindo uma relação positiva entre participação no projeto e rendimento acadêmico.

A análise das dificuldades apresentadas pelos demais estudantes matriculados em M01 reforça que persistem falhas significativas na formação básica em Matemática, especialmente em conteúdos como funções, conversão entre representações e resolução de situações-problema, aspectos já identificados por outros estudos na área, como citado em [5]. Tais lacunas, somadas à baixa resiliência e ao pouco engajamento demonstrado por parte dos ingressantes no enfrentamento das demandas do ensino superior, ajudam a explicar os elevados índices de reprovação e a necessidade de ações pedagógicas complementares.

Diante desse cenário, conclui-se que iniciativas de nivelamento, quando acompanhadas de frequência consistente e participação ativa, podem não apenas minimizar dificuldades conceituais, mas também favorecer o desenvolvimento de hábitos de estudo mais sólidos antes do ingresso na graduação. Ademais, os resultados apontam para a importância de fortalecer, ainda no ensino básico, práticas de ensino que promovam maior domínio de habilidades matemáticas essenciais, especialmente aquelas relacionadas à interpretação, modelagem e transição entre diferentes formas de representação.

Como perspectivas futuras, recomenda-se ampliar o número de participantes efetivos no curso, programar estratégias que promovam maior engajamento e investigar o impacto da

participação no projeto ao longo das disciplinas subsequentes, especialmente em Cálculo I.

Em síntese, o presente estudo confirma a relevância de ações de apoio e articulação entre ensino básico e superior como caminho para reduzir as dificuldades mitigar as reprovações e melhorar a trajetória formativa dos futuros engenheiros.

Agradecimentos

O presente trabalho foi realizado com o apoio da Fundação Instituto Nacional de Telecomunicações - Finatel e do Instituto Nacional de Telecomunicações - Inatel, os quais diligentemente promoveram os recursos necessários para o engendramento do projeto.

Referências

[1] A. Barreto. "O ensino do cálculo I nas universidades". Em: *Informativo da Sociedade Brasileira de Matemática (SBM)* 6 (1995), pp. 4–5.

[2] A. C. R. Cantaruti. "Ensino de Funções no Primeiro Ano do Ensino Médio: uma abordagem com ênfase no comportamento das funções e sua repercussão no Ensino Superior na disciplina de Cálculo I". Tese (Doutorado). Universidade Federal de São João del-Rei (UFSJ), 2017.

[3] R. Duval. "Registros de Representações Semióticas e Funcionamento Cognitivo da Compreensão em Matemática". Em: *Aprendizagem em Matemática: Registros de Representação Semiótica*. Ed. por Silvia D. A. Machado. Campinas: Papirus, 2003, pp. 11–33.

[4] L. Nasser, G. Sousa e M. Torraca. "Transição do Ensino Médio para o Superior: como minimizar as dificuldades em cálculo?" Em: *Atas do V Seminário Internacional de Pesquisa em Educação Matemática*. Petrópolis, RJ: SBEM, 2012.

[5] Marcia Maioli. "A contextualização na matemática do ensino médio". Disponível em: https://repositorio.pucsp.br/jspui/handle/handle/10922. Acesso em: 20 fev. 2025. Tese de dout. São Paulo: Pontifícia Universidade Católica de São Paulo (PUC-SP), 2012.

Autores



Ana Luísa Lyra Pavanelli
Graduanda no 10º período de Engenharia de Controle e Automação pelo Instituto Nacional de Telecomunicações (INATEL) e graduação sanduíche em Informática e Eletrônica de Sistemas Embarcados pela Universidade Polytech Grenoble - INP, França, com bolsa BRAFITEC fomentada pela CAPES. Já participou das monitorias acadêmicas nos cursos de Matemática e Cálculo I das graduações em engenharia pelo Inatel e atualmente realiza estágio no Inatel Competence Center (ICC) em Hardware e Software embarcados.



Felipe Ferreira de Carvalho Gabriel Pereira
Graduando no 7° período de Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel). Atualmente, monitor das disciplinas de Matemática e Calculo I e bolsista de Iniciação Científica na mesma instituição. Tem interesse nos seguintes temas: educação, matemática, estatística e machine learning.



Guilherme Muniz de O. Reis
Graduado em Engenharia de Software pelo INATEL – Instituto Nacional de Telecomunicações e mestrando em Ciências da Computação pela UNIFEI – Universidade Federal de Itajubá. Atuou como monitor em diversas disciplinas, como Matemática, Cálculo I e Algoritmos I, e participou de projetos de iniciação científica na área de Programação 1.0. Já trabalhou como analista de software e como professor de Matemática e Algoritmos. Possui interesse nas áreas de magistério, desenvolvimento de jogos, ciência de dados e desenvolvimento back-end.



Lucas do Lago Figueiredo
Graduando no 7º período de Engenharia Biomédica pelo Instituto Nacional de Telecomunicações, com experiência acadêmica e profissional nas áreas de matemática aplicada, ensino e desenvolvimento de software. Iniciou sua trajetória com trabalho voluntário no projeto Casa Viva, ministrando atividades de soft skills e eletrônica. Posteriormente, atuou como monitor das disciplinas de Matemática Aplicada à Engenharia, Cálculo I e Anatomia Humana, além de ter participado de um projeto de Iniciação Científica em Matemática no próprio instituto. Atualmente trabalha como desenvolvedor Front-End, aplicando conhecimentos de programação e desenvolvimento de interfaces web.



Maria Clara Bernardes Ferrer
Graduanda no 7° período de Engenharia Biomédica pelo Instituto Nacional de Telecomunicações (INATEL). Possui sólida trajetória em pesquisa científica, tendo atuado nos projetos: ”Desenvolvimento e avaliação de adequação de prótese impressa ativa para animal de pequeno porte amputado”, ”Matemática para Engenharias – Uma adequação aos conceitos e habilidades, necessários”e ”Dispositivo vestível para avaliação individual dos músculos do assoalho pélvico masculino”. Estagiou na área de Engenharia Clínica no Hospital Unimed Circuito das Águas. Atualmente, desenvolve pesquisas como bolsista de Iniciação Científica no eHealth - Innovation Center (INATEL).



Luiz Felipe Simões de Godoy
Possui graduação em Engenharia Elétrica pelo Instituto Nacional de Telecomunicações (1994), graduação em Licenciatura Plena em Matemática pela Universidade do Estado de Minas gerais (1997), pós-graduação Lato-Senso em Matemática pela PUC-Minas (1999) e mestrado em Educação Matemática pela PUC-SP. Atualmente é professor Adjunto do Instituto Nacional de Telecomunicações. Tem experiência na área de Física, Cálculo Diferencial e Integral e Álgebra. Tem interesse nos seguintes temas: ensino-aprendizagem de CDI, transtornos na aprendizagem da matemática, registros de representação semiótica.

Estudo para o Desenvolvimento de um Aplicativo de Alfabetização Matemática

R. F. S. Silva, E. L. J. Junior, G. F. Finazzi, L. G. Garcia, T. E. P. Oliveira, J. P. T. Dias, Wanderson E. Saldanha
Instituto Nacional de Telecomunicações - Inatel
rayana.fernanda@geb.inatel.br, giovani.f@ges.inatel.br, thais.emanuelle@ges.inatel.br,
leticia.geovana@geb.inatel.br, dias.joao@ges.inatel.br, wsaldanha@inatel.br

***Abstract*—This article presents a study on the development of AlfaMat, a gamified educational application designed to optimize the teaching and learning of mathematics. The research aims to evaluate the feasibility and potential of AlfaMat as an innovative tool to assist students in improving their mathematical skills. To achieve this, an in-depth literature review was conducted, covering concepts of gamification in mathematics education, as well as instructional design strategies that promote engagement and knowledge retention. Additionally, Artificial Intelligence-based approaches were analyzed to personalize the user experience, adapting content and challenges according to each student's individual progress. The expected results include increased student motivation, reduced learning difficulties, and a positive impact on academic performance, demonstrating the potential of AlfaMat as a unique and effective educational resource.**

***Index Terms*—AlfaMat, learning, teaching, gamification, Math.**

***Resumo*—Este artigo apresenta um estudo sobre o desenvolvimento do AlfaMat, um aplicativo educacional gamificado projetado para otimizar o ensino e a aprendizagem da matemática. A pesquisa busca avaliar a viabilidade e o potencial do AlfaMat como uma ferramenta inovadora para auxiliar estudantes no aprimoramento de suas habilidades matemáticas. Para isso, foi realizada uma revisão aprofundada da literatura, abrangendo conceitos de gamificação no ensino de matemática, bem como estratégias de design instrucional que favorecem o engajamento e a retenção do conhecimento. Além disso, foram analisadas abordagens baseadas em Inteligência Artificial para personalizar a experiência do usuário, ajustando conteúdos e desafios conforme o progresso individual de cada estudante. Os resultados esperados incluem maior motivação dos alunos, redução das dificuldades de aprendizado e um impacto positivo no desempenho acadêmico, demonstrando o potencial do AlfaMat como um recurso educacional diferenciado e eficaz.**

***Palavras chave*—AlfaMat, aprendizado adaptativo, ensino de matemática, gamificação, Inteligência Artificial. .**

I. INTRODUÇÃO

A matemática é uma disciplina essencial para diversas áreas do conhecimento e do desenvolvimento humano, sendo um dos pilares da educação formal. No entanto, muitos estudantes enfrentam dificuldades significativas em sua aprendizagem, o que pode levar à desmotivação e a um baixo desempenho acadêmico. Diante desse cenário, pesquisadores têm buscado novas abordagens pedagógicas que tornem o ensino mais dinâmico e atrativo.

A crescente dificuldade no aprendizado da matemática tem sido amplamente discutida em avaliações educacionais nacionais e internacionais, que apontam lacunas significativas na formação básica dos estudantes. Muitos alunos desenvolvem bloqueios cognitivos e emocionais relacionados à disciplina, resultando em baixa autoconfiança, evasão acadêmica e dificuldades em áreas que exigem raciocínio lógico estruturado. Nesse contexto, torna-se fundamental a adoção de estratégias pedagógicas inovadoras que não apenas transmitam conteúdos, mas que também promovam engajamento contínuo e autonomia no processo de aprendizagem.

Paralelamente, observa-se um aumento significativo do uso de tecnologias digitais no ambiente educacional, especialmente por meio de dispositivos móveis. Aplicativos educacionais vêm sendo utilizados como ferramentas complementares ao ensino tradicional, permitindo uma abordagem mais interativa, personalizada e acessível ao estudante. O avanço do mobile learning amplia as possibilidades de aprendizagem fora do ambiente formal da sala de aula, favorecendo práticas educativas centradas no aluno e adaptadas ao seu ritmo individual.

Diante desse cenário, iniciativas que integram gamificação, inteligência artificial e design instrucional estratégico apresentam grande potencial para transformar a forma como os estudantes interagem com o conhecimento matemático. Ao alinhar elementos tecnológicos com fundamentos pedagógicos consistentes, torna-se possível criar ambientes de aprendizagem mais motivadores, capazes de reduzir barreiras cognitivas e estimular o desenvolvimento progressivo das competências matemáticas.

Uma revisão bibliográfica conduzida por Rosa (2022) [1] analisou o impacto da gamificação no ensino de matemática, destacando sua eficácia na promoção do engajamento dos alunos e na melhoria do processo de aprendizado. A gamificação, ao incorporar elementos típicos de jogos, como desafios progressivos, recompensas e feedback imediato, tem se mostrado uma estratégia promissora para estimular o interesse e a participação ativa dos estudantes.

Com base nessa perspectiva, este trabalho tem como objetivo investigar como a gamificação pode influenciar positivamente o engajamento dos alunos e contribuir para a melhora do desempenho acadêmico na disciplina de matemática. Para isso, propõe-se o desenvolvimento do AlfaMat, uma ferramenta interativa baseada em tecnologia digital e princípios da gamificação. O AlfaMat será estruturado em módulos de aprendizado e incorporará mecânicas de jogabilidade, tornando o processo de ensino mais envolvente e motivador.

Este artigo abordará a concepção do AlfaMat, explorando seu design educacional, os fundamentos teóricos que embasam sua estrutura e as expectativas em relação aos resultados esperados. Ao examinar o potencial da gamificação no ensino

da matemática, este estudo busca contribuir para o avanço das práticas pedagógicas, fornecendo subsídios para uma melhor compreensão de como as tecnologias digitais podem ser utilizadas para aprimorar a aprendizagem em disciplinas fundamentais. Além disso, pretende-se discutir como ferramentas interativas podem impactar a autonomia dos estudantes, incentivando uma abordagem mais ativa e personalizada do aprendizado.

II. ARQUITETURA E IMPLEMENTAÇÃO

O desenvolvimento do AlfaMat baseia-se na ideia de uma trilha de aprendizado estruturada, onde os alunos são guiados por diferentes conceitos matemáticos de forma progressiva. Cada módulo aborda um tópico específico, permitindo que os estudantes construam seu conhecimento de maneira gradual. O sistema também incorpora elementos de gamificação, proporcionando recompensas a cada conquista, como pontos, distintivos e acesso a conteúdos exclusivos. Essa abordagem visa aumentar o engajamento e a motivação dos alunos, criando um ambiente de aprendizado estimulante e interativo.

Além da estrutura modular, o AlfaMat foi concebido considerando o fluxo completo de interação do usuário com o sistema. O processo inicia-se com o cadastro do estudante e a realização de uma avaliação diagnóstica inicial, cujo objetivo é identificar o nível de conhecimento prévio do usuário. A partir dessa análise, o sistema sugere uma trilha de aprendizado personalizada, garantindo que o aluno seja direcionado para conteúdos compatíveis com suas necessidades educacionais.

Durante a utilização do aplicativo, cada interação do aluno — como respostas corretas, erros recorrentes e tempo de resolução — é registrada e analisada. Esses dados permitem que o sistema adapte progressivamente o nível de dificuldade das atividades propostas, promovendo uma experiência de aprendizagem dinâmica e ajustada ao desempenho individual. Essa abordagem contribui para evitar tanto a frustração causada por desafios excessivamente complexos quanto a desmotivação gerada por tarefas muito simples.

Do ponto de vista estrutural, o AlfaMat organiza seus conteúdos em níveis sequenciais, nos quais determinados módulos exigem a consolidação de competências prévias para liberação de novos desafios. Essa lógica de progressão pedagógica favorece a construção gradual do conhecimento e reforça a retenção dos conceitos fundamentais ao longo do processo educativo.

A. Implementação Tecnológica

O AlfaMat foi desenvolvido utilizando o Kodular Creator, uma ferramenta de desenvolvimento visual que permite a criação de aplicativos educacionais sem necessidade de codificação extensiva. Além disso, a plataforma conta com um sistema de inteligência artificial baseado no TensorFlow, uma das bibliotecas mais utilizadas para aprendizado de máquina. Essa IA desempenha um papel fundamental na adaptação do aprendizado, analisando o desempenho dos alunos e ajustando automaticamente a dificuldade dos exercícios conforme a necessidade individual de cada usuário.

B. Mecânicas de Gamificação

Para manter os alunos engajados, o AlfaMat utiliza estratégias de gamificação que transformam o aprendizado em uma experiência dinâmica e motivadora. As principais mecânicas incluem:

Acúmulo de pontos para desbloquear novos módulos e desafios.

Distintivos virtuais para reconhecer conquistas específicas.

Sistema de progressão de níveis, refletindo o avanço do aluno.

Acesso a conteúdos exclusivos, como vídeos explicativos e exercícios adicionais.

Essa abordagem visa incentivar os alunos a persistirem no aprendizado, tornando o estudo mais dinâmico e recompensador.

C. Monitoramento e Aprendizado Adaptativo

A integração da inteligência artificial permite um monitoramento preciso do desempenho dos alunos. Com base nas respostas fornecidas, a IA ajusta dinamicamente a dificuldade das questões, garantindo que cada estudante receba desafios compatíveis com seu nível de conhecimento. Além disso, relatórios personalizados são gerados para auxiliar professores e alunos no acompanhamento da evolução acadêmica. Estudos como o de Sun, Xie e Anderman (2018) [2] destacam a importância da autorregulação do aprendizado (self-regulated learning) para o sucesso dos alunos em ambientes educacionais adaptativos, reforçando a relevância de estratégias que promovam a autonomia e a personalização no processo de ensino.

Essa abordagem inovadora visa personalizar o ensino da matemática, tornando-o mais acessível, eficiente e motivador para todos os usuários do AlfaMat.

III. A IMPORTÂNCIA DA MOTIVAÇÃO E DO ENGAJAMENTO

A matemática, muitas vezes vista como um desafio, pode se transformar em uma jornada de descobertas quando os alunos se sentem motivados e engajados. O AlfaMat surge como uma proposta inovadora, utilizando estratégias que unem tecnologia e educação para criar uma experiência de aprendizado mais interativa e significativa. Ao adotar práticas que incentivam a curiosidade e a participação ativa, o AlfaMat busca não apenas facilitar a compreensão dos conceitos matemáticos, mas também inspirar os estudantes a explorar a disciplina com confiança e entusiasmo.

Sob a perspectiva das teorias contemporâneas da aprendizagem, a motivação desempenha um papel central na construção do conhecimento significativo. Modelos como a Teoria da Autodeterminação destacam a importância de fatores como autonomia, competência e pertencimento para a manutenção do interesse dos estudantes ao longo do processo educacional. Ao oferecer desafios progressivos e feedback imediato, o AlfaMat cria condições favoráveis para o desenvolvimento dessas dimensões motivacionais.

Além disso, o conceito de estado de fluxo (flow), amplamente discutido em estudos sobre engajamento, evidencia que indivíduos tendem a apresentar maior concentração e satisfação quando são expostos a atividades que equilibram desafio e habilidade. A adaptação dinâmica do conteúdo, promovida pela

inteligência artificial do AlfaMat, contribui diretamente para a manutenção desse equilíbrio, potencializando o envolvimento cognitivo e emocional do estudante com a disciplina.

Ao alinhar fundamentos pedagógicos com recursos tecnológicos interativos, o aplicativo busca transformar a percepção tradicional da matemática, reduzindo a ansiedade associada à disciplina e incentivando uma postura mais ativa e investigativa por parte dos alunos.

Essa abordagem reconhece que, para muitos alunos, a matemática vai além de números e fórmulas; ela é uma ferramenta para desenvolver habilidades como raciocínio lógico, criatividade e resiliência. Ao integrar elementos que tornam o aprendizado mais envolvente e personalizado, o AlfaMat abre caminho para uma educação matemática que valoriza o progresso individual e o desenvolvimento de competências essenciais para o futuro.

IV. Protótipo

A. Linguagem e IDE

O desenvolvimento do AlfaMat foi realizado utilizando a linguagem de programação Python, uma escolha estratégica devido à sua versatilidade, facilidade de aprendizado e vasta gama de bibliotecas disponíveis. Python é amplamente reconhecido por sua sintaxe clara e legível, o que facilita a manutenção e a escalabilidade do projeto. Além disso, sua compatibilidade com frameworks e ferramentas modernas torna-o ideal para a criação de aplicativos interativos e dinâmicos, como o AlfaMat.

B. Interação aluno-aplicativo

Interface Amigável e Intuitiva O AlfaMat será projetado com uma interface simples e visualmente atraente, pensada para ser acessível a alunos de diferentes idades e níveis de conhecimento. Menus claros, ícones explicativos e uma navegação fluida garantirão que os estudantes consigam usar o aplicativo sem dificuldades.

C. Inteligência Artificial

A integração da Inteligência Artificial (IA) no AlfaMat oferece uma perspectiva promissora para otimizar o aprendizado de matemática. De acordo com estudos como o de Johnson2016[3], o uso do modelo de Aprendizado Supervisionado se mostra vantajoso devido à sua capacidade de analisar dados históricos do desempenho dos alunos e suas interações com o aplicativo. Isso permite que o sistema identifique padrões de aprendizado individuais e ajuste dinamicamente o conteúdo e as atividades conforme as necessidades específicas de cada usuário. Além disso, o aprendizado supervisionado possibilita a geração de feedback personalizado, o que contribui significativamente para uma experiência de aprendizado mais eficaz e envolvente. Com essa abordagem, o AlfaMat tem o potencial de proporcionar uma aprendizagem mais adaptativa e eficiente, atendendo melhor às necessidades individuais dos alunos e promovendo um engajamento mais profundo com o conteúdo matemático.

V. Conclusão

A divisão do conteúdo em doze módulos no AlfaMat permite uma abordagem estruturada e progressiva do ensino de matemática, garantindo que os alunos adquiram conhecimento de forma gradual e consolidada. Foi observado que o primeiro módulo é o mais longo, pois cobre uma ampla gama de conceitos fundamentais, incluindo frações, operações com frações, expressões numéricas, propriedades das operações matemáticas e introdução a conceitos algébricos. A solidez dessa base é essencial para que os alunos desenvolvam um raciocínio matemático estruturado e consigam progredir nos módulos subsequentes com maior segurança e compreensão.

Além da estruturação em módulos, a estratégia de gamificação implementada no AlfaMat desempenha um papel central no engajamento dos alunos. O sistema de recompensas utilizado busca ativar mecanismos do sistema de recompensa do cérebro humano, favorecendo a motivação intrínseca e extrínseca. Ao estimular os alunos por meio de desafios progressivos e recompensas por acertos, a plataforma promove um ambiente de aprendizado mais dinâmico e envolvente, reduzindo a aversão tradicional que muitos estudantes possuem em relação à matemática.

Outro aspecto relevante do AlfaMat é a capacidade de adaptação do conteúdo conforme o desempenho dos alunos. Essa abordagem personalizada permite que estudantes que apresentam dificuldades possam revisar e reforçar os conceitos antes de avançar para tópicos mais complexos, enquanto aqueles com maior facilidade podem seguir um ritmo mais acelerado, garantindo um aprendizado eficiente e adequado a diferentes perfis de alunos.

Espera-se que a implementação do AlfaMat, com sua metodologia gamificada e abordagem modular, contribua significativamente para o aprimoramento do ensino de matemática, tornando o processo de aprendizado mais acessível, interativo e eficaz. Além disso, o estudo do AlfaMat representa um avanço na exploração do potencial da gamificação no contexto educacional, sugerindo novas possibilidades para o desenvolvimento de ferramentas pedagógicas inovadoras. Pesquisas futuras poderão avaliar a longo prazo o impacto da plataforma no desempenho acadêmico dos alunos, bem como explorar novas formas de integração da tecnologia ao ensino da matemática, ampliando ainda mais as oportunidades de aprendizado de qualidade.

Referências

[1] Emiliana Faria Rosa, Gabriela de Moraes Chaves, José Antonio Ribeiro de Moura, Leandro Mayer, Lopes Neta e Natércia de Andrade. “Educação em pauta: desafios e possibilidades formativas”. Em: (2022).

[2] Zhiru Sun, Kui Xie e Lynley H Anderman. “The role of self-regulated learning in students’ success in flipped undergraduate math courses”. Em: *The internet and higher education* 36 (2018), pp. 41–53.

[3] Daniel Johnson, Sebastian Deterding, Kerri-Ann Kuhn, Aleksandra Staneva, Stoyan Stoyanov e Leanne Hides. “Gamification for health and wellbeing: A systematic review of the literature”. Em: *Internet interventions* 6 (2016), pp. 89–106.

Modelo estacionário de ondas com interferência destrutiva

Pedro Emanuel Almeida Bernardes Rodrigues – e-mail:p.emanuel@gea.inatel.br
Pedro Henrique Gaviolli – e-mail:pedro.gaviolli@gea.inatel.br
Felipe Emanoel Chaves – e-mail:felipe.chaves@inatel.br

***Abstract*—This work aims to develop a stationary model of sound waves featuring destructive interference through the application of active noise cancellation systems. The study integrates numerical simulations conducted in MATLAB with laboratory experiments employing two loudspeakers, sinusoidal tone generation via Audacity, and measurements across a discretized Cartesian plane. Initial evaluations considered configurations with both in-phase and phase-shifted sources, encompassing various source arrangements. The analysis further addresses discrepancies between the model and experimental results, which are attributed to variations in the speed of sound, simplified modeling of attenuation, and the use of transverse waves in the simulation.**

***Index Terms*—destructive interference; longitudinal wave; wave; ambient noise control; active noise cancellation.**

***Resumo*—Este trabalho busca desenvolver um modelo estacionário de ondas sonoras com interferência destrutiva aplicando sistemas de cancelamento ativo de ruído. O estudo combina simulações numéricas em MATLAB com experimentos em laboratório utilizando duas caixas de som, geração de tons senoidais no Audacity e medições em um plano cartesiano discretizado. Inicialmente, foram avaliadas configurações com fontes em fase e defasadas, incluindo diferentes disposições para as fontes. A análise discute também as discrepâncias entre modelo e experimento, atribuídas à variação da velocidade do som, à modelagem simplificada da atenuação e ao uso de ondas transversais na simulação.**

***Palavras-chave*—interferência destrutiva; onda longitudinal; ondulatória; controle de ruído ambiente; cancelamento ativo de ruído.**

I. Introdução

A exposição prolongada a ruídos de alta potência em ambientes industriais representa um desafio crítico para a saúde ocupacional, estando associada a distúrbios de estresse, perda de produtividade e danos auditivos irreversíveis [1]. Para mitigar esses efeitos, as estratégias de controle de ruído dividem-se, fundamentalmente, em duas abordagens: o Controle Passivo de Ruído (Passive Noise Control) e o Controle Ativo de Ruído (Active Noise Control – ANC).

O método passivo fundamenta-se na atenuação da onda sonora por meio de mecanismos de absorção e isolamento, em que a energia cinética da onda é convertida em calor através do atrito com materiais porosos ou impedida por barreiras físicas [2]. Em contrapartida, o sistema ANC baseia-se no princípio da superposição de ondas, em que uma fonte secundária emite um sinal com a mesma frequência e amplitude do ruído indesejado, porém em oposição de fase π rad. Esse processo resulta no fenômeno de interferência destrutiva, minimizando a pressão sonora resultante no ponto de interesse.

Através da integração de simulações computacionais e validações experimentais, este estudo busca estabelecer uma base de conhecimento para a implementação futura de protótipos capazes de operar em cenários industriais com alta complexidade sonora, mirando a manipulação de zonas de cancelamento de ruído não estacionário.

II. Fundamentos Teóricos

O som é uma onda longitudinal mecânica que se propaga em meios materiais elásticos por meio de compressões e rarefações periódicas das partículas do meio. Ao contrário das ondas transversais, nas quais o deslocamento ocorre perpendicularmente à direção de propagação, nas ondas longitudinais o deslocamento das partículas se dá paralelamente à direção de propagação [3]. A velocidade de propagação do som em um gás ideal é dada por

$$v = \sqrt{\frac{\gamma P_0}{\rho}}, \tag{1}$$

na qual γ é o coeficiente adiabático do gás, P_0 a pressão estática do meio e ρ sua massa específica [4]. No ar seco a 20 °C, esse valor corresponde a aproximadamente 343 m/s [5].

O som está sujeito a diversos fenômenos físicos, tais como reflexão, refração e interferência. O presente estudo concentra-se, em especial, na interferência destrutiva, fenômeno que ocorre quando duas ondas de mesma frequência e amplitude encontram-se em oposição de fase, resultando na redução ou até mesmo no cancelamento da intensidade sonora no ponto de superposição [3].

O comportamento da superposição de ondas é regido pelo Princípio da Superposição, que estabelece que o deslocamento resultante em qualquer ponto do meio é igual à soma algébrica dos deslocamentos individuais de cada onda naquele ponto e instante [4]. Para duas ondas longitudinais harmônicas, esse comportamento pode ser descrito matematicamente pela expressão:

$$S(x,t) = A_1 \sin\big(k_1 x - \omega t + \phi_1\big) + A_2 \sin\big(k_2 x - \omega t + \phi_2\big), \tag{2}$$

na qual A_1 e A_2 representam a amplitude da onda, k_1 e k_2 o número de onda, x a posição ao longo da direção de propagação, ω a frequência angular, t o tempo e ϕ_1 e ϕ_2 as fases iniciais de cada onda [3]. O número de onda k está relacionado ao comprimento de onda λ por

$$k = \frac{2\pi}{\lambda}, \tag{3}$$

e a frequência angular ω à frequência f por

$$\omega = 2\pi f. \tag{4}$$

Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí, Minas Gerais, Brasil.

Quando $k_1 = k_2 = k$, a interação entre os sinais resulta em interferência construtiva ou destrutiva, dependendo exclusivamente da diferença de fase $\Delta\phi = \phi_1 - \phi_2$ entre as ondas [4]. Para $\Delta\phi = 0$ (ondas em fase), as amplitudes se somam, gerando interferência construtiva com amplitude resultante máxima $A_{\text{res}} = A_1 + A_2$. Para $\Delta\phi = \pi$ rad (ondas em oposição de fase), as amplitudes se cancelam, resultando em interferência destrutiva com amplitude mínima $A_{\text{res}} = |A_1 - A_2|$; no caso particular de amplitudes iguais $A_1 = A_2$, o cancelamento é total [5].

Este princípio constitui a base física do controle ativo de ruído, técnica particularmente relevante em baixas frequências, nas quais soluções passivas tendem a exigir grande massa e volume para obter atenuação eficiente [6], [7].

Quando as frequências das ondas são próximas, mas não idênticas, observa-se o fenômeno ondulatório conhecido como batimentos, caracterizado por uma modulação periódica da amplitude resultante, perceptível auditivamente como uma variação rítmica de intensidade [3]. A frequência de modulação é dada por

$$f_{\text{bat}} = |f_1 - f_2|. \tag{5}$$

Para frequências significativamente distintas, o resultado da superposição tende a ser percebido como a soma de dois tons independentes, e não como uma modulação audível, ou seja, o conceito de batimentos deixa de se aplicar de forma prática [4].

III. Metodologia

Fundamentando-se nos conceitos de interferência destrutiva e construtiva apresentados no tópico II, a metodologia experimental consistiu na utilização de um plano cartesiano com escala de [1:0,05] e dimensões de 17x17 pontos discretizados, apresentado na Figura 1. Esta padronização permitiu a organização sistemática e o monitoramento pontual da coleta de dados, levando em conta a limitação do comprimento dos cabos das caixas de som, mas preservando a integridade dos dados medidos utilizando pontos a 5 cm de distância, já que dentro das frequências testadas não houve mudança abrupta nos valores de pressão sonora, em dB, nesse intervalo. O aparato experimental foi composto por duas caixas de som convencionais conectadas a um notebook equipado com o software Audacity para geração e manipulação de sinais. Para a medição do nível de dB, utilizou-se a aplicação Decibel X em um dispositivo móvel iPhone, cuja precisão foi previamente testada em laboratório, apresentando uma variação de apenas 1 a 2 dB em relação a um decibelímetro profissional, garantindo praticidade sem comprometer a integridade dos resultados. As caixas de som foram posicionadas sobre o plano cartesiano em diferentes arranjos para fins comparativos. O controle das variáveis acústicas fase, amplitude, frequência e direção foi executado integralmente via Audacity, conforme a Figura 2.

As caixas de som foram posicionadas sobre o plano cartesiano em diferentes arranjos para fins comparativos. O controle das variáveis acústicas fase, amplitude, frequência e direção foi executado integralmente via Audacity.

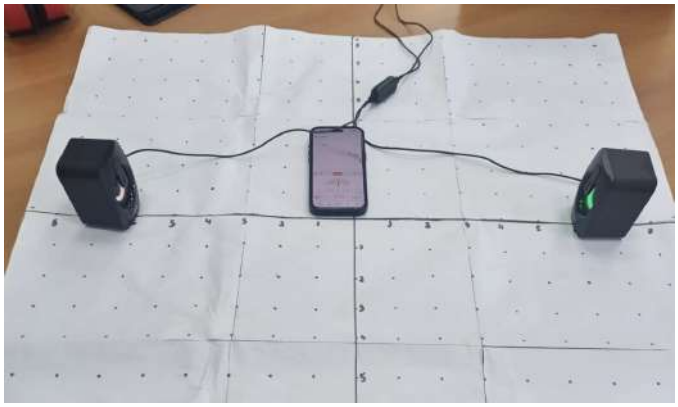


Figura 1: Arranjo usado para ondas em sentidos opostos.

Para a etapa computacional, utilizou-se o software MATLAB para simular as áreas de interação entre ondas senoidais em um ambiente controlado, adotando uma taxa de atenuação linear de 40% por metro. Ressalta-se que a simulação visou primordialmente a comparação das zonas de interferência em frequências específicas, as quais demonstraram padrões similares independentemente do tipo de onda. Discussões detalhadas sobre as discrepâncias de intensidade sonora e a natureza das ondas (transversais vs. longitudinais) são abordadas no tópico V.

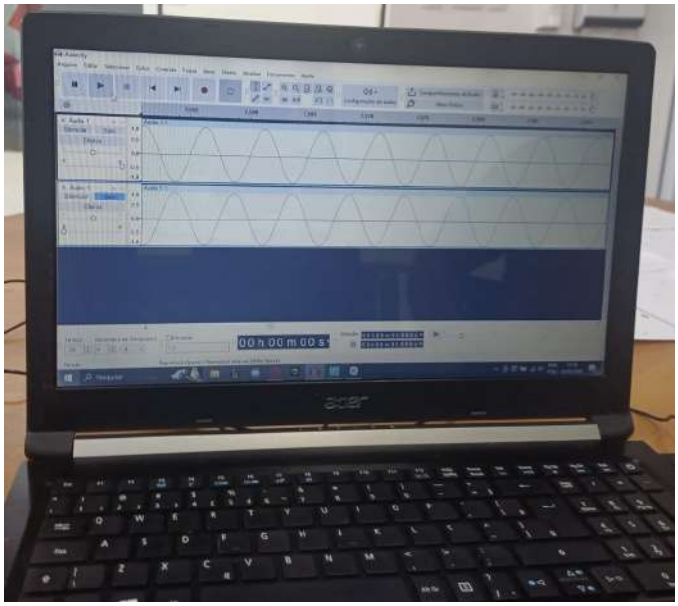


Figura 2: Audacity sendo usado no arranjo da Figura 1.

O processamento de dados seguiu um protocolo rigoroso: após o posicionamento e calibração das fontes, realizou-se a medição ponto a ponto em todo o plano cartesiano. Os valores obtidos experimentalmente foram processados no MATLAB para a geração de gráficos de superfície, permitindo a comparação direta com os modelos teóricos simulados.

As medições em testes, Figura 1, e simulações, Figura 3, focais foram realizadas a 200 Hz, com as fontes a 60 cm de distância e fases opostas. Adicionalmente, realizaram-se testes, Figura 4, e simulações, Figura 5, a 430 Hz com as fontes

separadas por $\lambda/2$ (40 cm) operando na mesma direção e sentido.

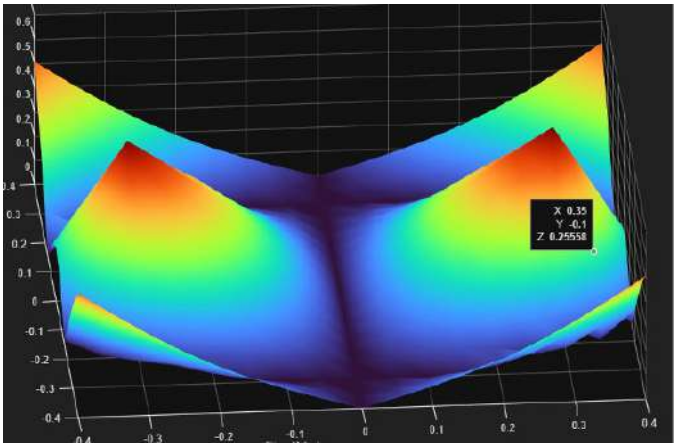


Figura 3: Simulação com 200 Hz (MATLAB).

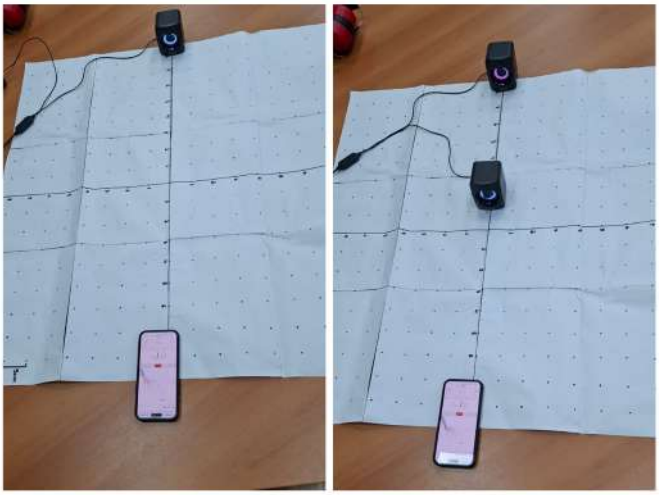


Figura 4: Arranjo usado para ondas no mesmo sentido

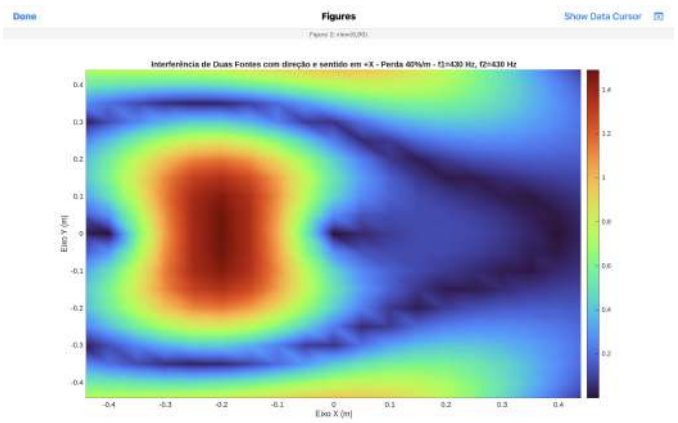


Figura 5: Simulação: 430 Hz e duas fontes no mesmo sentido

Para a obtenção de dados de referência, demonstrados na simulação vista na Figura 7, isolou-se o comportamento acústico de uma fonte única a 200 Hz, removendo-se uma das caixas do arranjo da Figura 1.

IV. Resultados e Análise

Nos experimentos iniciais, realizados com frequências fixas, observou-se que a configuração que apresentou maior atenuação do ruído foi aquela em que as caixas de som estavam sobrepostas, gerando uma área de cancelamento significativamente mais ampla do que em outras disposições testadas. Contudo, devido à inviabilidade prática de aplicar esse arranjo nos testes rentes ao plano R^2, foi testada a configuração com as fontes voltadas para a mesma direção e sentido e posicionadas a $\lambda/2$ de distância uma da outra e com as ondas em mesma fase, como exposto na Figura 4, e na simulação apresentada na Figura 5. O experimento demonstrou uma atenuação considerável, com valores de 10 dB até 18 dB, compatíveis com o observado nas simulações.

Mas, por limitações de alcance dos recursos, optou-se por prosseguir com os testes utilizando as caixas de som posicionadas a 60 cm de distância, voltadas uma para a outra, como pode ser observado na Figura 1.

Nessa configuração, identificaram-se zonas de interferência compatíveis com as simulações realizadas no MATLAB, incluindo regiões de cancelamento (vales) superiores a 15 dB de atenuação, ilustradas pelo gráfico simulado da Figura 5 e pelo gráfico baseado em medições reais da Figura 6.

Ao comparar os perfis acústicos apresentados na Figura 6 e na Figura 7 e ambos sobrepostos na Figura 8, observa-se uma atenuação de aproximadamente 7 dB nas zonas de interferência destrutiva. Paralelamente, identificou-se um incremento de 6 dB nas regiões de interferência construtiva, tomando como referência o nível de pressão sonora de uma fonte isolada.

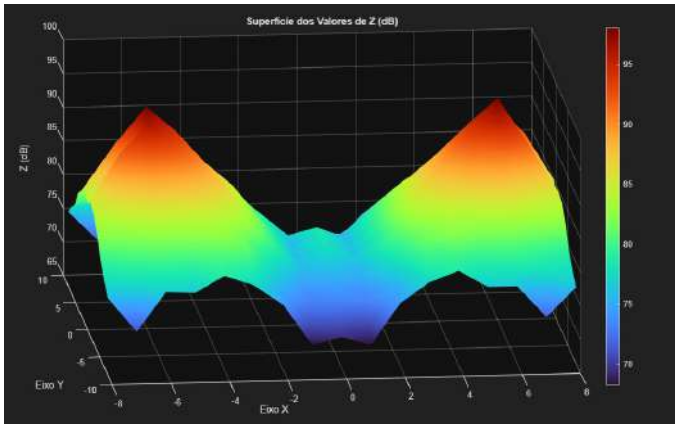


Figura 6: Gráfico baseado nos valores reais, escala [1:0,05].

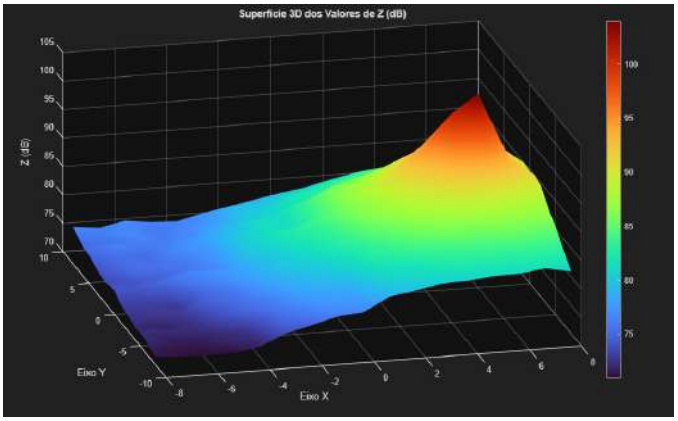


Figura 7: Propagação do som com fonte única de 200 Hz.

Tais resultados demonstram a eficácia do experimento em condições reais e a possibilidade da utilização de tal fenômeno em sistemas de atenuação, controle e cancelamento ativo de ruído ambiente. Embora a performance apresente desvios em relação ao modelo simulado devidos à ausência de mecanismos adaptativos para compensar reflexões em superfícies e a absorção atmosférica do meio, os experimentos demonstraram consistência na validação do fenômeno de cancelamento de fase.

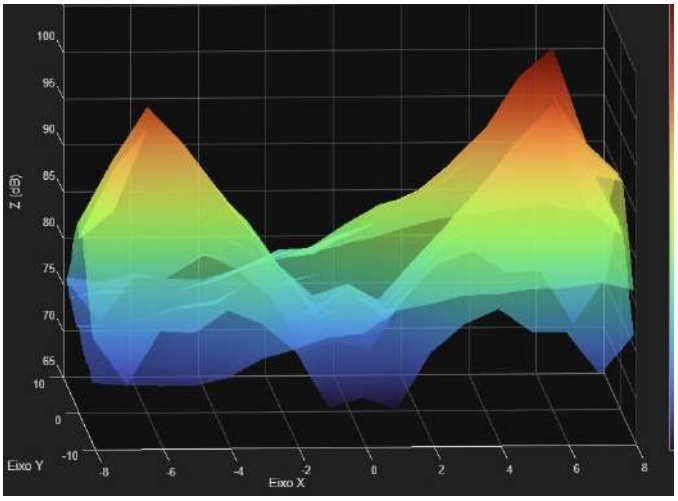


Figura 8: Sobreposição e comparação das Figuras 6 e 7

V. Discussão das Divergências entre Simulações e Medições

A diferença observada entre o gráfico simulado e os dados obtidos experimentalmente pode ser atribuída a duas hipóteses principais, ambas relacionadas às premissas adotadas no modelo computacional.

A. Variação nas Condições Atmosféricas e no Comprimento de Onda

A velocidade de propagação do som, definida por v, utilizada na simulação foi fixada em 343 m/s, valor padrão para seco a 20 °C ao nível do mar; entretanto, o experimento foi realizado em Santa Rita do Sapucaí, Minas Gerais, Brasil e tal região conta com condições ambientais mais quentes, cerca de 27 °C, e em maior altitude, 846 m do nível do mar segundo o CalcMaps. Sabe-se que a velocidade do som é aproximadamente proporcional à raiz quadrada da temperatura absoluta T; logo, em temperaturas ambientes elevadas, v pode se aproximar de 350 m/s. Como a frequência, definida por f, da fonte permanece constante, o aumento de v implica aumento proporcional do comprimento de onda, definida por λ, conforme

$$\lambda = \frac{v}{f}. \tag{6}$$

Esse desvio de λ entre o cenário real e o parâmetro fixado na simulação contribui para o desalinhamento de picos e vales observado entre os gráficos.

B. Imprecisão na Modelagem da Atenuação e do Decaimento do Sinal

A segunda hipótese envolve a metodologia de modelagem da atenuação acústica: a simulação emprega uma taxa de perda fixa e linear de 40% por metro (0,4 dB/m). Tal simplificação diverge do comportamento físico em campo livre, no qual a intensidade sonora decai predominantemente em função do inverso do quadrado da distância; em termos de nível em decibéis, a variação é dada por

$$\Delta L = 20 \log_{10}\left(\frac{R_2}{R_1}\right). \tag{7}$$

A não aplicação explícita desse modelo no ambiente de simulação (no MATLAB, por exemplo), somada a limitações numéricas para distâncias muito pequenas ($R \to 0$) e à adoção de uma forma simplificada baseada em $1/R$, impõe uma perda de sinal que não é fidedigna ao ambiente experimental. Consequentemente, o modelo simulado pode superestimar ou subestimar a intensidade em faixas específicas de distância, produzindo divergências significativas em relação aos dados medidos.

C. Tipos de ondas

Optou-se pela utilização de simuladores de ondas senoidais transversais no MATLAB em vez de ondas longitudinais devido à maior facilidade de controle e análise dos parâmetros envolvidos, como amplitude, fase e frequência. Esse tipo de simulação permite isolar e avaliar com maior precisão os efeitos da interferência construtiva e destrutiva que fundamentam o princípio do controle ativo de ruído, porém até a ínfima diferença entre o tipo de onda simulado e o tipo de onda do experimento também corroboram a tais pequenas variações entre o resultado simulado e o obtido nas mediçoes reais.

VI. Trabalhos Futuros

Como trabalhos futuros, propõe-se o aprofundamento nos estudos de algoritmos de atenuação de ruído, tais como LMS, FLMS e FxLMS, com foco em sua implementação e otimização voltadas ao presente projeto. Além disso, pretende-se investigar algoritmos de machine learning para melhorar a adaptação à

aleatoriedade dos ruídos ambientes, permitindo uma resposta mais robusta em cenários dinâmicos. Também se prevê a aplicação de arranjos com múltiplos sensores e atuadores, de forma a aumentar a eficácia, ampliar a faixa de operação e a área efetiva da região de cancelamento ativo de ruído. Por fim, almeja-se o desenvolvimento de um sistema aplicável de atenuação de ruído em ambientes controlados, que possa servir como prova de conceito para futuras extensões em contextos industriais ou de conforto acústico.

VII. Conclusões

Os resultados obtidos neste trabalho confirmam a viabilidade da interferência destrutiva como mecanismo de cancelamento ativo de ruído em configurações simples de duas fontes sonoras. As simulações em MATLAB mostraram boa concordância qualitativa com as medições experimentais em laboratório, evidenciando atenuações significativas nos vales de interferência e leves reforços nas regiões construtivas. Em especial, observou-se redução de nível de pressão sonora da ordem de alguns decibéis nas zonas de cancelamento, mesmo em ambiente não ideal e sujeito a reflexões e variações ambientais. Por outro lado, as discrepâncias entre o modelo e os dados reais destacam a importância de uma modelagem mais precisa da velocidade do som, da atenuação em função da distância e do tipo de onda considerado. Considerando o comportamento das zonas de atenuação observadas nos experimentos, a criação de zonas de cancelamento de ruído não estacionário utilizará algoritmos adaptativos e arranjos de múltiplos sensores e atuadores para viabilizar a expansão dessas áreas em situações mais complexas.

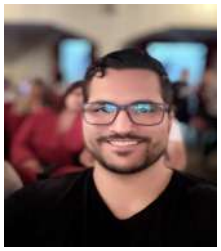
Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio financeiro por meio da concessão de bolsa, fundamental para o desenvolvimento deste trabalho. Agradecem, também, ao Prof. Dr. Felipe Emanoel Chaves pela orientação, dedicação e constante acompanhamento das atividades de pesquisa. Por fim, registram seu agradecimento ao Instituto Nacional de Telecomunicações (Inatel) pela infraestrutura disponibilizada e pelo acesso aos ambientes laboratoriais necessários à realização dos experimentos.

Referências

[1] L. Goines and L. Hagler, “Noise pollution: a modern plague,” Southern Medical Journal, vol. 100, no. 3, pp. 287–294, 2007, DOI: 10.1097/SMJ.0b013e3180318be5.
[2] M. L. Munjal, Acoustics of Ducts and Mufflers, 2nd ed. Chichester, U.K.: John Wiley & Sons, 2014.
[3] P. A. Tipler, Física, vol. 1, 4ª ed.
[4] H. M. Nussenzveig, Curso de Física Básica, 4ª ed.
[5] Sears & Zemansky; Young & Freedman, Física 2: Termodinâmica e Ondas, 10ª ed.
[6] S. J. Elliott and P. A. Nelson, “Active noise control,” IEEE Signal Processing Magazine, vol. 10, no. 4, pp. 12–35, Oct. 1993, DOI: 10.1109/79.248551.
[7] S. M. Kuo and D. R. Morgan, “Active noise control: A tutorial review,” Proceedings of the IEEE, vol. 87, no. 6, pp. 943–973, Jun. 1999, DOI: 10.1109/5.763310.

AUTORES



Pedro Emanuel Almeida Bernardes Rodrigues é graduando em Engenharia de Controle e Automação no Instituto Nacional de Telecomunicações (Inatel). Atua como pesquisador de Iniciação Científica no Health Innovation Center (Inatel).



Pedro Henrique Gaviolli é graduando em Engenharia de Controle e Automação no Instituto Nacional de Telecomunicações (Inatel) e formação técnica em informática pelo Instituto Federal Campus Pouso Alegre. Atua como pesquisador de Iniciação Científica no Health Innovation Center (Inatel)



Felipe Emanoel Chaves é graduado em Física pela Universidade Federal de Itajubá, possui mestrado em Física e Matemática Aplicada pela mesma instituição e doutorado em Matemática Pura pela Universidade Estadual de Campinas. Atualmente atua como professor no Instituto Nacional de Telecomunicações.

IMPACTO DO IF MAIS EMPREENDEDOR EM MICRO E PEQUENOS NEGÓCIOS

Jéssica dos Santos Leite Gonella & Pietro Luís Bertoldo Silva
Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais - Campus Inconfidentes
jessica.gonella@ifsuldeminas.edu.br, pietro.luis@alunos.ifsuldeminas.edu.br

Abstract

The TEC CIRCULAR Project: Competitiveness, Circular Economy and Sustainability presents the actions developed under the IF Mais Empreendedor Program, linked to the Federal Institute of Southern Minas Gerais – Campus Inconfidentes. The initiative aimed to strengthen the economic, social, and environmental sustainability of micro and small entrepreneurs (MSEs) and rural producers in the region through the integration of management tools, digital technologies, and circular economy strategies. Six local enterprises were assisted through structured diagnoses and the application of tools such as SWOT analysis (*Strengths, Weaknesses, Opportunities and Threats*), the GUT matrix (*Gravity, Urgency and Trend*), and the Ishikawa diagram. Based on these instruments, customized action plans were designed and implemented, including workshops, training sessions, technical consulting, and the development of digital materials. The results demonstrate advances in strategic organization, strengthened digital presence, and the adoption of sustainable practices. This study aims to analyze the impacts of these interventions, contributing to the debate on the practical application of managerial tools in micro and small businesses.

***Index Terms*:** Innovation; Sustainability; Circular economy;

RESUMO

O projeto TEC CIRCULAR: Competitividade, Economia Circular e Sustentabilidade apresenta as ações desenvolvidas no âmbito do programa IF Mais Empreendedor, vinculado ao Instituto Federal do Sul de Minas Gerais – Campus Inconfidentes. A proposta teve como objetivo fortalecer a sustentabilidade econômica, social e ambiental de micro e pequenos empreendedores (MPEs) e produtores rurais da região, por meio da integração de ferramentas de gestão, tecnologias digitais e estratégias de economia circular. Foram atendidos seis empreendimentos locais, a partir de diagnósticos estruturados e aplicação de ferramentas como análise *SWOT* (*Strengths, Weaknesses, Opportunities and Threats*), matriz GUT (Gravidade, Urgência e Tendência) e diagrama de *Ishikawa*. Com base nesses instrumentos, foram elaborados e executados planos de ação personalizados, incluindo workshops, oficinas, assessorias técnicas e desenvolvimento de materiais digitais. Os resultados evidenciam avanços na organização estratégica, fortalecimento da presença digital e incorporação de práticas sustentáveis. O estudo tem como objetivo analisar os impactos dessas intervenções, contribuindo para o debate sobre a aplicação prática de ferramentas gerenciais em micro e pequenos negócios.

Palavras-chave:Inovação;Sustentabilidade;Economia circular;Empreendedorismo

I. INTRODUÇÃO

O projeto IF Mais Empreendedor – TEC Circular: Transformação, Economia Circular e Competitividade, desenvolvido no IFSULDEMINAS – Campus Inconfidentes, teve como propósito central apoiar Micro e Pequenos Empreendedores (MPEs), fortalecendo suas práticas de gestão, sustentabilidade e transformação digital. Dessa forma, o objetivo deste estudo é analisar os impactos do projeto IF Mais Empreendedor na gestão, sustentabilidade e presença digital de micro e pequenos negócios da região do Sul de Minas Gerais. A clareza desse objetivo permite alinhar os diagnósticos realizados, os métodos aplicados e as conclusões obtidas, garantindo coerência e rigor acadêmico. A proposta buscou integrar ensino, pesquisa e extensão, conectando a comunidade acadêmica às demandas reais do território e oferecendo soluções inovadoras para empreendimentos locais. Ao todo, foram atendidos seis negócios: Parque Ecológico Cachoeira dos Félix, Brechó Manas, Vanessa *Make-up*, *3D Inc*, Casa da Criança e Estilo Confecção, cada qual com desafios específicos e potencialidades próprias. No Brasil, os micro e pequenos empreendimentos representam uma alta porcentagem das empresas ativas e são responsáveis por parcela significativa da geração de emprego e renda. Apesar de sua relevância econômica e social, essas organizações enfrentam limitações estruturais, tais como restrições de capital, baixa qualificação gerencial e dificuldades de acesso a tecnologias digitais e inovação. Tais desafios dificultam a incorporação de práticas sustentáveis e reduzem sua competitividade em mercados cada vez mais dinâmicos. Nesse cenário, o programa IF Mais Empreendedor, desenvolvido pelos Institutos Federais de Educação, Ciência e Tecnologia, surge como estratégia de articulação entre conhecimento acadêmico e demandas territoriais. No âmbito do Instituto Federal do Sul de Minas Gerais – Campus Inconfidentes, foi implementado o projeto TEC Circular, com o objetivo de promover transformação sustentável em micro e pequenos negócios por meio de consultorias estruturadas, fundamentadas em princípios da economia circular e ferramentas de gestão estratégica. A economia circular foi um dos eixos centrais da proposta, entendida como uma alternativa ao modelo linear tradicional e voltada à regeneração e reutilização de recursos, reduzindo desperdícios e impactos ambientais [1,2]. Neste contexto, buscou-se aplicar também estratégias de gestão como a análise *SWOT*, a matriz GUT e o diagrama de *Ishikawa*, articulando teoria e prática em um processo de diagnóstico e intervenção colaborativa [3].

II. FUNDAMENTAÇÃO TEÓRICA

A fundamentação do projeto IF Mais Empreendedor se apoia em três pilares: economia circular, gestão estratégica e inovação digital. A economia circular constitui modelo econômico regenerativo que busca dissociar o crescimento econômico do consumo de recursos finitos. Esse paradigma baseia-se em três princípios fundamentais: eliminar resíduos e

poluição desde o design, manter produtos e materiais em uso pelo maior tempo possível e regenerar sistemas naturais. Diferentemente da economia linear, a circularidade pressupõe redesenho sistêmico das cadeias produtivas, estimulando inovação, cooperação e responsabilidade compartilhada. Para micro e pequenos negócios, a adoção de práticas circulares pode resultar em redução de custos, fortalecimento de marca, diferenciação competitiva e acesso a novos nichos de mercado. De acordo com MacArthur [1], trata-se de uma economia restaurativa e regenerativa por design. Geissdoerfer *et al*. [2], complementam que a circularidade deve estar integrada à sustentabilidade econômica, ambiental e social, sendo uma alternativa prática ao modelo linear. Para o diagnóstico organizacional, utilizou-se a análise *SWOT* [3], que permite identificar forças, fraquezas, oportunidades e ameaças, proporcionando uma visão sistêmica dos negócios. A priorização de problemas foi feita por meio da matriz GUT, que leva em consideração a gravidade, urgência e tendência de cada desafio. Já o diagrama de *Ishikawa* possibilitou o mapeamento das causas dos problemas organizacionais, auxiliando na construção de soluções com base na lógica de causa e efeito. Por fim, o *Easy Strategic Canvas* foi adotado como ferramenta auxiliar na elaboração dos planos de ação, especialmente nas ações relacionadas ao posicionamento digital e estratégias de comunicação visual [4].

III. MATERIAL E MÉTODOS

A equipe foi composta por bolsistas dos cursos técnicos e superiores do IFSULDEMINAS, sob coordenação de docente com experiência em administração. Os seis empreendimentos atendidos foram selecionados a partir de critérios como vulnerabilidade econômica, potencial de transformação e interesse em práticas sustentáveis. A metodologia adotada caracterizou-se como pesquisa aplicada, de natureza qualitativa e com delineamento de estudo de caso múltiplo, associada a uma abordagem interventiva extensionista. A escolha desse desenho metodológico justifica-se pela necessidade de compreender fenômenos organizacionais em seu contexto real, considerando as especificidades de cada micro e pequeno empreendimento atendido. Participaram do projeto seis empreendimentos da região do Sul de Minas Gerais, selecionados com base em critérios previamente definidos, incluindo vulnerabilidade econômica, potencial de inovação, interesse na adoção de práticas sustentáveis e disponibilidade para participação ativa nas atividades propostas. O percurso metodológico envolveu cinco etapas: diagnóstico inicial com uso da análise SWOT para identificar pontos fortes, fracos, oportunidades e ameaças de cada negócio [3]; priorização de problemas com a matriz GUT; aprofundamento das causas com uso do diagrama de Ishikawa; construção de planos de ação com base na metodologia 5W2H; e execução e avaliação das ações, com aplicação de indicadores de desempenho econômico, social e ambiental [5].A avaliação dos resultados ocorreu por meio de indicadores qualitativos e quantitativos, tais como melhoria na organização estratégica, ampliação do engajamento digital, fortalecimento da identidade de marca, implementação de práticas sustentáveis e percepção dos gestores quanto ao impacto das ações. A coleta de dados pós-intervenção foi realizada por meio de novos encontros avaliativos e aplicação de instrumentos de feedback estruturado. A triangulação entre diagnóstico inicial, execução das ações e avaliação posterior permitiu maior confiabilidade na análise dos resultados e compreensão dos impactos gerados pelo projeto.

IV. RESULTADOS E DISCUSSÃO

As ações desenvolvidas no âmbito do projeto evidenciaram impactos concretos e estratégicos nos micro e pequenos empreendimentos participantes, demonstrando o potencial transformador da aproximação entre instituição pública de ensino e setor produtivo local. Mais do que a execução de atividades pontuais, o projeto contribuiu para o fortalecimento da capacidade gerencial, da competitividade e da sustentabilidade de negócios que exercem papel central na dinâmica econômica regional. Os micro e pequenos empreendimentos representam a base do ecossistema empreendedor brasileiro, sendo responsáveis por significativa geração de emprego, circulação de renda e manutenção da economia local. Nesse sentido, apoiar esse segmento significa fortalecer estruturas econômicas essenciais, ampliar oportunidades e estimular a permanência de negócios competitivos no mercado. Os resultados observados indicam que intervenções técnicas direcionadas podem produzir efeitos relevantes mesmo em organizações com recursos limitados.De forma geral, verificou-se avanço na organização estratégica dos empreendimentos atendidos. Ferramentas como análise *SWOT*, matriz GUT e diagrama de *Ishikawa* permitiram que problemas antes percebidos de maneira difusa fossem convertidos em prioridades claras de ação. Esse processo ampliou a capacidade de planejamento dos gestores, favorecendo decisões mais assertivas e orientadas por diagnóstico técnico.Outro impacto expressivo ocorreu na transformação digital dos negócios participantes. Em diferentes casos, constatou-se fortalecimento da presença em redes sociais, melhoria da comunicação institucional e maior profissionalização da imagem das marcas. Para micro e pequenas empresas, cuja competitividade depende cada vez mais de visibilidade e relacionamento com clientes, tais avanços representam vantagem estratégica relevante.Também foram identificados ganhos associados à sustentabilidade e à inovação. Empreendimentos como o Brechó Manas tiveram sua proposta de valor reforçada a partir da lógica da moda circular e do consumo consciente, enquanto negócios ligados à tecnologia e ao turismo sustentável passaram a comunicar melhor seus diferenciais ao mercado. Isso demonstra que práticas sustentáveis, quando articuladas à gestão, deixam de ser apenas discurso e passam a atuar como elemento concreto de competitividade.No Parque Ecológico Cachoeira dos Félix, a produção de conteúdos digitais e o reposicionamento comunicacional ampliaram a visibilidade do empreendimento e fortaleceram sua imagem vinculada ao
ecoturismo e à responsabilidade ambiental. No Brechó Manas, as ações de capacitação contribuíram para consolidar a identidade da marca e ampliar o reconhecimento de seu propósito sustentável. Já na *3D Inc*, o aprimoramento da comunicação digital e da apresentação comercial favoreceu maior exposição de produtos inovadores e novas possibilidades de conexão com clientes. Além dos efeitos diretos sobre os empreendimentos, o projeto gerou impacto sistêmico ao fortalecer a cultura empreendedora regional. Ao oferecer suporte técnico qualificado às MPEs, a iniciativa contribuiu para reduzir barreiras históricas enfrentadas por pequenos negócios, como ausência de planejamento,

limitações mercadológicas e baixa inserção digital. Em outras palavras, fortaleceu-se a hegemonia empreendedora local ao consolidar micro e pequenas empresas como protagonistas do desenvolvimento econômico.

V. CONCLUSÃO

O estudo evidencia que programas de extensão tecnológica, como o IF Mais Empreendedor, podem desempenhar papel relevante no fortalecimento de micro e pequenos negócios, ao aproximar conhecimento acadêmico das demandas concretas do território. Mais do que melhorias operacionais pontuais, a experiência demonstrou que a aplicação articulada de ferramentas gerenciais favorece processos de tomada de decisão mais consistentes, amplia a capacidade adaptativa dos empreendimentos e estimula práticas alinhadas ao desenvolvimento sustentável. Do ponto de vista científico, os achados contribuem para reduzir a lacuna existente sobre intervenções aplicadas em micro e pequenas empresas, especialmente em contextos regionais, nos quais há menor disponibilidade de suporte técnico especializado.Como limitação, destaca-se o número restrito de empreendimentos analisados e o período reduzido de acompanhamento pós-intervenção. Dessa forma, recomenda-se que pesquisas futuras adotem recortes longitudinais e ampliem a amostra investigada. Conclui-se que o modelo aplicado apresenta potencial de replicabilidade em diferentes contextos territoriais, constituindo alternativa viável para promoção da competitividade, inovação e sustentabilidade em pequenos negócios.



Fig. 1. Workshop de moda circular promovido pelos bolsistas do projeto



Figura 2: Registro da primeira reunião com os bolsistas do projeto TEC Circular

REFERÊNCIAS

[1] MACARTHUR, F. E. What is a circular economy? A framework for an economy that is restorative and regenerative by design. Disponível em: https://www.ellenmacarthurfoundation.org/circular-economy/concept

[2] GEISSDOERFER, Martin; SAVAGET, Paulo; BOCKEN, Nancy M. P.; HULTINK, Erik Jan. The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production*, v. 143, p. 757-768, 2017. https://doi.org/10.1016/j.jclepro.2016.12.048

[3] GÜREL, Emet; TAT, Merba. SWOT Analysis: A Theoretical Review. *Journal of International Social Research*, v. 10, n. 51, p. 994-1006, 2017. https://www.sosyalarastirmalar.com/cilt10/sayi51_pdf/6iksisat_kamu_isletme/gurel_emet.pdf

[4] SOARES, Ricardo. Ferramentas Estratégicas para Gestão de Projetos: Easy Strategic Canvas. São Paulo: Editora Exemplo, 2021.

[5] SCARPELLINI, S. *et al.* Definition and measurement of the circular

economy's regional impact. *Journal of Environmental Planning and Management*, 2019.
https://doi.org/10.1080/09640568.2019.1609517

[6] SEBRAE. Impactos da pandemia nos pequenos negócios: retrato até agosto de 2020. Brasília: Sebrae, 2020. https://www.sebrae.com.br/sites/PortalSebrae/artigos/pesquisa-de-impacto-da-pandemia-nos-pequenos-negocios-ate-agosto,584177bce7919710VgnVCM1000004c00210aRCRD

[7] TIMMERS, Paul. Business Models for Electronic Markets. *Electronic Markets*, v. 8, n. 2, p. 3-8, 1998.
https://doi.org/10.1080/10196789800000016

AUTORES

Jéssica dos Santos Leite Gonella é Doutora em Engenharia de Produção pela Universidade Federal de São Carlos (UFSCar),



com pesquisa voltada para Gestão, Sustentabilidade e Economia Circular. É autora da tese intitulada *"Peoples Awareness of Circular Economy: Scale Development and Relationship With Social Influence, Psychological Barriers, and Favourable Evaluation of Sustainable Development"*. Atua como docente de Administração no Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IFSULDEMINAS) – Campus Inconfidentes, ministrando disciplinas e orientando projetos na área de gestão sustentável, inovação e economia circular. Desenvolve projetos de pesquisa e extensão financiados pelo CNPq e pela FAPEMIG, com foco em sustentabilidade organizacional e desenvolvimento regional.

Pietro Luís Bertoldo Silva é discente do 3º ano do Curso Técnico em Informática no Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais (IFSULDEMINAS) – Campus Inconfidentes, com ênfase em lógica computacional, desenvolvimento de sistemas e



programação. Atua em projetos de pesquisa e extensão voltados à inovação, sustentabilidade e empreendedorismo, com participação ativa no programa IF Mais Empreendedor. É autor de dois artigos científicos na área de Ciências Sociais Aplicadas e desenvolve iniciativas acadêmicas voltadas à integração entre tecnologia e gestão sustentável. Possui certificações nas áreas de tecnologia e sustentabilidade, além de ser sócio de empresa incubada com foco em gestão de resíduos e inovação ambiental.

Estudo da Operação ZVS em Fonte Chaveada Aplicada a Circuito Ressonante

Felipe Coser, Jean Patrick Prigol, Renato Gregolon Scortegagna
Universidade do Oeste de Santa Catarina - UNOESC
felipe.coser@unoesc.edu.br, jean.prigol@unoesc.edu.br, renato.scortegagna@unoesc.edu.br

***Abstract*—This paper presents the development and analysis of a serie-serie (SS) resonant converter with magnetic coupling, supplied by a half-bridge inverter. The main objective is to investigate the conditions required to achieve zero-voltage switching (ZVS) in the semiconductor switches. Simulations were performed by varying the switching frequency, duty cycle, and resonant tank capacitor, evaluating their influence on the voltage-current phase shift and overall system behavior.**

***Index Terms*—SS resonant converter, ZVS, duty cycle, switching frequency, magnetic coupling.**

***Resumo*—Este artigo apresenta o desenvolvimento e a análise de um circuito ressonante série-série (SS) com acoplamento magnético, alimentado por um inversor meia-ponte com modulação por largura de pulso assimétrica (APWM). O principal objetivo é investigar as condições necessárias para a obtenção da comutação suave (Zero Voltage Switching – ZVS) nas chaves semicondutoras. Foram realizadas simulações com variações na frequência de chaveamento, na razão cíclica e no capacitor do tanque ressonante, avaliando seus impactos sobre a defasagem entre tensão e corrente e sobre o comportamento do sistema.**

***Palavras chave*—Circuito ressonante SS, ZVS, razão cíclica, frequência de chaveamento, acoplamento magnético.**

I. Introdução

A transmissão de energia sem fio (*Wireless Power Transfer* – WPT) tem sido objeto de estudo desde os experimentos pioneiros de Nikola Tesla no final do século XIX, quando se buscava a transmissão de energia elétrica a longas distâncias por meio de campos eletromagnéticos. Embora as limitações tecnológicas da época tenham impedido sua consolidação prática, o avanço da eletrônica de potência, dos materiais magnéticos e das técnicas de compensação ressonante possibilitou, nas últimas décadas, a retomada e o amadurecimento dessa tecnologia [1].

Atualmente, sistemas WPT são empregados em diversas aplicações comerciais e industriais. No segmento de dispositivos portáteis, destacam-se carregadores para smartfones, relógios inteligentes e equipamentos eletrônicos de pequeno porte. Na área médica, são utilizados para alimentação de implantes biomédicos, reduzindo riscos associados a conexões físicas e aumentando a confiabilidade do sistema. Em aplicações industriais, a tecnologia é empregada em ambientes agressivos ou selados, onde a ausência de contato elétrico minimiza desgaste mecânico e falhas por corrosão. Mais recentemente, a mobilidade elétrica impulsionou pesquisas voltadas ao carregamento estático e dinâmico de veículos elétricos, ampliando o interesse por sistemas de média e alta potência [1].

A tecnologia de transferência de energia sem fio (WPT) abrange um conjunto diversificado de técnicas, que podem ser classificados em duas categorias principais: sistemas radiativos (*far-field*), baseados na propagação de ondas eletromagnéticas como micro-ondas ou laser, e sistemas não-radiativos (*near-field*), que utilizam campos elétricos ou magnéticos, vide Figura 1 [2].

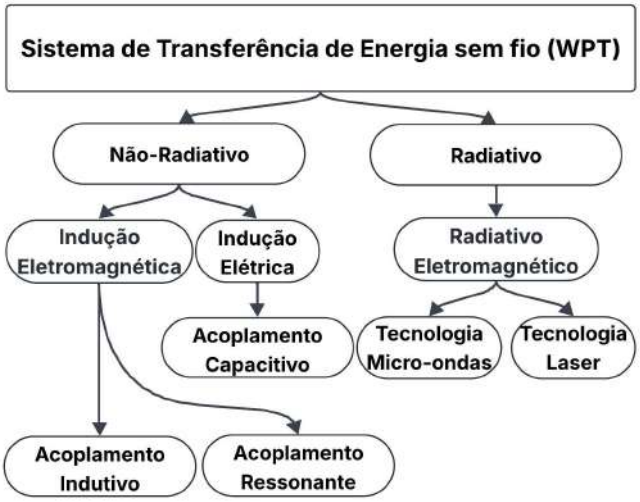


Fig. 1. Diagrama de métodos WPT

Dentre as abordagens não-radiativas, a Transferência Indutiva de Potência (*Inductive Power Transfer* – IPT) destaca-se por utilizar o acoplamento magnético entre as bobinas primária e secundária, eletricamente isoladas, similares a transformadores com núcleo de ar [1]. O desempenho do sistema IPT depende do fator de acoplamento magnético, da frequência de operação e da topologia de compensação empregada [3]. Topologias de compensação, como a série-série (SS), permitem minimizar os efeitos das reatâncias do sistema, melhorar a transferência de potência e aumentar a eficiência global [4].

Apesar dos avanços obtidos com a compensação ressonante, as perdas associadas à comutação dos dispositivos semicondutores ainda representam um desafio relevante, especialmente em aplicações de alta frequência. Nessas condições, as capacitâncias parasitas dos dispositivos podem provocar perdas significativas durante as transições de chaveamento, reduzindo a eficiência do sistema e elevando o estresse térmico, conforme discutido na literatura sobre conversores ressonantes aplicados a MOSFETs [5]. Além disso, elevados gradientes de tensão durante as comutações contribuem para o aumento da interferência eletromagnética (EMI) um aspecto crítico em aplicações industriais e automotivas.

Nesse contexto, técnicas de comutação suave tornam-se

fundamentais para reduzir perdas de chaveamento, reduzir esforços elétricos nos semicondutores e melhorar o desempenho global do conversor. A obtenção da comutação suave está diretamente relacionada ao comportamento dinâmico do circuito ressonante, à defasagem entre tensão e corrente, e às condições de operação impostas pela frequência de chaveamento e pela razão cíclica [5].

Diante desse cenário, este trabalho apresenta o desenvolvimento e a análise de um conversor ressonante série-série aplicado a sistemas IPT, com ênfase na investigação das condições necessárias para a obtenção da comutação suave. São avaliadas as influências da frequência de chaveamento, da razão cíclica e dos capacitores ressonantes sobre a defasagem entre tensão e a corrente no primário na obtenção da condição de comutação suave, especialmente sob tensão nula (*Zero Voltage Switching* - ZVS).

II. Comutação Suave em Conversores Ressonantes

Em conversores estáticos operando em alta frequência, as perdas por comutação representam um dos principais fatores limitantes para o aumento da densidade de potência e da eficiência energética. Durante as transições de chaveamento, a sobreposição entre tensão e corrente nos semicondutores resulta em dissipação de potência, proporcional à frequência de operação, à tensão do barramento e à corrente comutada [4, 6].

Em aplicações que operam em frequências com algumas dezenas de kHz, como sistemas de transferência indutiva de potência (IPT), conversores ressonantes e fontes chaveadas, as perdas por comutação tornam-se críticas, podendo inviabilizar termicamente o sistema se não forem adequadamente dimensionadas. Neste contexto, as técnicas de comutação suave surgem como alternativa para reduzir tais perdas, permitindo a operação em frequências elevadas com maior eficiência [5, 6].

A. Conceito de Comutação Suave

A comutação suave compreende um conjunto de técnicas que visam reduzir ou eliminar a sobreposição entre tensão e corrente durante as transições das chaves semicondutoras. As duas principais modalidades são:

- *Zero Voltage Switching* - ZVS: a chave é acionada quando a tensão sobre ela já é zero (ou próxima de zero), eliminando as perdas por comutação durante o acionamento;
- *Zero Current Switching* - ZCS: a chave é desligada quando a corrente através dela é zero, eliminando as perdas por comutação durante o desligar.

Dentre estas, a técnica ZVS é particularmente adequada para conversores com MOSFETs, uma vez que estes dispositivos apresentam capacitâncias parasitas de saída que podem ser aproveitadas para promover a descarga antes do acionamento [5].

B. Aplicações da Comutação Suave

A técnica ZVS é amplamente empregada em diversas aplicações onde a eficiência e a densidade de potência são requisitos fundamentais:

- Conversores ressonantes para WPT/IPT: sistemas de transferência de energia sem fio operam tipicamente em frequências de dezenas a centenas de kHz, onde as perdas por comutação seriam inviáveis sem técnicas de comutação suave [4, 1];
- Fontes chaveadas compactas: em fontes para eletrônicos portáteis e equipamentos de TI, a comutação suave permite reduzir o tamanho dos dissipadores e aumentar a densidade de potência;
- Inversores para veículos elétricos: sistemas de alta potência exigem máxima eficiência para otimizar a autonomia, tornando a comutação suave essencial;
- Conversores para telecomunicações: aplicações que demandam alta confiabilidade e eficiência em regime contínuo se beneficiam significativamente da redução de perdas;
- Dispositivos médicos: em aplicações como implantes biomédicos (marca-passos, desfibriladores implantáveis e bombas de infusão) e equipamentos de diagnóstico por imagem, a comutação suave é fundamental para garantir a eficiência energética, minimizar o aquecimento e assegurar a biocompatibilidade, além de contribuir para a longevidade das baterias em dispositivos implantáveis [4].
- Redução da interferência eletromagnética - EMI: a comutação suave reduz as interferências geradas pelos conversores, pois opera com variações suaves de tensão e corrente, diferentemente da comutação dissipativa que produz ruídos que afetam equipamentos próximos. Isto é essencial em aplicações automotivas, aeronáuticas e médicas, que exigem conformidade com normas rigorosas de compatibilidade eletromagnética.

C. Condições para Obtenção da Comutação Suave

A obtenção da comutação ZVS está diretamente relacionada à defasagem entre a tensão aplicada ao circuito ressonante e a corrente resultante. Conforme discutido em [5], para que a capacitância parasita da chave seja completamente descarregada antes do acionamento, é necessário que haja corrente circulando no sentido inverso ao da condução da chave durante o intervalo de tempo morto.

Em conversores ressonantes, esta condição é tipicamente obtida quando o circuito opera em regime levemente indutivo. Nesta condição, a corrente apresenta atraso em relação à tensão, garantindo a circulação de corrente no diodo antiparalelo à chave durante o tempo morto, promovendo a descarga da capacitância antes do próximo acionamento [5].

Pesquisas recentes têm explorado diferentes estratégias para estender a faixa de operação com ZVS em conversores ressonantes. Destacam-se:

- Modulação por largura de pulso assimétrica (APWM): permite controlar a tensão de saída enquanto mantém condições favoráveis à comutação suave [5].
- Ajuste dinâmico da frequência: Conforme discutido em [2], a operação com frequência variável é característica de conversores quase-ressonantes (QRCs - *Quasi-Resonant Converters*), onde a frequência de chaveamento é ajustada para manter a condição de comutação suave em diferentes condições de carga.
- Projeto otimizado do circuito ressonante: dimensionamento dos componentes reativos para garantir corrente

suficiente durante o tempo morto em toda a faixa operacional. Em [5], o projeto cuidadoso dos elementos ressonantes é apresentado como fator determinante para a obtenção de ZVS em uma ampla faixa de operação.

Apesar dos avanços, os desafios persistem, particularmente em condições de carga leve e em conversores com ampla faixa de tensão de entrada, onde a manutenção da condição ZVS em todas as situações operacionais nem sempre é possível ou requer estratégias de controle mais sofisticadas.

O objetivo deste trabalho é, por meio de simulações computacionais, avaliar o comportamento de um inversor meia-ponte sob aspecto da comutação suave (ZVS) aplicado a um circuito ressonante com acoplamento magnético. Busca-se verificar as condições em que a chave de potência realiza a comutação com tensão próxima de zero, visando à redução das perdas por chaveamento e ao aumento da eficiência do sistema.

Além disso, pretende-se avaliar a influência da variação dos capacitores do circuito ressonante, da razão cíclica da fonte e da frequência de chaveamento na obtenção da comutação suave. As simulações têm como finalidade compreender como esses parâmetros afetam o desempenho do sistema e identificar faixas de operação favoráveis à comutação suave.

III. Elementos do Projeto

Para a realização das análises propostas, foi desenvolvido um circuito ressonante série-série (SS) com acoplamento magnético, alimentado por um inversor meia-ponte com modulação por largura de pulso assimétrica (APWM). O sistema de alimentação meia-ponte é apresentado na Figura 2 e seus subsistemas são detalhados a seguir.

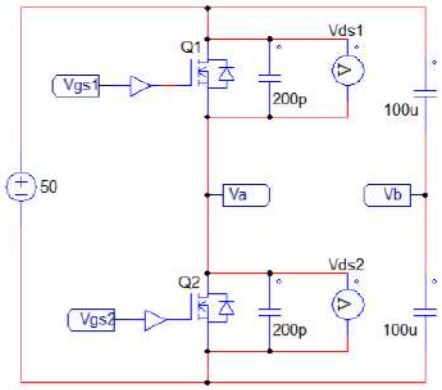


Fig. 2. Sistema de Alimentação Meia-Ponte

O inversor meia-ponte é composto por dois MOSFETs acionados de forma complementar e responsáveis por gerar uma tensão alternada de alta frequência no ponto médio do circuito. Os terminais V_a e V_b representam os terminais de saída do inversor, onde conecta-se o circuito ressonante transmissor IPT. Os capacitores intrínsecos aos MOSFETs (C_{ds1} e C_{ds2}) representam as capacitâncias parasitas, enquanto os capacitores de barramento garantem a estabilidade da tensão DC durante o chaveamento.

A. Gerador de APWM

Este trabalho utiliza um gerador de modulação por largura de pulso assimétrica(APWM), vide Figura 3, desenvolvido para o acionamento da fonte chaveada. O gerador é baseado na comparação entre um sinal de referência contínuo, responsável pela definição da razão cíclica, e uma onda portadora periódica, cujo é possível alterar a frequência de chaveamento. A partir dessa comparação, obtém-se um sinal APWM com frequência desejada e razão cíclica ajustável.

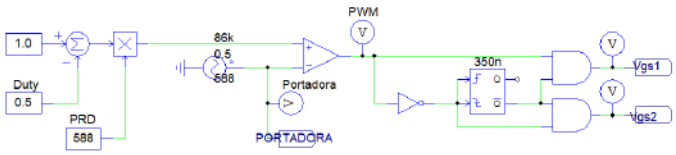


Fig. 3. Gerador de APWM

O circuito inclui uma etapa responsável pela geração de sinais de acionamento e pela inserção de tempo morto (*dead-time*) entre os pulsos de comando, evitando a condução simultânea das chaves do conversor. Esse recurso é fundamental para a operação segura da fonte. Os sinais resultantes são aplicados diretamente aos terminais de *gate* dos MOSFETs, garantindo sincronismo e controle do chaveamento.

O gerador permite a análise do comportamento da fonte sob diferentes condições de operação, possibilitando a variação da razão cíclica e da frequência de comutação.

B. Circuito Ressonante série-série

Na Figura 4 observa-se um circuito ressonante do tipo série–série (SS) com acoplamento magnético, aplicado a sistemas de transferência indutiva de potência. O circuito é composto por um estágio inversor responsável por aplicar uma tensão alternada de alta frequência ao circuito ressonante primário, formado por um capacitor e um indutor em série. A operação próxima à frequência de ressonância contribui para a melhoria da eficiência da transferência de energia.

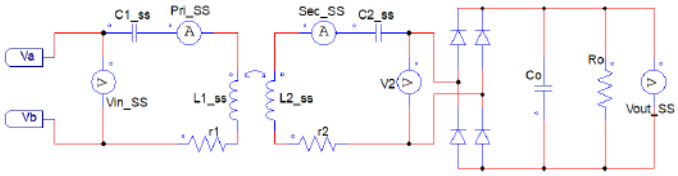


Fig. 4. Circuito Ressonante série-série

A transferência de potência ocorre por meio do acoplamento magnético entre os indutores do primário e do secundário, proporcionando isolamento entre os estágios. O circuito ressonante no secundário, sintonizado na mesma frequência do primário, reduz os efeitos indutivos e capacitivos do sistema, contribuindo para uma transferência de energia mais eficiente e menos sensível às variações de carga. A tensão alternada induzida é convertida em tensão contínua por meio de um retificador em ponte de diodos e, posteriormente, filtrada por um capacitor, permitindo o fornecimento de energia à carga.

IV. Funcionamento do Sistema

O funcionamento baseia-se na modulação APWM, que permite larguras de pulso diferentes para os semiciclos, controlando a tensão média aplicada ao circuito ressonante. Os sinais

APWM gerados acionam alternadamente Q_1 e Q_2, produzindo no ponto médio uma tensão alternada de alta frequência.

Esta tensão é aplicada ao circuito ressonante primário, formado por C_1 e L_1. Quando a frequência de chaveamento é ajustada adequadamente em relação à frequência natural do circuito, a corrente I_{L1} pode apresentar defasagem em relação à tensão de entrada. Em particular, a operação levemente acima da frequência de ressonância torna o circuito indutivo, com a corrente atrasada em relação à tensão, condição que favorece a descarga das capacitâncias parasitas antes do acionamento das chaves, permitindo a comutação suave.

A energia é transferida magneticamente ao secundário através do acoplamento entre L_1 e L_2. O circuito ressonante secundário (C_2 e L_2), sintonizado na mesma frequência, minimiza as reatâncias do sistema. A tensão alternada induzida em L_2 é retificada pela ponte de diodos e filtrada por C_o, resultando em tensão contínua aplicada à carga R_o.

V. Cálculos Realizados

Para a realização dos cálculos e simulações, foram fixados alguns valores que estão contidos na Tabela I.

TABELA I
PARÂMETROS FIXADOS DO PROJETO

Parâmetro	Circuito SS
Frequência de Cálculo (f_s)	85 kHz
Carga na Saída (R_o)	62 Ω
Fator de acoplamento (k)	0,235
Bobina transmissora (L_{1ss})	122 μH
Bobina receptora (L_{2ss})	163 μH
Tempo Morto	350 ns
R_1	28 $m\Omega$
R_2	170 $m\Omega$
C_o	10 μF

Para o dimensionamento dos capacitores ressonantes $C1_{\text{ss}}$ e $C2_{\text{ss}}$, foram empregadas as equações (1) e (2), respectivamente, conforme metodologia apresentada em [4].

$$C1_{ss} = \frac{1}{(2\pi \cdot f_s)^2 \cdot L_1} = 28,737nF \tag{1}$$

$$C2_{ss} = \frac{1}{(2\pi \cdot f_s)^2 \cdot L_2} = 21,508nF \tag{2}$$

VI. Simulações

Durante o desenvolvimento do projeto, investigam-se os efeitos da variação da razão cíclica, do valor do capacitor do circuito ressonante primário ($C1_{ss}$) e da frequência de chaveamento da fonte. Essas variáveis exercem impacto direto sobre o comportamento do circuito, afetando a forma de onda das tensões e correntes, bem como o desempenho do chaveamento.

Por meio de simulações computacionais com o software PSIM [7], busca-se analisar o comportamento do sistema sob diferentes condições de operação, permitindo compreender a relação entre os parâmetros ajustáveis e o desempenho global do conversor. A abordagem adotada contribui para a identificação de condições operacionais mais adequadas e fornece subsídios para o projeto e a otimização de sistemas ressonantes aplicados à transferência de potência em alta frequência.

No ambiente de simulação PSIM, as chaves Q_1 e Q_2 foram modeladas por meio do bloco *MOSFET switch* com configuração *Ideal*. Para esse modelo, adotou-se resistência de condução $R_{DS(on)} = 4{,}8\,m\Omega$, tensão direta do diodo igual a $0\,V$ e resistência do diodo de $1\,m\Omega$. As capacitâncias parasitas de saída dos semicondutores não foram consideradas embutidas no modelo da chave, sendo representadas explicitamente no circuito pelos capacitores C_{ds1} e C_{ds2}, ambos com valor de $200\,pF$.

Neste trabalho, o critério adotado para caracterizar a ocorrência de comutação suave (ZVS) foi a verificação de tensão nula sobre a chave no instante de seu acionamento.

A. Circuito Ressonante

Aplicando os valores obtidos nos cálculos anteriores, observa-se na Figura 5 que o circuito opera em condição de ressonância, evidenciada pelo alinhamento em fase entre a tensão V_{in}, aplicada à entrada do circuito ressonante, e a corrente medida no enrolamento primário, confirmando o comportamento esperado do sistema. Essa condição é fundamental para a obtenção da máxima eficiência do sistema.

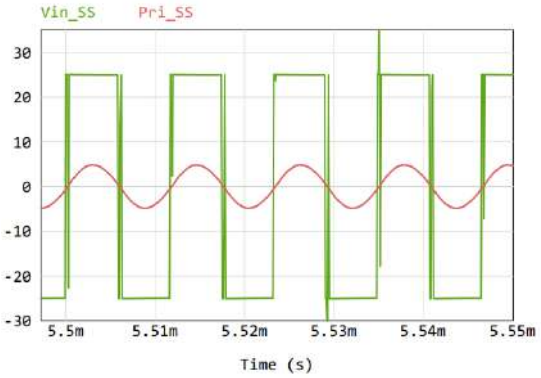


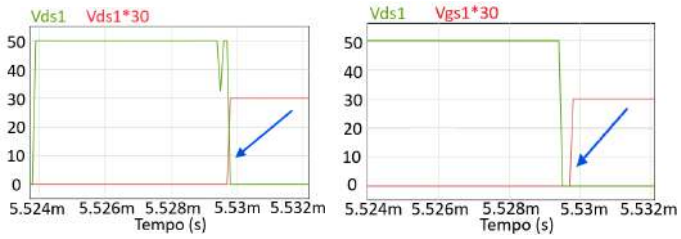
Fig. 5. Onda ressonante do Circuito

Nessa condição, embora o circuito esteja operando em ressonância, não foi possível obter a condição de comutação suave, conforme observado na Figura 6. Para facilitar a visualização dos sinais, o sinal de *gate-source* das chave 1 (V_{gs1}) foi multiplicado por 30, sem alterar seu significado físico, apenas com o objetivo de destacar graficamente o instante de acionamento em relação à tensão (V_{ds1}). No momento em que o sinal de *gate* é aplicado às chaves, a tensão V_{ds} ainda apresenta valor elevado, caracterizando a ocorrência de comutação sob tensão. Dessa forma, há sobreposição entre tensão e corrente durante o acionamento, resultando em aumento das perdas de comutação, maior dissipação de potência e redução da eficiência do sistema.

Isso ocorre porque a corrente e a tensão encontram-se em fase, não havendo corrente suficiente no instante de comutação para promover a descarga completa das capacitâncias parasitas das chaves. Para solucionar esse problema, foi necessário aumentar em 3 nF o valor do capacitor $C1_{ss}$, deslocando o ponto de operação do circuito para uma condição levemente indutiva. Essa alteração provoca um atraso da corrente em

relação à tensão, conforme apresentado na Figura 6. Esse comportamento está em concordância com a análise clássica de conversores ressonantes operando acima da frequência natural, onde o comportamento indutivo garante a defasagem de corrente necessária para descarga das capacitâncias parasitas das chaves, conforme discutido em [5].

Adicionalmente, a análise das formas de onda de tensão e corrente na chave permitiu avaliar a potência dissipada associada ao processo de comutação. Na condição sem ZVS, observou-se valor de 17,94 W, enquanto, na condição com ZVS, esse valor foi reduzido para 4,83 mW. Essa diferença evidencia a redução expressiva das perdas de chaveamento proporcionada pela comutação suave.



(a) Operação sem ZVS. (b) Operação com ZVS.

Fig. 6. Comparação entre a operação sem ZVS e com ZVS.

B. Razão Cíclica

Mantendo-se esse ajuste do capacitor série ($C1_{ss}$) com um aumento de 3 nF, investiga-se o comportamento do circuito frente a variações da razão cíclica e o impacto na comutação suave. As simulações realizadas indicam que, para valores de razão cíclica abaixo de 0,32, ocorre a perda da condição de ZVS na chave Q_2, enquanto a chave Q_1 mantém a comutação suave com ampla margem. A variação da razão cíclica também resulta em redução da tensão de saída do circuito ressonante série–série, evidenciando seu impacto direto no desempenho do sistema. Observa-se ainda que, para valores de razão cíclica superiores a 0.68, ocorre uma inversão na condição de operação dos interruptores, onde a chave Q_2 passa a apresentar comutação suave com margem, enquanto a chave Q_1 perde a condição de comutação suave

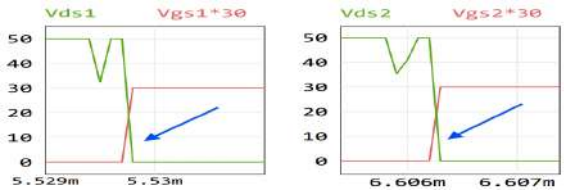
C. Frequência de Chaveamento

A Tabela II apresenta o comportamento do conversor para diferentes frequências de chaveamento, mantendo a razão cíclica fixa em 0,5 e considerando a frequência de ressonância igual a 85 kHz. Observa-se que, para frequências inferiores à ressonância (80 kHz a 85 kHz), o circuito não apresenta a condição de comutação em tensão zero (ZVS). Nessa região, a corrente primária apresenta maiores amplitudes e a tensão de saída é mais elevada, indicando operação em regime predominantemente capacitivo, no qual a corrente encontra-se adiantada em relação à tensão. Essa condição impede a descarga adequada das capacitâncias parasitas das chaves no instante de comutação, podendo ser observada pelas setas azuis do primeiro gráfico da Figura 7. Cabe ressaltar que, neste momento, o acréscimo de 3 nF no capacitor $C1_{ss}$ ainda não foi considerado.

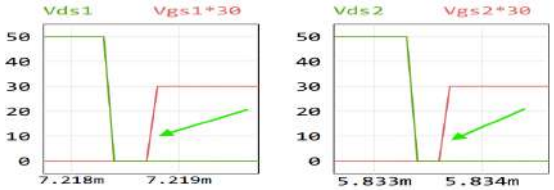
TABELA II
ANÁLISE DO COMPORTAMENTO DO CONVERSOR EM FUNÇÃO DA FREQUÊNCIA DE CHAVEAMENTO

Frequência (kHz)	Correntes (A) Primário	Correntes (A) Secundário	V_{out} (V)	ZVS
80	5,43	1,88	71,6	Não
81	5,25	1,85	70,5	Não
82	4,96	1,80	68,4	Não
83	4,53	1,68	63,6	Não
84	4,14	1,54	58,4	Não
85	3,76	1,42	53,3	Não
— Limiar de ZVS —				
86	3,40	1,29	48,5	Sim
87	3,11	1,18	44,2	Sim
88	2,83	1,07	40,5	Sim
89	2,61	0,98	37,2	Sim
90	2,42	0,90	34,3	Sim

A partir de 86 kHz, ou seja, acima da frequência de ressonância, verifica-se a obtenção da condição de ZVS. Nesse ponto, o circuito passa a operar em regime levemente indutivo, caracterizado pelo atraso da corrente em relação à tensão (setas verdes da Figura 7), favorecendo a descarga natural das capacitâncias parasitas antes do acionamento das chaves. Observa-se também que, à medida que a frequência aumenta, ocorre redução progressiva da corrente primária e da tensão de saída, comportamento típico de conversores ressonantes SS operando acima da frequência natural, onde a impedância equivalente do circuito aumenta e a potência transferida diminui.



(a) Chaves Q_1 e Q_2 operando sem ZVS



(b) Chave Q_1 e Q_2 operando com ZVS

Fig. 7. Análise de um circuito capacitivo x indutivo.

D. Relação da Razão Cílica x Frequência de Chaveamento

Para razão cíclica igual a 0,31 e frequência de chaveamento de 85 kHz (frequência de ressonância), observou-se a perda da condição de ZVS na chave Q_2, enquanto a chave Q_1 manteve a comutação suave com margem adequada. Ressalta-se que essa análise considera a configuração original do capacitor série ($C1_{ss}$), ou seja, sem o acréscimo de 3 nF. A redução da razão cíclica introduz assimetria na condução das chaves, alterando o instante e a amplitude da corrente ressonante no momento da comutação. Em condição de ressonância, tensão e corrente encontram-se praticamente em fase, não havendo defasagem suficiente para garantir corrente adequada no instante crítico de comutação, especialmente sob operação assimétrica imposta

pela razão cíclica. Como consequência, uma das chaves passa a comutar quando a corrente é insuficiente para promover a descarga completa de sua capacitância parasita, comprometendo a obtenção da condição de ZVS.

Ao elevar a frequência para 86 kHz, o circuito passa a operar em regime levemente indutivo, caracterizado pelo atraso da corrente em relação à tensão. Essa defasagem favorece a presença de corrente no instante de comutação, permitindo a descarga das capacitâncias parasitas antes do acionamento das chaves e restabelecendo a condição de ZVS na chave Q_2. Tal resultado evidencia a forte dependência da comutação suave em relação ao perfil temporal da corrente no circuito ressonante e demonstra que a operação acima da frequência de ressonância aumenta a robustez do sistema quanto à obtenção da comutação suave.

VII. Conclusão

O presente artigo apresentou o desenvolvimento e a análise de um circuito ressonante série-série (SS) com acoplamento magnético, alimentado por um inversor meia-ponte, com foco na obtenção da comutação suave (ZVS). A partir das simulações realizadas, foi possível verificar que a condição de ZVS está diretamente relacionada à defasagem entre a tensão aplicada pelo inversor e a corrente do primário, sendo favorecida em regime levemente indutivo. Observou-se, ainda, que variações na razão cíclica e na frequência de chaveamento influenciam diretamente o comportamento do sistema e a obtenção da comutação suave.

Os resultados obtidos contribuem para a compreensão do comportamento do circuito série-série sob diferentes condições operacionais, auxiliando no correto dimensionamento e na definição da estratégia de controle. Entretanto, este estudo apresenta como limitação o fato de os resultados terem sido obtidos exclusivamente por meio de simulações computacionais, sem validação experimental em bancada.

Como continuidade da pesquisa, propõe-se a implementação experimental do conversor, de modo a validar os resultados observados nas simulações. Além disso, sugerem-se estudos voltados à análise da eficiência energética do sistema em diferentes condições de operação, bem como à comparação com outras topologias de compensação, como LCC e série-paralelo (SP), visando ampliar a compreensão sobre o desempenho do sistema.

Referências

[1] Zhen Zhang, Hongliang Pang, Apostolos Georgiadis e Carlo Cecati. “Wireless Power Transfer—An Overview”. Em: *IEEE Transactions on Industrial Electronics* 66.2 (fev. de 2019). Invited Paper, pp. 1044–1058. DOI: 10.1109/TIE.2018.2835378.

[2] Masood Rehman, Sohrab Mirsaeidi, Nursyarizal Mohd Nor, Mohsin Ali Koondhar, Muhammad Ammirul Atiqi Mohd Zainuri, Zuhair Muhammed Alaas, Elsayed Tag-Eldin, Nivin A. Ghamry e M. M. R. Ahmed. “A Review of Inductive Power Transfer: Emphasis on Performance Parameters, Compensation Topologies and Coil Design Aspects”. Em: *IEEE Access* 11 (2023), pp. 144798–144830. DOI: 10.1109/ACCESS.2023.3344041.

[3] Sunkyu Kong, Myunghoi Kim, Kyoungchoul Koo, Seungyoung Ahn, Bumhee Bae e Joungho Kim. “Analytical Expressions for Maximum Transferred Power in Wireless Power Transfer Systems”. Em: *2011 IEEE International Symposium on Electromagnetic Compatibility*. Long Beach, CA, USA, 2011, pp. 379–383. DOI: 10.1109/ISEMC.2011.6038353.

[4] Alexandre Hotz Moret. “Estudo e projeto de um sistema de transferência de energia elétrica sem fio com compensação capacitiva e baseado no transformador de bobinas em espirais planas fracamente acopladas”. Dissertação (Mestrado) - Escola Politécnica. Diss. de mestr. São Paulo, Brazil: Universidade de São Paulo, 2018.

[5] Darryl J. Tschirhart e Praveen K. Jain. “A CLL Resonant Asymmetrical Pulsewidth-Modulated Converter With Improved Efficiency”. Em: *IEEE Transactions on Industrial Electronics* 55.1 (2008), pp. 114–122. DOI: 10.1109/TIE.2007.906176.

[6] Xu-Feng Cheng, Chenyang Liu, Dianlong Wang e Yong Zhang. “State-of-the-Art Review on Soft-Switching Technologies for Non-Isolated DC-DC Converters”. Em: *IEEE Access* 9 (2021), pp. 119235–119249. DOI: 10.1109/ACCESS.2021.3107861.

[7] Altair Engineering. *PSIM Software - Simulação de Eletrônica de Potência e Controle*. 2025. URL: https://web.altair.com/pt-br/psim_hubpage.

Autores







Felipe Coser Acadêmico em Engenharia Elétrica pela Universidade do Oeste de Santa Catarina - UNOESC. Possui técnico em automação industrial pelo Instituto Federal Catarinense - IFC(2018). Técnico em Eletrotécnica pelo Instituto Federal Catarinense - IFC(2020). Atualmente trabalha como supervisor de produção na empresa Baterias Pioneiro.

Jean Patrick Prigol, engenheiro eletricista graduado pela UNOESC (2011), Mestre e doutor em Engenharia Elétrica pela UTFPR (2017 e 2026) . É professor nos cursos de Engenharia da UNOESC Joaçaba-SC com atuação nas áreas de circuitos elétricos, projetos elétricos e sistemas eletrônicos.

Renato Gregolon Scortegagna Engenheiro Eletricista graduado pela UNOESC (2011), Mestre e Doutor em Engenharia Elétrica pela UTFPR (2015 e 2022). É membro da Associação Brasileira de Eletrônica de Potência (SOBRAEP). É professor e coordenador dos Cursos de Engenharia Elétrica e Engenharia de Computação da UNOESC Joaçaba-SC.

Computer Vision Applied to a Robotic Arm : A Case Study

Fernando Barbosa Mota Júnior, Diogo Antonio Castro dos Santos,
Carl Betsa Costa, Raphael Henrique de Oliveira Centenário
João Paulo Carvalho Henriques & Renzo Paranaiba Mesquita
Instituto Nacional de Telecomunicações - Inatel
fernando.b@ges.inatel.br, diogoantonio@ges.inatel.br,
carl.betsa@ges.inatel.br, raphael.h@ges.inatel.br,
joao.paulo@inatel.br & renzo@inatel.br

***Abstract*— Industry 4.0 has stimulated the robotic systems to become smarter and more independent. With the advent and popularization of information systems and IoT devices, robots have had the opportunity to have its capabilities more easily enhanced these days. For exemple, by using Computer Vision and cost-benefit cameras, it is possible to allow pre-programmed robots to execute new functions in a more independent way, once it will be able to better understand its surrounded environment. This work aims to "give eyes" to the Staubli TS60 Arm Robot, a robot that natively does not have embedded Computer Vision capabilities, in order to allow it to execute experiments that go beyond its native capabilities and that make use of Artificial Intelligence techniques. Results demonstrate that integrating algorithms and low cost hardware devices, the Staubli TS60 Arm Robot was successfully able to detect and manipulate objects in a controlled environment, with a 0.4cm margin of error.**

***Index Terms*— Autonomous Robots, Computer Vision, Image Processing, Staubli TS60 Arm Robot**

I.. Introduction

The Industry 4.0 is taking production, revenue, speed, security and integration of the worldwide manufacturing to a higher level than before[1] . This is only possible with the implementation of new technologies and methods of production, all centered in these nine pillars: Big Data, Autonomous Robots, Simulation, Internet of Things, Augmented Reality, Additive Manufacturing, Cybersecurity, Cloud Computing, Horizontal, and Vertical System Integration [2]. This new concept can create smart factories that are connected with the digital world and can take advantage of modern solutions.

On the field of Autonomous Robots (AR), one topic has been crucial to the development of smart robots, that is Computer Vision (CV). CV is a mix of techniques to extract useful information from images, using image processing and pattern recognition [3]. So, when robots are combined with CV, it is possible to create machines that can learn and understand their environment more similar to how a human can, turning them into more flexible and generalist tools. Machines integrated with CV can detect, recognize, and manipulate objects in much more complex and new scenarios. On top of all that, CV enables more security and safety for human-robot collaboration, by allowing the robot to detect people inside a danger zone and adjust its movements in order to prevent accidents.

Alongside the power of CV, the necessary hardware to implement it is becoming cheaper each day, with affordable cameras being capable of taking pictures with enough quality to extract the necessary data for basic tasks that don't require much detail. In addition, there are small minicomputers like the Raspberry Pi [4], that can handle simple image processing, providing some extra support to robots that naturally don't have CV capabilities.

Robotic arms are electronic devices made up of interconnected parts controlled by motors, sensors, and actuators. They are capable of performing movements in different directions and manipulating objects, similar to the movements of the human arm. These devices can be used in various scenarios such as space exploration, industrial areas, healthcare, domestic use, construction, activities with a high risk of unhealthiness (high risk to life), and aid in increasing productivity [5].

Despite of being useful devices to bring more productivity in different fields, some robotic arms do not have cameras embedded on it, as is the case of the Staubli SCARA TS60 Robot. This robot arm is useful to teach students important fundamentals regarding control and automation operations at the National Institute of Telecommunications (Inatel), an institute focused in engineering courses in Minas Gerais, Brazil. However, it was proposed to embed CV on it in order to take it to another level, leveraging its capabilities to perform autonomous operations.

Considering all that, the goal of this work is to build a cost-benefit integration among a basic camera, a minicomputer, and the aforementioned robotic arm using Machine Learning (ML) techniques, aiming to evolve the robot from the principles of Industry 3.0 to those of Industry 4.0. This way, it allows to avail the benefits of Industry 4.0 without the need to buy a new robotic arm.

The rest of this article is organized as follows: Section 2 provides an overview about the CV technique and presents related works that inspired the development of this work. Section 3 gives details about the materials and methods employed to develop this study. Section 4 presents the tests and results of the proposed solution. Finally, in Section 5, conclusions and future directions are presented by the authors.

II.. Background

A. Industry 4.0

Industry 4.0, often referred as the Fourth Industrial Revolution, represents a significant transformation in manufacturing processes, integrating advanced digital technologies to create smart and highly adaptable factories. This evolution is characterized by the incorporation of cyber-physical systems, the Internet of Things (IoT), cloud computing and artificial intelligence, enabling seamless communication between machines, systems and products along the entire value chain. The concept of Industry 4.0 was initially introduced in Germany at the Hannover Messe in 2011 as part of the "Industrie 4.0"

strategy, which aimed to strengthen the competitiveness of German manufacturing through digital transformation. Since then, this approach has been adopted globally, promoting the digitalization and automation of industrial processes to increase efficiency, flexibility and customization in production [6].

B. Computer Vision

CV started in the early 1970s, with the goal of making computers able to describe images as good as humans could do. Although a few decades have passed and a lot of contribution has been made, the goal of the field is basically the same: make computers understand and extract information from images, using algorithms, mathematics, and engineering.

With many years of progress, CV is becoming more powerful each day, resulting in many different areas that are taking advantage of it. For example, in the medical field, there are machine learning models using CV to perform classification and diagnostics of patients [7]. In biology, researchers are transforming the study of insects and other animals with deep learning and CV [8]. Another important examples of CV applications are the self-driving cars relying only on cameras [9], improved control the of of the food quality[10], and the usage of robotics and computer vision for education [11].

Another area that has greatly benefited from CV is, of course, the Industry 4.0, especially in quality-related purposes. For example, industry CV is being implemented as a way to detect printing defects and mistakes [12]. The security has also been enhanced by CV, as machines controlled by optical sensors can work safely alongside humans [13]. There has also been an evolution in Industrial Internet of Things (IIoT) with the advent of CV by enabling an efficient way to collect and analyze data from several machines, ensuring high-quality production with reduced costs [14]. In addition, Industry 4.0 also benefits from the advancements in robotics and automation driven by CV, since it allows a production system with greater adaptability [15].

In order to benefit from CV we need to meet certain requirements [16, 17], such as:

- Implementation of computer vision software since they are needed to allow machines to understand images, which are the basis of a CV powered system. Techniques like Object Detection and Image Segmentation allow the machine to accurately identify, classify, and locate elements within an image;
- Classic models such as Region-based Convolutional Neural Network (R-CNN) [18] and its variations like Fast or Faster R-CNN [19, 20] contributed a lot to the advancement of object detection, while other approaches like You Only Look Once (YOLO) [21] improved real-time applications efficiency. As for image segmentation, techniques like Fully Convolutional Networks (FCN) [22] have enhanced the precision to pixel-level, allowing the differentiation of individual objects from other instances;
- CV does not rely only on software though; it also depends on hardware capable of efficiently performing parallel computing. Graphics Processing Units (GPUs) are usually the main option since their high parallel processing capacity makes them ideal to train neural networks such as Convolutional Neural Networks (CNNs) [23];
- Selecting the appropriate camera is essential once various computer vision tasks demand specific types of visual data. For example, high-resolution cameras are needed for applications such as facial recognition, where it is important to capture detailed facial features with accuracy. On the other hand, lower-resolution cameras are suitable for tasks like monitoring queues, which more focus on broader patterns rather than fine details.

C. Robotics and Robotic Arm

The origin and use of the term “robot” emerged in the 20th century, derived from the Slavic language, specifically Czech, “robota,” which means “work” [24]. Many science fiction films depicted robots imitating human behavior, inspiring young individuals to research and develop robots in real life. Various industries have benefited from the development of robotics, fostering production and eliminating the need for humans to perform dangerous tasks [25].

Robotics covers themes across multiple engineering fields, including mechanics, mechatronics, and especially computing, which are integrated to robots [25]. Robots excel in activities that pose significant risks to human operators, such as repetitive tasks whose quality might be compromised by human effort, and tasks demanding dynamism, precision, and production standardization [25].

Benefits from these applications include implementing collaborative robots (cobots) in workplace safety to reduce physical strain and prevent accidents caused by human fatigue [26], and using robots in healthcare for millimetric surgeries, sparking debates about surgical risks, aesthetic results, physical and emotional recovery, and technical training [27]. Technology’s integration into human lives is reshaping traditional ethical concepts [28].

Although the term “robot” is relatively new, the concept of autonomous machines has existed for centuries. In the 12th century, Ismail al-Jazari, known as the “father of robotics” and modern engineering, authored The Book of Knowledge of Ingenious Mechanical Devices, detailing 50 mechanical inventions [29]. Robotics can be explored through interdisciplinary fields such as mechatronics, mechanics, computing, control, and others.

According to Edione Cirilo Pereira, the classic definition of an industrial robot refers to a standalone machine operating as specific components of the industrial process, performing repetitive activities requiring physical effort and precision. Benefits of this application include the increasing global use of industrial robots and their decreasing acquisition costs, even as human labor costs continue to rise. The continuous evolution of robotic arms, more particularly, has enabled them to play critical roles in addressing extremely dangerous or prohibitive tasks that humans cannot perform [30].

D. Related Works

In order to better understand how CV has been applied in Industry 4.0, specially in robotics, some related works in the literature are presented. They were helpful to show the authors

some great possibilities to be applied in the robotic arm that is the object of this study.

In the work proposed by Malburg et al. [31], the authors showed how object detection can be implemented in factories to complement existing sensors at the task of monitoring pieces, identifying errors and extract useful information to adjust and synchronize the work process. For the simulation it was also used a "learning factory" allied with a mini personal computer Raspberry Pi. The paper compared three models for fast and precise detection: YOLOv3, YOLOv4, and PANet CSPResNeXt50 [32]. Among the models, YOLOv4 presented the best detection accuracy with a hit rate of 97.13% and the worst of them is YOLOv3, with an assertiveness rate of 95.28%.

The objective of the work proposed by Mallick, Pobil, and Cervera [33] was to make the robot recognize objects in disorganized environments and make it pick up the correct object in these same environments. To implement the robotic vision and make the tests, two algorithms were used: RESNET [34] and YOLO. RESNET was more accurate with an Intersection Over Union (IoU) of 0.49 and a Visual Grasping Score (ViG) of 0.94, proving semantic segmentation to be a powerful tool for detecting objects in a messy environment. While YOLO with its bounding boxes achieved an IoU of 0.255 and a ViG of 0.405,these results are significantly lower than those achieved by RESNET. According to the authors, this discrepancy was due to most objects mostly having irregular shapes, which made semantic segmentation-based methods provide better results.

The project proposed by Aguiar [35] involved developing an integrated computer vision system for controlling an industrial robotic arm. The vision system, tasked with classifying three different parts and determining their positions within a work area marked by various colors, was designed to consolidate the image data related to the parts and transmit it to the robotic arm for action. Python was used for programming, with the OpenCV and Socket libraries facilitating image processing and data transmission, respectively. The PDL2 language was employed to script the robotic arm's movements based on the information provided by Python. The hardware setup included a notebook, a smartphone, and the robotic arm's control center. After using Template Matching it achieved an accuracy of 100% and with no false positives. These results demonstrate a significant success in achieving the primary goals of classifying parts, locating them, and directing the robotic arm to pick up the parts and place them in a designated final position based on their type.

This project the focus was on integrating CV with the SCARA TS60 robot, in order to upgrade it from the principles of Industry 3.0 to those of Industry 4.0. This approach not only modernizes the robot but also provides a cheap way of introducing future students to some key concepts of Industry 4.0 since instead of buying a new robot we reutilize the existing one. To achieve this goal were used NumPy for complex mathematical operations and array manipulation, OpenCV to process the images, and YOLOv8 for the object detection. The model was trained using pictures of its environment resulting in a better accuracy.

III.. Materials

In the following subtopics we will briefly explain the main technologies and methods that we used to develop the proposed solution.

A. Python

Python is a high-level open source programming language, with easy syntax and readability. It is for general purpose and supports structured, object-oriented, and functional paradigms. Nowadays it's been widely used for data science and machine learning, mainly because of its powerful libraries, that help and speed up the development of these solutions. In this project, a few Python libraries were considered and used: for data manipulation and test metrics NumPy, Pandas and Scikit-Learn were used; for image manipulation and recognition models OpenCV [36] and Ultralytics YOLOv8 [37].

Numpy is a powerful library which provides n-dimensional arrays and efficient ways to operate on it. It adds powerful data structures to Python and guarantees efficient calculations with arrays and matrices even on high-level mathematical functions. Pandas is an open source Python library which is widely used for working with datasets. It has functions to analyze, explore, and manipulate data. Scikit-Learn, or sklearn, is a popular library to implement ML models for classification, clustering and others, it also provides some simple and efficient tools for data mining. [38]

At the beginning of the project we opted to use OpenCV as the main tool for image recognition. Later, we decided to migrate our system to YOLOv8 once it presented results for real-time item recognition. When we started training using with our own dataset, the YOLOv8 model did a great job on detecting the elements. Even though OpenCV is no longer our main tool, it is still used to perform basic functions and operate the images, while YOLOv8 is responsible for identification-related work.

B. RoboDK

RoboDK is a simulator for industrial robots and robot programming. It allows to test and develop projects outside the production environment. Once it works as a computer software, it is easier to integrate with other software, like Python scripts. In the early stage of this project, RoboDK was used to validate the idea of using information extracted of an image to control the movement of a robot arm.

We started by implementing single-axis movement tests to better understand how the robot arm works. Once we became more familiar with it, we started applying ML to be able to use CV techniques. Later, we stopped using RoboDK, since the real goal was to connect with the real physical robot. [39].

C. Staubli TS60 Arm Robot

The Staubli brand robotic arm, model SCARA TS60 [40] was developed in 2011 and has 4 high-speed axes, with a load limit tolerance of up to 2.0 kg. It has a closed architecture and uses the VAL3 programming language to set configure of its movements. It also has Ethernet ports, fieldbus, digital

inputs and outputs, and serial connections available. Although its movements can be programmed, it can also be manually controlled by a Manual Control Pendant (MCP), a human-machine interface (HMI) that works as a joystick [41].

As mentioned before, this model does not come with an embedded camera and CV capabilities. Therefore, the goal of this project is to literally "give eyes" to it in order to leverage its capabilities and allow it to go beyond its natural functions available.

D. Camera

The camera used in this project was a Web Knup model KP-CW100 [42]. It is a popular, plug and play, and cost-benefit camera that is generally used to make video conferences. It uses an USB 2.0 interface to connect to a computer and has a 1280px x 720px (HD) resolution.

IV.. Experiments and Results

A. Methodology

The Block Diagram depicted by Figure 1 shows the connections of the main elements of the proposed solution.

The solution consists of a camera that monitors a board and the position of the parts located over it. The camera is connected via an USB cord to a laptop, that is responsible to process the image with the YOLOv8 model and make the calculations to find the correct destination position with Python and OpenCV.

These information are passed via Ethernet cable and socket connection to the Robotic Arm Staubli, that will move the parts to a specific position via the Val3 script, the native language of the arm.

For the proposed solution to work, first we had to train the Yolov8; for that, we created our own dataset with images of the piece and the white board. It contains 134 images, split between: 96 for test, 24 for validation and 14 for test.

The specifications of the models are: nano, object detection with Oriented Bounding Boxes, and no prior training. With 300 epochs the model obtained a mAP50 of 0.995 and a mAP50-95 of 0.866. Once the training has finished, OpenCV is used to take pictures with the camera and the now trained Yolov8 can detect the pieces and the board, generating the positions in pixels.

After getting the position in pixels, we calculate the euclidean distance to the center of the board, and then use the real measurement of the board height and width to approximate a proportionality between pixels ad millimeters. To get the final coordinates, we apply the proportionality to the distance.

The communication with the arm robot was made with a socket connection; for that the library socket was imported, the host was set to '192.168.0.240', the port to 7772 and the cable was connected to the J205 physical port of the robot.

For every message, the computer opens a new connection, sends the data and then closes. The data is formatted like this: "GX—Y+++E" where — is the x coordinate in millimeter and +++ is the y coordinate, both must be integer numbers.

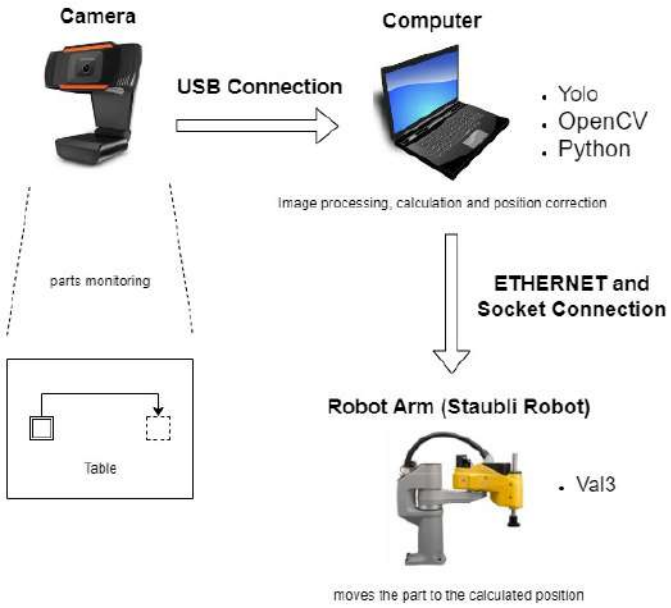


Fig. 1. Block diagram of the proposed solution.

B. Experiments

In order to use CV, a metal structure was built above the robot, where the camera was positioned, as shown in the blue circle at the top of Figure 2. The computer is connected to the camera via a USB cable and connected to the arm robot via an Ethernet cable.

To validate the system, this experiment was conducted dozens times: a part was positioned on the board in an random coordinate, as shown in the Figure 3. After adding the part to the board, the software recognizes the image, identifies the position of the part and sends the information to the software, which processes the image, as shown in Figure 4.

After receiving the coordinates sent by the software, the first movement the robot makes is towards the center of the board, which is its reference point. Once it reaches the center of the board, the robot moves to the coordinates of the part to pick it up (Figure 5).

At this stage, an error in the calculated position was identified, as further way of the center of the board, the bigger the difference between the real position to the calculated, with the max error being 0.4 cm.

After reaching the coordinates of the part, the robot makes the movement to pick it up and transport it to the designated location, previously registered, and makes the movement to release the part.

Once added to the location, the robot finishes the movement and returns to the waiting position, as showed by Figure 6.



Fig. 2. The structure built for the experiment.

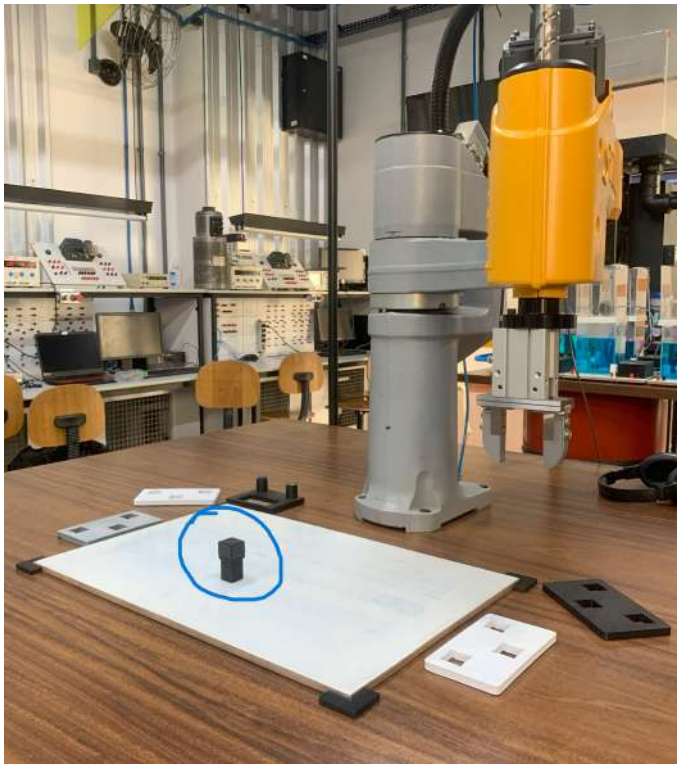


Fig. 3. Part positioned on the board.



Fig. 4. Part and board identified by the software



Fig. 5. Robot moving the part.



Fig. 6. Robot at the waiting position.

C. Results

With the tests completed, we can evaluate the results. First, the YOLO model with our training achieved an mAP50 accuracy of 0.995, higher than the pre-trained YOLOv11, which had an mAP50 of 0.784 [43]. During the experiments, it successfully identified the desired elements. The coordinates calculation showed a maximum error of 0.4 cm on the X-axis and 0.3 cm on the Y-axis, with an average error of 0.25 cm on the X-axis and 0.17 cm on the Y-axis. Communication and control of the robot did not present measurable errors. Even with some stages showing considerable errors, the robotic arm controlled by computer vision was able to locate and manipulate the pieces satisfactorily.

V.. Conclusions

Our main goal with this article was to prove the existence of a cheaper way to automate industrial processes through CV. A mid-range SCARA robot starts at $40.000 and a high-end model goes for more than $60.000 [44]. As our solution could be obtained with these hardware: $30 camera [45] and $500 laptop [46], the total cost of the project is significant lower than a new robot. Once our project proved reliable to detect and interact with pieces placed randomly around the table, our main goal was covered. This proved a more efficient way of implementing CV since robots with built in cameras are far more expensive, and our model can still grant similar results.

As we followed with the project, we had two major challenges: the first one was finding a way to make all parts from our system communicate, since we had no previous experiences with the Staubli system. The second one was improving the calculation of the coordinates, since initially we had a high error margin when detecting the pieces.

Even though we got some good results, at the end we still got a small error on getting the real position for the objects. We found out it was due to our camera distortion. The improvement of this calibration would be one of our suggestions of future work. Detecting people's presence and using it to interrupt the arm's movement, improving the safety of any worker around the arm, would be the second suggestion. Finally, we also suggest that experiments involving detections of multiple parts could be performed.

References

[1] Xun Xu, Yuqian Lu, Birgit Vogel-Heuser, and Lihui Wang. "Industry 4.0 and Industry 5.0—Inception, conception and perception". In: *Journal of Manufacturing Systems* 61 (2021), pp. 530–535. DOI: 10.1016/j.jmsy.2021.10.006.

[2] Gizem Erboz. "How to define industry 4.0: Main pillars of industry 4.0". In: *Managerial trends in the development of enterprises in globalization era* 761 (2017), pp. 761–767.

[3] Victor Wiley and Thomas Lucas. "Computer vision and image processing: a paper review". In: *International Journal of Artificial Intelligence Research* 2.1 (2018), pp. 29–36. DOI: 10.29099/ijair.v2i1.42.

[4] Hirak Dipak Ghael, L Solanki, and Gaurav Sahu. "A review paper on raspberry pi and its applications". In: *International Journal of Advances in Engineering and Management (IJAEM)* 2.12 (2020), pp. 225–227. DOI: 10 . 35629 / 5252 - 0212225227. URL: https : / / ijaem.net/issue_dcp/A%20Review%20Paper%20on%20Raspberry%20Pi%20and%20its%20Applications.pdf.

[5] N. J. Espindola. *"Braço robótico: aplicações gerais e a interação homem-máquina"*. Porto Alegre, 2023.

[6] F. R. Lima and R. Gomes. "Conceitos e tecnologias da Indústria 4.0: uma análise bibliométrica". In: *Revista Brasileira De Inovação* 19 (2020). DOI: 10.20396/rbi.v19i0.8658766.

[7] Eyad Elyan, Pattaramon Vuttipittayamongkol, Pamela Johnston, Kyle Martin, Kyle McPherson, Chrisina Jayne, Mostafa Kamal Sarker, et al. "Computer vision and machine learning for medical image analysis: recent advances, challenges, and way forward." In: *Artificial Intelligence Surgery* 2 (2022). DOI: 10.20517/ais.2021.15.

[8] Toke T. Høye, Johanna Ärje, Kim Bjerge, Oskar L. P. Hansen, Alexandros Iosifidis, Florian Leese, Hjalte M. R. Mann, Kristian Meissner, Claus Melvad, and Jenni Raitoharju. "Deep learning and computer vision will transform entomology". In: *Proceedings of the National Academy of Sciences* 118.2 (2021), e2002545117. DOI: 10.1073/pnas.2002545117.

[9] Bhaskar Barua, Clarence Gomes, Shubham Baghe, and Jignesh Sisodia. "A Self-Driving Car Implementation using Computer Vision for Detection and Navigation". In: *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*. 2019, pp. 271–274. DOI: 10.1109/ICCS45141.2019.9065627.

[10] Vijay Kakani, Van Huan Nguyen, Basivi Praveen Kumar, Hakil Kim, and Visweswara Rao Pasupuleti. "A critical review on computer vision and artificial intelligence in food industry". In: *Journal of Agriculture and Food Research* 2 (2020), p. 100033. DOI: 10.1016/j.jafr.2020.100033.

[11] Aphrodite Sophokleous, Panayiotis Christodoulou, Lefteris Doitsidis, and Savvas A. Chatzichristofis. "Computer Vision Meets Educational Robotics". In: *Electronics* 10.6 (2021). DOI: 10.3390/electronics10060730.

[12] Javier Villalba-Diez, Daniel Schmidt, Roman Gevers, Joaquín Ordieres-Meré, Martin Buchwitz, and Wanja Wellbrock. "Deep Learning for Industrial Computer Vision Quality Control in the Printing Industry 4.0". In: *Applied Sciences* (2019). DOI: 10.3390/s19183987.

[13] Jaiprakash Bhamu Pramod Kumar Dharmendra Singh. "Machine Vision in Industry 4.0: Applications, Challenges and Future Directions". In: *Machine Vision for Industry 4.0*. 2022, pp. 263–284.

[14] M. Tzampazaki, C. Zografos, E. Vrochidou, and G.A. Papakostas. "Machine Vision—Moving from Industry 4.0 to Industry 5.0". In: *Applied Sciences* (2024). DOI: 10.3390/app14041471.

[15] Jaimin G. Pandya and Nirav P. Maniar. “Computer Vision-Guided Human–Robot Collaboration for Industry 4.0: A Review”. In: *Recent Advances in Mechanical Infrastructure*. Ed. by Ajit Kumar Parwani, PL. Ramkumar, Kumar Abhishek, and Saurabh Kumar Yadav. Singapore: Springer Nature Singapore, 2022, pp. 147–155.

[16] Xin Feng, Youni Jiang, Xuejiao Yang, Ming Du, and Xin Li. “Computer vision algorithms and hardware implementations: A survey”. In: *Integration* 69 (2019), pp. 309–320. DOI: 10.1016/j.vlsi.2019.07.005.

[17] Abirami Vina. *“Computer Vision Cameras and Their Applications”. https://www.ultralytics.com/blog/computer-vision-cameras-and-their-applications. Accessed: 10-17-2024*. 2024.

[18] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation”. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 2014, pp. 580–587. DOI: 10.1109/CVPR.2014.81.

[19] Ross Girshick. “Fast R-CNN”. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 1440–1448. DOI: 10.1109/ICCV.2015.169.

[20] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.6 (2017), pp. 1137–1149. DOI: 10.1109/TPAMI.2016.2577031.

[21] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. “You Only Look Once: Unified, Real-Time Object Detection”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 779–788. DOI: 10.1109/CVPR.2016.91.

[22] Jonathan Long, Evan Shelhamer, and Trevor Darrell. “Fully convolutional networks for semantic segmentation”. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015, pp. 3431–3440. DOI: 10.1109/CVPR.2015.7298965.

[23] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324. DOI: 10.1109/5.726791.

[24] Andre Luiz Gioia Morrell, Alexander Charles Morrell-Junior, Allan Gioia Morrell, Jose Mauricio Freitas Mendes, Francisco Tustumi, LUIZ GUSTAVO DE-OLIVEIRA-E-SILVA, and Alexander Morrell. “The history of robotic surgery and its evolution: when illusion becomes reality”. In: *Revista do Colégio Brasileiro de Cirurgiões* 48 (2021), e20202798. DOI: 10.1590/0100-6991e-20202798.

[25] Valdemir Carrara. “Introdução à robótica industrial”. In: *São José dos Campos: Instituto Nacional de Pesquisas Espaciais-Inpe* (2015).

[26] Elcio Shinji Yamada. “Cobots-Impacto na segurança do trabalhador”. In: (2024).

[27] Yasmin Mendes Rodrigues dos Santos, Ciro Ferrari de Souza, Tainan Gomes Ferreira, Stephany Bellelis, and Bárbara San de Carvalho. “ASPECTOS ÉTICOS DA ROBÓTICA NA PRÁTICA MÉDICA CONTEMPORÂNEA”. In: *Revista Ibero-Americana de Humanidades, Ciências e Educação* 10.12 (2024), pp. 826–840. DOI: 10.51891/rease.v10i12.17280.

[28] Anna Puzio. “Not relational enough? Towards an eco-relational approach in robot ethics”. In: *Philosophy & Technology* 37.2 (2024), p. 45.

[29] Aldo Brassetti, Alberto Ragusa, Francesco Tedesco, Francesco Prata, Loris Cacciatore, Andrea Iannuzzi, Alfredo Maria Bove, Umberto Anceschi, Flavia Proietti, Simone D’Annunzio, et al. “Robotic surgery in urology: history from PROBOT® to HUGOTM”. In: *Sensors* 23.16 (2023), p. 7104. DOI: 10.3390/s23167104.

[30] Edione Cirilo Pereira. “Desenvolvimento de um Robô Manipulador SCARA”. In: (2023).

[31] Lukas Malburg, Manfred-Peter Rieder, Ronny Seiger, Patrick Klein, and Ralph Bergmann. “Object Detection for Smart Factory Processes by Machine Learning”. In: *Procedia Computer Science* 184 (2021). The 12th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops, pp. 581–588. DOI: 10.1016/j.procs.2021.04.009.

[32] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. “CSPNet: A New Backbone That Can Enhance Learning Capability of CNN”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. June 2020.

[33] Arijit Mallick, Angel P. del Pobil, and Enric Cervera. “Deep Learning Based Object Recognition for Robot Picking Task”. In: *Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication*. IMCOM ’18. Langkawi, Malaysia: Association for Computing Machinery, 2018. DOI: 10.1145/3164541.3164628.

[34] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep Residual Learning for Image Recognition”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[35] M. J. L. S. Aguiar. *“Sistema De Classificação E Separação De Peças Por Visão Computacional Usando O Braço Robótico Comau Ns 16 1.65”*. Recife, 2019.

[36] Gary Bradski. *The OpenCV Library*. 2000. URL: https://opencv.org/.

[37] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. *Ultralytics YOLOv8*. Version 8.0.0. 2023. URL: https://github.com/ultralytics/ultralytics.

[38] Charles R. Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[39] Albert Nubiola and Rafael Romero. *RoboDK: Software for Simulation and Offline Programming of Industrial Robots*. Available at: https://robodk.com/. 2023.

[40] Staubli. *Datasheet Staubli TS60*. URL: https://robodk.com/robot/Staubli/TS60-FL-400 (visited on 12/10/2024).

[41] João Paulo Carvalho Henriques, Egidio Raimundo Neto, João Pedro Magalhães de Paula Paiva, Letícia Carneiro de Souza, Alexandre Baratella Lugli, Flávio Miranda Florentino, and Giuliano Clayton de Andrade Carvalho. "Trajectory Generation Using RoboDK for a Staubli SCARA TS 60 Robot". In: *2023 11th International Conference on Control, Mechatronics and Automation (ICCMA)*. 2023, pp. 121–126. DOI: 10.1109/ICCMA59762.2023.10374649.

[42] Knup KP-CW100. *Datasheet KP-CW100*. URL: https://sulcell.com.br/index.php/webcam-hd-720p-c-microfone-usb-2-0-kp-cw100.html (visited on 12/10/2024).

[43] Ultralytics. *Yolo/OBB*. URL: https://docs.ultralytics.com/tasks/obb/#visual-samples (visited on 12/12/2024).

[44] StandarBots. *standardbots/Scara*. URL: https://standardbots.com/blog/scara-robot-price (visited on 12/12/2024).

[45] Logitech. *lgitech/notebook*. URL: https://www.logitechstore.com.br/camera-webcam-hd-logitech-c270/ (visited on 12/12/2024).

[46] Acer. *acer/notebook*. URL: https://br-store.acer.com/notebook-acer-a515-57-58w1-ci512450h-8gb-256gb-ssd-gutta-gold-15-6-fhd-nx-kngal-001-p/p (visited on 12/12/2024).

AUTHORS



Fernando Mota is an undergraduate student in Software Engineering at INATEL - Instituto Nacional de Telecomunicações. Served as a teaching assistant for the Object-Oriented Programming course and is the founder of INATEL's game development group, OctoDevs. Has interests in the fields of Technology, Game Development, and Programming.



Diogo Castro is a Telecommunications Management Technologist (Inatel) and an undergraduate student in Software Engineering (Inatel). He works as a Software Developer at Scoder Tech Studio. He is interested in software development, data analysis, and machine learning.



Carl Betsa is an undergraduate student in Software Engineering at INATEL - Instituto Nacional de Telecomunicações. Has interests in the fields of Artificial Intelligence, Game Development, and Data Science.



Raphael Henrique de Oliveira Centenário is an undergraduate student in Software Engineering at INATEL - Instituto Nacional de Telecomunicações. Internship at CDG (Coders, Developers, and Gamers) Hub - INATEL. Has interests in the fields of Game Development, Data Science and Web Development.



Renzo Paranaíba Mesquita is a Computer Engineer (INATEL), Specialist and Master in Computing (UNIFEI) and PhD candidate in Production Engineering (UNIFEI). He works as a professor and coordinator of the Computer Engineering and Software Engineering courses at the National Institute of Telecommunications (INATEL). He is interested in the areas of Software Development, Data Analysis and Computational Simulation.

Estudo para Desenvolvimento de um Robô para a Categoria Perseguidor de Linha

Alexandre B. Lugli, Eugênio M. J. da Silva, Egídio R. Neto, Giovanna G. Fernandino
Luiz Fernando D. R. dos Santos, Maria Fernanda P. Castro, Wanderson E. Saldanha
Instituto Nacional de Telecomunicações - Inatel
baratella@inatel.br, eugenio@inatel.br, egidioneto@inatel.br, giovanna.gama@gea.inatel.br
luizf.santos@inatel.br, maria.castro@gea.inatel.br, wsaldanha@inatel.br

***Abstract*—This work presents the study, development, and evaluation of an autonomous robot designed for the Follow Line - Chaser category, a competition in which two robots simultaneously follow a symmetric track guided by a high-contrast line. The project integrates electronics, mechanics, and software , employing optical sensors, high-speed motors, suction-based downforce systems (fans) , and a high-performance microcontroller. The robot's structure is based on a dual-layer PCB chassis , optimizing weight and rigidity. Track tests and simulated matches demonstrated stability, fast curve response, and high acceleration capability. Results highlight the relevance of efficient control algorithms (PID) and robust line-detection strategies to maximize performance in their respective competitions.**

***Index Terms*—Mobile robotics, Follow Line - Chaser, PID control**

***Resumo*—Este trabalho apresenta o estudo, desenvolvimento e avaliação de um robô autônomo projetado para competir na categoria Perseguidor de Linha, uma modalidade na qual dois robôs percorrem simultaneamente uma pista simétrica guiada por uma linha de alto contraste. O projeto integra eletrônica, mecânica e software, empregando sensores ópticos, motores de alta rotação, ventoinhas para aumento da força normal e um microcontrolador de alto desempenho. A construção baseia-se em uma PCB estrutural com duas placas integradas, otimizando peso e rigidez. Testes em pista demonstraram estabilidade, rápida resposta a curvas e elevada capacidade de aceleração. Os resultados reforçam a importância de algoritmos eficientes de controle (PID) e estratégias robustas de leitura de linha para maximizar desempenho em competições da categoria.**

***Palavras chave*—robótica móvel, Perseguidor de Linha, controle PID**

I. Introdução

A robótica móvel tem sido amplamente utilizada como plataforma de experimentação em áreas como controle, percepção sensorial e sistemas embarcados [1]. Em ambientes educacionais, competições de robótica oferecem um contexto prático para aplicação desses conceitos, permitindo que estudantes desenvolvam soluções que integram eletrônica, mecânica e software em sistemas autônomos.

Entre essas modalidades, destaca-se a categoria Perseguidor de Linha, na qual dois robôs percorrem simultaneamente uma pista simétrica guiada por uma linha de alto contraste. Diferentemente de seguidores de linha convencionais, essa categoria impõe desafios adicionais relacionados ao controle em altas velocidades, estabilidade dinâmica e detecção rápida da linha, uma vez que o objetivo da disputa é alcançar o robô adversário ao longo do percurso.

Para atingir bom desempenho nesse cenário, é necessário que o robô seja capaz de interpretar rapidamente os dados dos sensores, aplicar correções eficientes na trajetória e manter estabilidade mesmo em curvas acentuadas. Dessa forma, aspectos como a qualidade da leitura dos sensores, o desempenho dos algoritmos de controle e o projeto mecânico do robô tornam-se fatores determinantes para o desempenho competitivo.

Este trabalho tem como objetivo o estudo e desenvolvimento de um robô autônomo para a categoria Perseguidor de Linha, abordando desde os fundamentos teóricos até a aplicação prática e a análise dos resultados obtidos. O projeto visa aprimorar o desempenho em competições de robótica e contribuir para a consolidação de conhecimentos em sensores, controle e dinâmica de movimento.

Este artigo está estruturado em cinco seções principais. Seção I (Introdução), seção II (Fundamentação teórica), seção III (Materiais e Aplicação Prática), seção IV(Testes e Resultados) e finalizando com a seção V (Conclusão).

II. Fundamentação Teórica

A. Regras

A categoria Perseguidor de Linha segue regulamentos definidos pela RoboCore® [2], que impõem restrições dimensionais e operacionais aos robôs participantes. Entre essas restrições, destacam-se os limites máximos de 20 cm de largura, 30 cm de comprimento e 20 cm de altura, além da obrigatoriedade de acionamento remoto sem contato físico direto.

As competições são realizadas em uma pista composta por um manto de borracha preto com uma linha branca de alto contraste, o que define diretamente o tipo de sensor utilizado e o processo de calibração do sistema de detecção.

Essas restrições influenciam diretamente o projeto do robô, impactando escolhas mecânicas, eletrônicas e de software, especialmente no que diz respeito à detecção da linha, estabilidade e controle em altas velocidades.

A categoria Perseguidor de Linha segue um conjunto de regulamentos específicos estabelecidos pela RoboCore® , que influenciam diretamente as decisões de projeto. As partidas são disputadas entre duas equipes, cada uma composta por até três participantes, embora apenas dois possam manipular o robô na área de competição.

Dimensionalmente, o robô deve respeitar as medidas máximas de 20 cm de largura, 30 cm de comprimento e 20

cm de altura, parâmetros que limitam o tamanho do chassi e a distribuição dos componentes. Além disso, é obrigatório o acionamento remoto, sem contato físico direto, o que exige a integração de um módulo de partida para a participação na competição [2].

Como apresentado na figura 2, a pista utilizada em competições brasileiras é composta por um tapete de borracha preto com uma linha branca de alto contraste. Esse padrão de cores determina o tipo de sensor a ser empregado e orienta o processo de calibração. Assim, as regras não apenas definem o formato da disputa, mas também direcionam as escolhas eletrônicas, mecânicas e de software do robô, afetando diretamente sua arquitetura e seu desempenho.

B. Robô

O robô da categoria Perseguidor de Linha pode ser compreendido como um sistema mecatrônico composto por três subsistemas principais[3] — eletrônica, mecânica e software — que atuam de forma integrada para garantir precisão, estabilidade e velocidade. Em uma modalidade em que dois robôs percorrem trajetórias simétricas simultaneamente, a eficiência na detecção da linha e na execução dos comandos de movimento torna-se fundamental para maximizar o desempenho competitivo.

b.1) Eletrônica

A eletrônica constitui o núcleo funcional do robô, responsável pela aquisição de dados, processamento e atuação. O microcontrolador executa a leitura dos sensores e aplica as rotinas de controle que determinam a velocidade e a trajetória. Em competições de Perseguidor de Linha, a principal fonte de percepção é o array de sensores ópticos infravermelhos, que identifica as diferenças de refletância entre a linha branca e o fundo preto da pista, permitindo a localização precisa da trajetória.

A atuação é realizada por motores controlados por pontes H, que modulam a velocidade e o sentido de rotação por meio de PWM. A etapa de alimentação deve ser cuidadosamente projetada para garantir estabilidade mesmo sob altas correntes, especialmente em robôs que utilizam sistemas auxiliares como ventoinhas para aumento de força normal.

Conforme exigido pelas regras[2], o robô deve integrar um módulo de partida remoto, responsável por garantir segurança e sincronização nas disputas. Assim, o subsistema eletrônico não apenas sustenta o funcionamento básico do robô, mas também atende aos requisitos normativos da categoria.

b.2) Mecânica

A estrutura mecânica é projetada para oferecer rigidez, leveza e estabilidade, garantindo que o robô suporte acelerações elevadas sem perda de controle. O chassi deve posicionar os sensores na região frontal, próximos ao solo, assegurando leituras precisas mesmo em variações de refletividade ou pequenas irregularidades da pista.

Os motores, pneus e a geometria escolhidos influenciam diretamente a capacidade de realizar curvas rápidas e de manter tração em altas velocidades. Os pneus devem maximizar a aderência, enquanto a disposição do centro de gravidade reduz oscilações e melhora a repetibilidade do comportamento.

Sistemas de aumento de força normal, como ventoinhas de alta rotação, podem ser empregados para elevar a aderência ao piso sem violar as regras. Quando bem integrados, esses mecanismos contribuem significativamente para a estabilidade em curvas, possibilitando acelerações mais agressivas e maior controle sobre a trajetória.

b.3) Software

O software é responsável por transformar as leituras dos sensores em ações de controle, definindo o comportamento dinâmico do robô. O algoritmo de navegação baseia-se na detecção da posição da linha por meio de um array de sete sensores ópticos infravermelhos.

A posição da linha é estimada a partir de uma média ponderada das leituras dos sensores, na qual cada elemento do array recebe um peso proporcional à sua posição, variando de 0 a 3000. O valor de referência adotado corresponde ao centro do array, definido como $x_{\text{ref}} = 1500$.

$$e(t) = x_{\text{medido}}(t) - x_{\text{ref}} \tag{1}$$

O erro de posição, definido na Eq. 1, representa o desvio lateral do robô em relação ao centro da linha e é utilizado como entrada para o sistema de controle.

Para correção da trajetória, foi implementado um controlador do tipo Proporcional-Derivativo (PD), cujo sinal de controle é dado por:

$$u(t) = K_p e(t) + K_d \frac{de(t)}{dt} \tag{2}$$

O sinal de controle descrito na Eq. 2 é aplicado como diferencial de velocidade entre os motores, permitindo ajustes dinâmicos da trajetória durante o movimento.

O termo integral não foi utilizado, pois, em aplicações de alta velocidade, sua inclusão pode introduzir acúmulo de erro e instabilidade. Além disso, não foi observado erro estacionário significativo que justificasse sua utilização.

Os ganhos K_p e K_d foram ajustados empiricamente por meio de testes em pista, buscando um compromisso entre rapidez de resposta e estabilidade. O ganho proporcional foi responsável pela intensidade das correções, enquanto o ganho derivativo contribuiu para a redução de oscilações, resultando em uma resposta mais suave e estável.

III. Materiais e Aplicação Prática

A. Materiais utilizados

O robô desenvolvido, que pode ser visto na figura 1, utiliza como base estrutural a própria placa de circuito impresso (PCB), que funciona simultaneamente como chassi e suporte para os componentes eletrônicos. A arquitetura é composta por duas PCBs integradas:

- uma placa principal, destinada aos sistemas de processamento, potência e acionamento,
- e uma placa frontal dedicada exclusivamente ao array de sensores ópticos.



Fig. 1. FlashBull - robô Perseguidor de Linha.

Essa configuração proporciona maior rigidez ao conjunto, reduz peso e facilita a manutenção, além de permitir o posicionamento ideal dos sensores próximos ao solo, garantindo leituras mais precisas durante a navegação.

A placa principal abriga o microcontrolador Teensy 3.6, escolhido pela elevada capacidade de processamento e pelo grande número de interfaces disponíveis, características essenciais para leitura rápida dos sensores e execução eficiente das rotinas de controle. O acionamento dos motores é realizado pela ponte H TB6612FNG, enquanto os sistemas auxiliares — como as ventoinhas — são alimentados a partir de uma bateria LiPo 3S (450 mAh).

A alimentação do robô é distribuída por meio de um buck converter, responsável por fornecer 9 V aos motores, e dois reguladores lineares que estabilizam as tensões de 5 V e 3.3 V para os demais módulos. Os motores utilizados são do tipo N20 com encoder, originalmente de 6 V, mas alimentados em 9 V para aumentar a velocidade e o torque, garantindo aceleração rápida e resposta eficiente durante a perseguição.

Para navegação, foi utilizada uma matriz frontal de sete sensores QRE1113, dispostos de forma compacta e posicionados a poucos milímetros do solo. Esse arranjo permite monitoramento contínuo da linha de contraste da pista, fornecendo dados estáveis ao controlador mesmo em situações de alta velocidade ou variações de iluminância.

A mobilidade do robô é complementada por rodas produzidas em impressão 3D (PLA), revestidas com borracha para aumentar a aderência. Além disso, o robô conta com quatro ventoinhas de 8000 kV, acionadas por uma ESC, cuja função é gerar força normal adicional e melhorar a tração em curvas de alta velocidade — recurso essencial para competições de perseguidor, onde a estabilidade dinâmica influencia diretamente o desempenho.

O robô incorpora ainda um módulo de partida infravermelho, compatível com o protocolo padrão da categoria, possibilitando o acionamento remoto obrigatório em competições da RoboCore®.

O software foi desenvolvido na Arduino IDE, utilizando bibliotecas da Pololu para leitura e calibração dos sensores, além de funções específicas para controle dos motores, leitura dos encoders e gerenciamento dos parâmetros de navegação.

B. Aplicação Prática e Testes

A etapa prática iniciou-se com a montagem completa do robô a partir das duas PCBs estruturais. A placa principal recebeu o microcontrolador Teensy 3.6, a ponte H TB6612FNG, os reguladores de tensão, a bateria e o conjunto de ventoinhas, enquanto a placa frontal foi dedicada exclusivamente aos sete sensores QRE1113. As duas placas foram fixadas por um suporte rígido e interligadas por um cabo flat, garantindo organização interna e reduzindo interferências elétricas.

Após a montagem física, procedeu-se à integração elétrica dos motores, sensores, ventoinhas e reguladores. Foram verificadas as tensões de cada trilha, garantindo que os motores recebessem 9 V regulados e que os demais módulos operassem de forma estável em 5 V e 3.3 V. Em seguida, o módulo de partida infravermelho foi acoplado ao sistema principal, permitindo o acionamento remoto conforme as regras da competição.

No software, foram implementadas rotinas essenciais para o funcionamento do robô, incluindo leitura dos sensores, calibragem automática, controle dos motores via PWM e comunicação com o módulo de partida. A escolha da Teensy 3.6 permitiu leituras rápidas e processamento eficiente das correções de trajetória.

Essa etapa concluiu a integração mecatrônica do robô, preparando-o para os testes de desempenho descritos na Seção IV.

IV. Testes e Resultados

A etapa de testes foi conduzida com o objetivo de validar o desempenho do robô no contexto competitivo do Perseguidor de Linha, avaliando individualmente cada subsistema e, posteriormente, o comportamento integrado em condições de disputa simulada. Os testes foram realizados em pista padronizada conforme as especificações da RoboCore®, utilizando manto de borracha preto e linha branca de alto contraste.

A. Procedimentos de Teste

Os testes foram conduzidos com o objetivo de validar o desempenho do robô no contexto competitivo da categoria Perseguidor de Linha. Os ensaios foram realizados em pista padronizada, conforme ilustrado na Fig. 2, composta por manto de borracha preto com linha branca de alto contraste.



Fig. 2. Pista Perseguidor de Linha - IronCup 2025

A matriz frontal de sete sensores QRE1113 foi previamente calibrada utilizando bibliotecas da Pololu, definindo limites de leitura adequados às condições da pista. Foram avaliadas situações típicas de operação, como variações de iluminância, microdesvios laterais e aproximações rápidas da linha em curvas, com o objetivo de garantir robustez nas leituras mesmo em altas velocidades.

Após a validação dos sensores, foram realizados testes de integração do sistema completo. Nessas etapas, foram analisados trechos de aceleração, curvas acentuadas, mudanças bruscas de direção e a capacidade de recuperação após pequenas perturbações.

Por fim, foram conduzidas simulações de disputa, nas quais dois robôs percorriam trajetórias simultâneas, permitindo avaliar aspectos como aproximação, estabilidade e comportamento dinâmico em condições próximas ao ambiente competitivo real.

B. Resultados Obtidos

Os testes foram realizados em pista padronizada, por meio de execuções repetidas, com o objetivo de avaliar o comportamento dinâmico do robô em diferentes condições de operação. A análise considerou, principalmente, o impacto do sistema de ventoinhas na estabilidade e na aderência.

Observou-se que a utilização das ventoinhas aumentou a força normal aplicada ao solo, reduzindo derrapagens e melhorando a estabilidade em curvas. Sem esse sistema, o robô apresentou maior incidência de microdesvios e instabilidade em mudanças rápidas de direção.

O controlador PD, ajustado empiricamente, proporcionou correções rápidas e recuperação eficiente da trajetória, com baixa tendência a oscilações. Nos testes completos, o robô manteve-se majoritariamente centrado na linha, apresentando comportamento consistente mesmo sob pequenas variações de iluminação e irregularidades da pista.

Em simulações de disputa, verificou-se capacidade consistente de aproximação em relação ao adversário, associada à elevada aceleração e estabilidade em altas velocidades. O desempenho observado está diretamente relacionado à estratégia de controle adotada, baseada na estimativa contínua da posição da linha e na aplicação de um controlador PD, conforme descrito na Seção II.

O desempenho também foi validado em ambiente competitivo, com obtenção de colocações de destaque.

Ressalta-se que medições quantitativas, como tempo de volta e velocidade média, não foram realizadas nesta etapa, sendo propostas como continuidade do trabalho.

V. Conclusão

Este trabalho apresentou o desenvolvimento de um robô autônomo para a categoria Perseguidor de Linha, contemplando a integração de sistemas mecânicos, eletrônicos e de controle voltados à operação em altas velocidades. Foram exploradas soluções de projeto com foco em desempenho, destacando-se o uso de um chassi em placa de circuito impresso (PCB), a aplicação de ventoinhas com controle via ESC para aumento da força normal e a adoção de um controlador PD baseado na estimativa contínua da posição da linha.

Os testes realizados indicaram que o sistema foi capaz de manter estabilidade e seguir a trajetória proposta de forma consistente, mesmo sob condições desafiadoras, como curvas acentuadas e variações de iluminação. A utilização das ventoinhas contribuiu para o aumento da aderência, enquanto o controlador PD possibilitou respostas rápidas e suaves na correção da trajetória.

O desempenho do robô também foi validado em ambiente competitivo, evidenciando sua robustez e adequação ao contexto da categoria. Esses resultados demonstram a eficácia das soluções adotadas e reforçam o potencial do projeto como plataforma para estudos em robótica móvel e sistemas embarcados. O projeto também apresenta caráter educacional relevante, contribuindo para o aprendizado prático de conceitos de controle, instrumentação e robótica móvel.

Como limitações, destaca-se a ausência de medições quantitativas detalhadas, como tempo de volta e velocidade média, devido à indisponibilidade do protótipo na fase final do trabalho. Como continuidade, propõe-se a instrumentação do sistema para aquisição desses dados, bem como a investigação de diferentes estratégias de controle e otimização de parâmetros visando ganhos adicionais de desempenho.

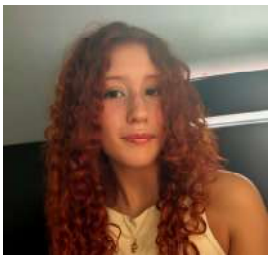
Referências

[1] R. Siegwart, I.R. Nourbakhsh e D. Scaramuzza. *Introduction to Autonomous Mobile Robots, second edition*. MIT Press, 2011.

[2] RoboCore®. *Regras Perseguidor de Linha*. URL: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Perseguidor+de+Linha.pdf (acesso em 03/11/2025).

[3] Devdas Shetty e Richard A Kolk. *Mechatronics system design*. CL-Engineering, 2010.

Autoras



Giovanna Gama Fernandino é estudante do 3º período de Engenharia de Controle e Automação no Instituto Nacional de Telecomunicações (INATEL). Integra a equipe RobotBulls, atuando na categoria Hockey, e atualmente é bolsista de iniciação científica.



Maria Fernanda Paula Castro é graduanda do 4º período de Engenharia de Controle e Automação no Instituto Nacional de Telecomunicações (INATEL). Membro da equipe Robotbulls, tem se destacado nas categorias Seguidor de Linha e Perseguidor de Linha, além de atuar como bolsista de Iniciação Científica no Laboratório RobotBulls.

Estudos de Desenvolvimento de um Robô para a Categoria Seguidor de Linha Pro

Alexandre B. Lugli, Caroline Ferreira de Souza, Eugênio M. J. da Silva, Egídio R. Neto, Luiz Fernando D. R. Santos, Pedro Paulo P. J., Wanderson E. Saldanha
Instituto Nacional de Telecomunicações (INATEL)
baratella@inatel.br, caroline.souza@ges.inatel.br, eugenio@inatel.br, egidio.neto@inatel.br, luizf.santos@inatel.br, pedro.junqueira@gec.inatel.br, wsaldanha@inatel.br

***Resumo*—O presente artigo apresenta o estudo, desenvolvimento e validação de um robô autônomo competidor da categoria Seguidor de Linha Pro. O trabalho aborda desde os fundamentos teóricos da modalidade, incluindo sensores ópticos de refletância, controladores PD e técnicas para aumento de aderência, até a implementação prática do robô FlashBull, desenvolvido pela equipe RobotBulls do INATEL. A metodologia empregada integra arquitetura eletrônica compacta, chassi leve, sistema de mapeamento de pista e ventoinhas para geração de downforce. Diversos testes experimentais foram conduzidos para avaliar o desempenho do controle de velocidade, a eficiência do sistema de sucção e o impacto do mapeamento automático no tempo total de volta. Os resultados mostram reduções significativas no tempo de percurso e maior estabilidade em curvas, evidenciando a eficácia das soluções aplicadas. Por fim, o artigo apresenta limitações identificadas e propostas de melhorias para futuras versões do robô, reforçando o potencial de evolução contínua dentro da categoria Pro.**

***Index Terms*—Seguidor de Linha Pro, Controle PD, Downforce, RobotBulls, Mapeamento.**

I. Introdução

O desenvolvimento de sistemas robóticos autônomos constitui um campo de pesquisa e aplicação prática de crescente relevância na Engenharia, servindo como catalisador para a integração de conhecimentos em controle, programação e eletrônica. Nesse cenário, competições de robótica emergem como ambientes cruciais para a validação de projetos e o aprimoramento tecnológico. A equipe Inatel RobotBulls, formada por membros de formação interdisciplinar, atua com destaque nacional nessas competições, posicionando protótipos como o FlashBull entre os melhores do país no portal oficial da RoboCore.

A categoria Seguidor de Linha Pro impõe o desafio específico de completar, no menor tempo possível, um circuito desconhecido delimitado por uma linha branca sobre uma superfície de borracha preta. Para obter êxito nesse cenário, o principal problema técnico do projeto reside em garantir uma navegação autônoma precisa, superar a perda de estabilidade em curvas de alta velocidade e executar um controle de trajetória rigoroso sob dinâmicas não lineares. Para compreender as soluções aplicadas a esses desafios, a seção a seguir detalha os princípios teóricos que fundamentam a arquitetura de controle e o projeto mecânico adotados.

A. A Categoria Seguidor de Linha Pro

A modalidade Seguidor de Linha Pro, promovida pela Robocore, estabelece um desafio específico na área de navegação autônoma e controle de velocidade. O principal objetivo da competição consiste em completar um circuito predefinido no menor tempo possível. O critério de vitória é determinado pela capacidade do robô de seguir integralmente uma linha guia e cessar o movimento precisamente no ponto final de marcação da pista, momento em que o tempo oficial de prova é registrado.

A pista é caracterizada por ser confeccionada em borracha fosca de cor preta, sobre a qual é aplicada uma linha de cor branca que serve como referência de trajeto para o robô. Um elemento fundamental do desafio reside no fato de que o layout exato do circuito é mantido em sigilo até o dia da competição, exigindo que os sistemas de controle do robô sejam robustos e adaptáveis a diferentes geometrias.

Tecnicamente, o robô seguidor de linha opera por meio de um sistema de percepção baseado em sensores infravermelhos. Estes sensores são empregados para medir o índice de refletividade da superfície, permitindo que o sistema eletrônico diferencie a linha branca (alta refletividade) da pista preta (baixa refletividade). O controle de navegação é executado por meio de um algoritmo de correção de trajetória que, baseado nas leituras dos sensores, modula a velocidade e a direção dos motores de forma contínua, garantindo que o robô permaneça sobre a linha guia mesmo em altas velocidades.

O presente artigo tem por objetivo apresentar os estudos de desenvolvimento e a arquitetura de controle implementada no robô Inatel Robotbulls, visando a otimização de seu desempenho na categoria Seguidor de Linha Pro. A Figura 1 apresenta um exemplo de percurso na pista do seguidor. [5]

II. Fundamentação Teórica

A. Regras da Categoria Seguidor de Linha Pro

A categoria Seguidor de Linha Pro, regulamentada pela RoboCore, consiste em percorrer autonomamente um circuito delimitado por uma linha branca sobre manta preta, realizando três tentativas sem intervenção externa e sem modificações de hardware ou firmware durante as rodadas. O trajeto apresenta retas, curvas com diferentes raios e cruzamentos de 90°, acompanhados de marcações laterais que permitem a leitura de trechos pelo robô. As dimensões máximas permitidas são 250 mm × 250 mm × 200 mm, sem restrição de peso. [5]

B. Sensores Ópticos de Refletância

Os sensores ópticos de refletância utilizam um emissor infravermelho e um fototransistor para distinguir regiões

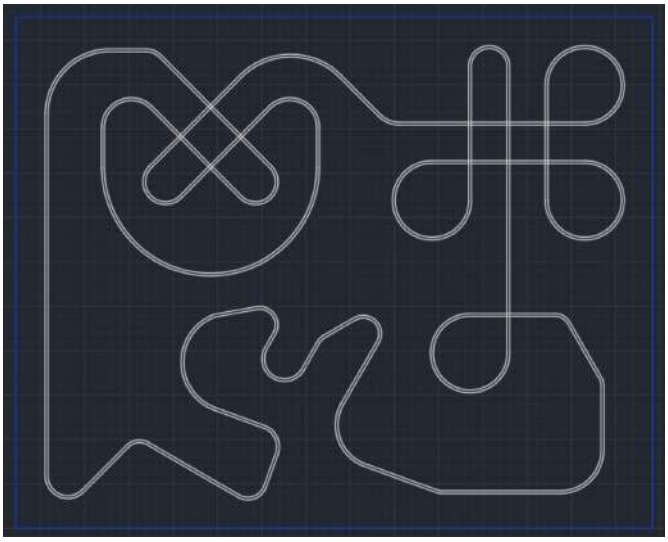


Figura 1. Pista do Follow Line

escuras e claras da pista. A manta preta absorve a luz IR enquanto a linha branca a reflete, permitindo ao robô calcular o erro posicional e ajustar sua trajetória. Este princípio é amplamente utilizado em robôs competitivos da categoria Seguidor de Linha.

C. Motores de Corrente Contínua e Encoders

Motores de corrente contínua com caixa de redução são controlados por PWM, enquanto encoders incrementais permitem medir velocidade e deslocamento por meio da contagem de pulsos. A relação temporal entre os pulsos permite identificar RPM e velocidade linear, possibilitando modelagem dinâmica para desenvolvimento de controladores de velocidade.

D. Controladores PID e PD

Controladores PID ou PD corrigem o erro lateral do robô em relação à linha. A ação proporcional reduz o erro, enquanto a derivativa minimiza oscilações. Em robôs de alta velocidade, a ação integral é frequentemente reduzida devido ao risco de instabilidade. O controlador PD é o mais eficiente na navegação nos robôs da equipe RobotBulls.

E. Força Normal e Downforce

A aderência depende da força de atrito entre pneus e pista. A adição de massa aumenta a força normal, porém prejudica aceleração. O uso de ventoinhas para criação de downforce aumenta a força normal sem comprometer a dinâmica e a massa, permitindo curvas mais rápidas e maior estabilidade, sendo uma solução adotada nos robôs de competição.

III. Materiais e Metodologia

O desenvolvimento de um sistema robótico para a categoria "Seguidor de Linha Pro"exige uma abordagem de engenharia mecatrônica que integre, de forma coesa, o design mecânico, a eletrônica embarcada de alto desempenho e algoritmos de controle em tempo real. A metodologia adotada neste projeto priorizou a otimização da relação peso-potência e a eficiência aerodinâmica. Abaixo, detalham-se as etapas de manufatura, a arquitetura de hardware e as estratégias de software implementadas.

A. Desenvolvimento Mecânico do Robô

O ciclo de vida do projeto iniciou-se pela definição da estrutura física (form factor). Diferentemente de robôs convencionais, onde a eletrônica é adaptada a um chassi pré-existente, neste projeto a estrutura mecânica foi desenvolvida para determinar os limites físicos do sistema. A geometria foi modelada visando garantir um centro de gravidade baixo e centralizado, essencial para a estabilidade vetorial em curvas de alta velocidade. A concepção mecânica definiu as áreas de fixação dos motores, o posicionamento da turbina de sucção e os pontos de apoio no solo, garantindo que o robô respeitasse as dimensões máximas permitidas pelo regulamento da categoria.

B. Projeto Eletrônico e Fabricação da PCB

Após a validação da estrutura física, procedeu-se ao projeto do esquemático eletrônico e ao layout da Placa de Circuito Impresso (PCB). A PCB foi desenhada para atuar simultaneamente como circuito lógico e chassi estrutural do robô. O contorno da placa e os furos de fixação foram definidos estritamente de acordo com o formato mecânico desenvolvido na etapa anterior. A Figura 2 apresenta a placa de PCB do FlashBull 2.1.

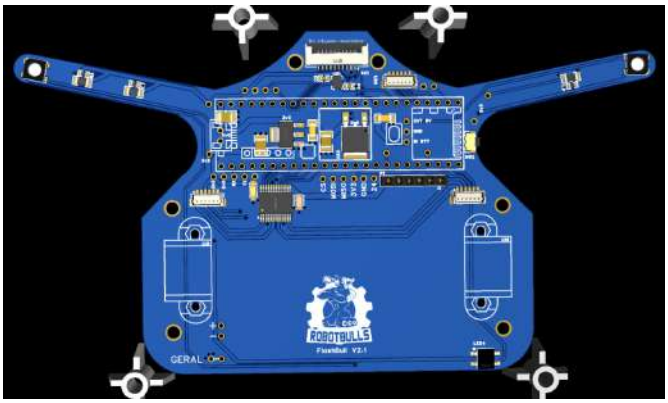


Figura 2. Placa PCB da versão 2.1

C. Sistema de Sensoriamento e QTR

A percepção do ambiente é realizada por um arranjo de sensores baseados na tecnologia de refletância óptica. O projeto integra sensores QTR (Quantum Tunneling Composite ou fototransistores similares) diretamente na PCB principal.

- Sensores Frontais (7 unidades): Dispostos em arco ou linha na extremidade dianteira, responsáveis por calcular o erro posicional do robô em relação à linha branca. A Figura 3 apresenta a disposição dos sensores.

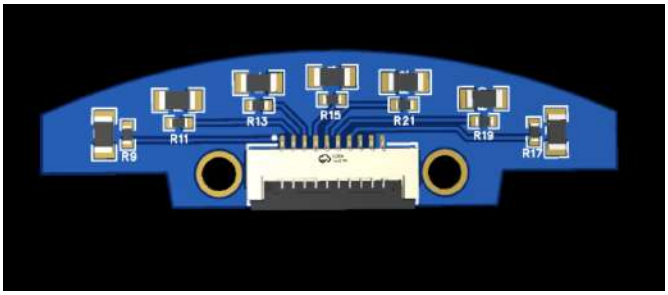


Figura 3. Sensores frontais

- Sensores Laterais (3 unidades): Posicionados estrategicamente para identificar marcadores de pista e auxiliar na frenagem preditiva. A Figura 4 apresenta o modelo do sensor usado na lateral.



Figura 4. Sensor lateral

- Sensor Infravermelho (1 unidade): Utilizado para calibração de distância ou detecção de obstáculos.

O posicionamento destes componentes foi definido na etapa de design da PCB para minimizar a inércia rotacional e garantir leituras livres de interferência mecânica.

D. Sistema de Locomoção (Motores, Encoders, Rodas)

A propulsão utiliza uma configuração de tração diferencial. Dois motores de corrente contínua (CC) de alta rotação são acoplados a suportes projetados para dissipação térmica e rigidez mecânica. O contato com a pista é feito por quatro rodas revestidas com pneus de silicone de alta aderência, minimizando o deslizamento. A Figura 5 mostra o modelo usado para as rodas e motores.

Para garantir o controle em malha fechada, cada motor é monitorado por um encoder de quadratura. Este dispositivo



Figura 5. Motor N20 e roda

fornece feedback da velocidade angular e deslocamento das rodas, permitindo que o microcontrolador compense diferenças físicas entre os motores e garanta uma trajetória retilínea precisa. A estabilidade do conjunto é complementada por uma esfera frontal (roda omnidirecional), que reduz o atrito de arrasto durante as mudanças de direção.

E. Sistema de Sucção e Ventoinhas

Uma característica distintiva da categoria "Pro"é o uso de downforce ativo. O robô é equipado com um sistema de ventoinhas, acionado por motores sem escova e um controlador ESC. Este sistema cria uma zona de baixa pressão sob o chassi, aumentando a força normal sobre os pneus e, consequentemente, a força de atrito disponível, sem a adição significativa de massa. Adicionalmente, esse sistema gerante o fluxo de ar para o arrefecimento dos drivers de potência e baterias. O sistema das ventoinhas é apresentado na Figura 6.

F. Controle Eletrônico e Driver de Potência

O acionamento dos motores é intermediado por uma Ponte H (Driver de Motores). Este componente é crítico para converter os sinais lógicos de baixa potência do microcontrolador em correntes elevadas necessárias para o torque dos motores. A Ponte H permite o controle bidirecional e a modulação de velocidade através de sinais PWM (Pulse Width Modulation), garantindo respostas rápidas às demandas do algoritmo de controle. A Figura 7 apresenta o componente.

G. Desenvolvimento do Software (Arduino/C++)

O firmware do robô foi desenvolvido em linguagem C++, utilizando a IDE e as bibliotecas do ecossistema Arduino [8] pela sua robustez e portabilidade. O código é estruturado para realizar a leitura dos sensores, o processamento do sinal e a atualização dos atuadores em ciclos de tempo determinísticos (tempo real).



Figura 6. Sistema de ventoinhas com motores *brushless*

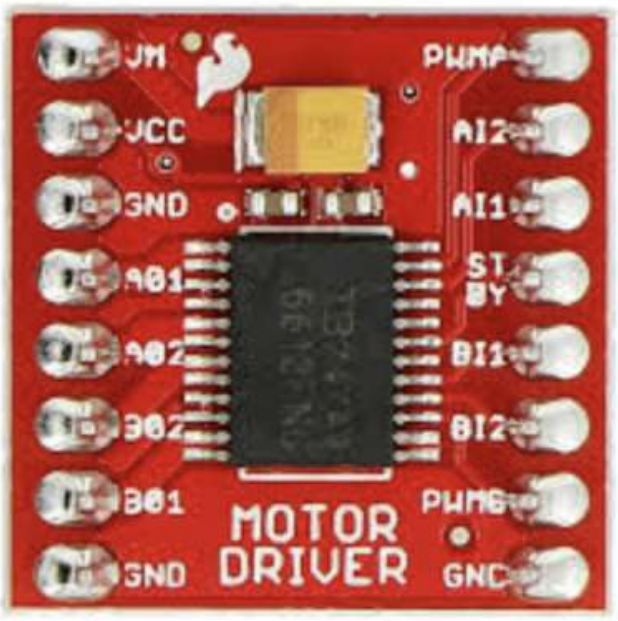


Figura 7. Driver Ponte H

Atualmente, o robô descrito atua como plataforma de validação (testbed). Os dados coletados nestes ensaios estão fundamentando o desenvolvimento de uma nova iteração do robô, projetada para estrear na competição IronCup. Um trecho do código é apresentado na Figura 8.

H. Estratégia de Controle PID

O núcleo da navegação autônoma baseia-se no algoritmo de controle PID (Proporcional, Integral e Derivativo) [7].

- **Proporcional (P):** Corrige o erro atual (distância do centro da linha).



Figura 8. Trecho de código em C++

- **Integral (I):** Compensa erros acumulados ao longo do tempo (pouco utilizado em seguidores de alta velocidade devido ao risco de instabilidade).
- **Derivativo (D):** Prediz o comportamento futuro do erro com base em sua taxa de variação, reduzindo oscilações.

A sintonia dos ganhos foi realizada utilizando uma abordagem baseada no método de Ziegler–Nichols em malha fechada [3], seguida de refinamento empírico em pista. Inicialmente, os ganhos integral e derivativo foram anulados ($K_i = 0$, $K_d = 0$), e o ganho proporcional foi aumentado gradualmente até que o sistema apresentasse oscilações sustentadas, definindo o ganho crítico K_u e o período de oscilação T_u.

A partir desses parâmetros, foram estimados os ganhos iniciais do controlador. Entretanto, devido às características não lineares do sistema e à alta velocidade do robô, optou-se por utilizar um controlador do tipo PD, descartando o termo integral. Assim, os ganhos foram ajustados experimentalmente da seguinte forma:

- Ajuste de K_p para obter resposta rápida sem saturação excessiva;
- Ajuste de K_d para reduzir overshoot e melhorar a estabilidade em curvas;
- Manutenção de $K_i \approx 0$ para evitar acúmulo de erro e instabilidade.

Esse processo iterativo permitiu encontrar um compromisso entre tempo de resposta e estabilidade, sendo validado por testes práticos em pista.

O sistema é exemplificado na Figura 9.

I. Fonte de Energia e Distribuição Elétrica

A alimentação é provida por duas baterias (provavelmente LiPo) de alta taxa de descarga, conectadas em série ou paralelo dependendo da tensão nominal dos motores. A segurança operacional é garantida por uma chave de ligamento mestre. Para a eletrônica lógica, um regulador de tensão converte a energia da bateria para níveis estáveis (geralmente 5V

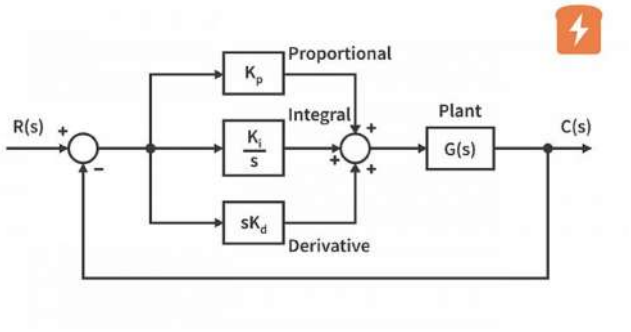


Figura 9. Algoritmo de controle PID

ou 3.3V), protegendo o microcontrolador e sensores contra flutuações geradas pelos motores. A Figura 10 apresenta um modelo de bateria utilizado.



Figura 10. Bateria LiPo

J. Materiais Utilizados

O desenvolvimento do sistema exigiu uma abordagem mecatrônica focada na otimização da relação peso-potência e na eficiência aerodinâmica. A estrutura mecânica foi modelada para garantir um centro de gravidade baixo, e a própria Placa de Circuito Impresso (PCB) foi projetada para atuar simultaneamente como circuito lógico e chassi estrutural.

A percepção do ambiente utiliza um arranjo de sensores frontais e laterais para o cálculo de erro e leitura de marcadores. O processamento ocorre em tempo real via firmware em C++, executando a estratégia de controle PD baseada no erro medido. Os principais componentes utilizados no robô e suas respectivas funções estão detalhados na Tabela I. [9]

IV. Testes e Resultados

A. Aumento de Força Normal por Ventoinhas

Com 70% da potência aplicada às ventoinhas, o robô FlashBull apresentou aumento de aproximadamente 168 g na força normal, valor praticamente equivalente ao seu peso real (162 g). Esse aumento de downforce proporcionou maior aderência e reduziu o tempo de volta em cerca de 16%, demonstrando que o uso da ventoinha é decisivo para curvas de alta velocidade. [7]

Tabela I
Lista de Materiais e Componentes do Robô

Componente	Qtd.	Função	Justificativa / Impacto
Placa de PCB	1	Estrutura e circuito lógico	Reduz peso e elimina fiação interna.
Microcontrolador	1	Processamento central	Alta frequência de *clock* para PID.
Ponte H	1	Controle de potência	Corrente de pico e resposta rápida ao PWM.
Regulador de Tensão	1	Estabilização	3.3V/5V estáveis para evitar *reset*.
Ventoinhas	4	Arrefecimento ativo	Evita superaquecimento em alta carga.
Sistema de Sucção	1	*Downforce*	Aumenta aderência nas curvas.
Motor da Sucção	1	Acionamento da turbina	Alta rotação para gerar vácuo.
Suporte de Sucção	1	Fixação	Mantém vedação e proximidade ao solo.
Motores	2	Tração	Alto torque e RPM.
Encoders	2	Velocidade	Permite controle fechado e odometria.
Rodas / Pneus	4	Contato com o solo	Silicone para alta tração.
Esfera frontal	1	Apoio passivo	Reduz atrito frontal.
Suportes de Motor	2	Fixação	Alinhamento e absorção de vibração.
Sensores frontais	7	Erro da linha	Alta resolução para PID suave.
Sensores laterais	3	Marcadores	Identificação de início/fim e cruzamentos.
Cabo Flat	1	Conexão	Organização e integridade do sinal.
Sensor Infravermelho	1	Auxiliar	Detecção/calibração.
Baterias	2	Energia	Alta densidade e C–rate.
Chave de Ligamento	1	Segurança	Desligamento imediato em falhas.

B. Teste do Controle de Velocidade

Os testes do controle de velocidade foram realizados em dois cenários distintos: [7]

- Ventoinha Desligada:
  - Tempo de acomodação: 0.65 s
  - Overshoot: 0%
  - Resposta mais rápida, porém com menor aderência
- Ventoinha Ligada:
  - Tempo de acomodação: 1.25 s
  - Overshoot: 18%
  - Maior estabilidade em curvas devido ao aumento de downforce

C. Mapeamento Automático da Pista

O robô executa uma volta inicial para identificar marcações laterais, classificando os trechos em curva aberta, curva média, curva fechada, reta curta, média e longa. Para cada tipo de trecho, parâmetros como Kp, Kd, velocidade máxima e

potência das ventoinhas são ajustados automaticamente. Isso reduz oscilações e melhora a estabilidade geral. [7]

D. Comparação das Configurações do Robô

Tabela II
COMPARAÇÃO DE CONFIGURAÇÕES DO FLASHBULL

Configuração	Ventoinha	Mapeamento	Tempo (s)
SpeedBull	Não	Sim	12.45
FlashBull	Não	Não	11.83
FlashBull	Não	Sim	10.39
FlashBull	Sim	Não	9.94
FlashBull (final)	**Sim**	**Sim**	**8.76**

A combinação de ventoinha + mapeamento resultou no menor tempo já registrado pela equipe RobotBulls. [2] [7].

V. CONCLUSÃO

A. Principais Resultados

O desenvolvimento do robô FlashBull evidenciou avanços significativos tanto na arquitetura mecânica quanto nos sistemas de controle, sensoriamento e otimização de desempenho. O conjunto de testes demonstrou que a integração entre sensores QTR [3] de alta densidade, controladores PD ajustados para a dinâmica do robô e o uso estratégico de ventoinhas permitiu alcançar tempos de volta substancialmente menores quando comparados às versões anteriores do robô.

Além disso, o emprego de ventoinhas para geração de *downforce* apresentou impacto direto na estabilidade lateral e na capacidade de realizar curvas em alta velocidade. Mesmo com o aumento do tempo de acomodação no controle de velocidade, o ganho de aderência compensou esse efeito, resultando em um comportamento mais previsível e seguro ao longo da pista.

B. Trabalhos Relacionados

Diversos estudos abordam robôs seguidores de linha voltados para competição, destacando o uso de controle PD e otimizações mecânicas [7]. Estratégias como mapeamento de pista e ajuste dinâmico de parâmetros têm sido exploradas para redução do tempo de percurso e aumento da estabilidade.

C. Limitações do Sistema

Apesar dos resultados positivos, algumas limitações foram identificadas. Observou-se que o acionamento das ventoinhas altera a dinâmica do robô de forma não linear, especialmente em trechos de aceleração rápida, exigindo ajustes constantes nos parâmetros dos controladores. Além disso, a ausência de sensores inerciais impede que o sistema identifique variações abruptas de inclinação ou trepidação, o que poderia contribuir para um mapeamento mais robusto da pista.

Outro ponto crítico está relacionado ao módulo de mapeamento baseado apenas em sensores ópticos. Embora eficiente, o sistema é sensível a condições de iluminação e sujeira, o que pode comprometer a classificação de curvas e retas em ambientes fora do padrão competitivo.

D. Trabalhos Futuros

Para melhorar o desempenho do FlashBull e evoluir o projeto, recomenda-se:

- Adoção de ventoinhas de sucção, capazes de gerar *downforce* mais estável e independente da velocidade do ar;
- Uso de sensores inerciais (IMU) [1] para detecção de vibração, aceleração e inclinação, aprimorando o mapeamento da pista;
- Encoders de maior resolução, permitindo estimar velocidade com menor ruído e melhorando a malha fechada;
- Implementação de controle adaptativo ou baseado em ganhos variáveis, ajustando dinamicamente Kp e Kd conforme o trecho da pista;
- Desenvolvimento de algoritmos preditivos, capazes de antecipar curvas com base na leitura histórica dos sensores;
- Análise térmica dos componentes, otimizando o posicionamento das ventoinhas e assegurando confiabilidade em longos períodos de operação.

Por fim, conclui-se que o projeto alcançou seus objetivos principais ao apresentar um robô estável, rápido e tecnicamente alinhado às exigências da categoria Seguidor de Linha Pro. Os resultados obtidos demonstram o potencial de evolução contínua da plataforma FlashBull, consolidando-se como base para novas iterações e pesquisas futuras dentro da equipe RobotBulls. [4] [2]

REFERÊNCIAS

[1] BAZANELLA, Alexandre Sanfelice. *Sistemas de controle: princípios e métodos de projeto* Porto Alegre: Ed. UFRGS, 2005. ISBN 8570258496.
[2] INATEL. *Robotbulls: Equipe de Robótica do Inatel*. Santa Rita do Sapucaí: Instituto Nacional de Telecomunicações, 2023.
[3] OGATA, Katsuhiko. *Engenharia de controle moderno*. 5. ed. São Paulo: Pearson Prentice Hall, 2010.
[4] POLOLU. *QTR Reflectance Sensors User's Guide*. Las Vegas: Pololu Robotics and Electronics, 2023.
[5] ROBOCORE. *Ranking Nacional de Robôs: Categoria Seguidor de Linha Pro*. 2024.
[6] ROBOCORE. *Regras da Categoria Seguidor de Linha*. 2023.
[7] SOUZA, LEANDRO MUCCIARONI DE. *Robôs Seguidores de Linha: Desenvolvimento e Implementações para Competição*. 2025.
[8] ARDUINO. *Arduino Reference Documentation*. 2025.
[9] TOSHIBA. *TB6612FNG Driver IC Datasheet*. Toshiba Electronic Devices & Storage Corporation, 2014.

AUTORES



Caroline Ferreira de Souza é graduanda em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel). Atua como membro da equipe de robótica RobotBulls, onde desenvolve pesquisas voltadas para a implementação de software e algoritmos de controle em robôs de competição. Possui experiência em projetos acadêmicos e atividades de extensão que integram o desenvolvimento de sistemas e soluções tecnológicas, com ênfase em lógica de programação e robótica móvel.



Pedro Paulo de Paiva Junqueira é graduando em Engenharia da Computação pelo Instituto Nacional de Telecomunicações (Inatel). Atua como bolsista da equipe de robótica RobotBull, onde realiza pesquisas para robôs de alta perfomance e monitor da disciplinas de Probabilidade e Processos Estocásticos e Física I, II e III. É entusiasta da tecnologia, automação e inovação, direcionando suas pesquisas nessas áreas.

Estudo para novos Sistemas de Navegação e Controle para Robôs Autônomos Trekking

Fuad Murad e Mayara do Prado Almeida
Alexandre Baratella Lugli, Egídio Raimundo Neto, Eugênio Moreira Jonas da Silva
Luiz Fernando Domiciano Ribeiro dos Santos e Wanderson Eleutério Saldanha

***Abstract*— The RobotBulls team was founded 13 years ago at the National Institute of Telecommunications, an institution recognized for its strong support and encouragement of robotics competitions.**

Within the RobotBulls team, the Trekking category focuses on the study and development of new navigation and control systems for autonomous robots. The goal is to create solutions that enable the robots to operate independently, efficiently, and safely in complex courses, exploring new technological approaches and advanced automation strategies.

***Index Terms*— Robotics, Autonomous Robots, Navigation Systems, Robot Control, Trekking.**

***Resumo*— A equipe RobotBulls foi criada há 13 anos no Instituto Nacional de Telecomunicações, instituição reconhecida pelo forte apoio e incentivo às competições de robótica. Na equipe RobotBulls, a categoria Trekking dedica-se ao estudo e desenvolvimento de novos sistemas de navegação e controle para robôs autônomos. O objetivo é criar soluções que permitam aos robôs operar de maneira independente, eficiente e segura em percursos complexos, explorando novas abordagens tecnológicas e estratégias de automação avançadas.**

***Palavras Chave*— Robótica, Robôs Autônomos, Sistemas de Navegação, Controle de Robôs, Trekking.**

I. Introdução

A. Equipe de Robótica - RobotBulls

A equipe de robótica Robotbulls, vinculada ao Instituto Nacional de Telecomunicações (Inatel), localizado em Santa Rita do Sapucaí, Minas Gerais, é uma equipe competitiva formada por estudantes de engenharia que atuam no desenvolvimento de robôs para diferentes modalidades de competições. A equipe conta com o apoio e orientação de professores, que auxiliam no processo de aprendizagem técnica e no desenvolvimento dos projetos, contribuindo para a formação prática dos alunos na área de robótica.

Fundada em 2013, a Robotbulls consolidou-se ao longo dos anos como uma equipe ativa no cenário nacional de robótica competitiva, participando de diversas competições e desenvolvendo soluções em diferentes áreas da engenharia. Atualmente, a equipe é dividida em várias categorias, abrangendo diferentes modalidades de competição, como drones, robôs de sumô e de combate, entre outras áreas que exploram diferentes desafios técnicos e estratégias de controle.

Dentro desse conjunto de categorias, destaca-se a modalidade Trekking, voltada ao desenvolvimento de robôs autônomos para navegação em ambientes estruturados e não estruturados. É nessa categoria que se insere o objeto de estudo deste trabalho, no qual será abordado o desenvolvimento de um robô Trekking, incluindo seus principais sistemas e soluções adotadas para atender aos requisitos da competição.

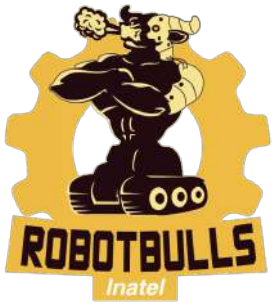


Fig. 1. Robotbulls

II. O que é o Trekking

A categoria *Trekking* é uma modalidade de competições de robótica voltada à navegação autônoma em ambientes desafiadores. Nela, o robô deve percorrer um trajeto previamente definido, passando por pontos de controle e lidando com diferentes tipos de obstáculos, sem qualquer intervenção humana. Esse tipo de desafio simula situações reais em que sistemas robóticos precisam tomar decisões e se deslocar de forma independente.

Nesse contexto, a robótica autônoma vem ganhando destaque na engenharia devido à necessidade de sistemas capazes de operar em ambientes complexos sem intervenção humana. O desenvolvimento de veículos terrestres autônomos, como os robôs Trekking, envolve a integração de tecnologias como inteligência artificial, visão computacional e diferentes tipos de sensores.

As competições acadêmicas de robótica têm um papel essencial nesse processo, pois permitem que estudantes apliquem, na prática, conhecimentos adquiridos em sala de aula. No Brasil, existem eventos consolidados que impulsionam esse desenvolvimento, como a IronCup, realizada no Instituto Nacional de Telecomunicações (Inatel), em Santa Rita do Sapucaí (MG), a RSM Challenge, em Mogi das Cruzes (SP), e a RoboCore Experience (RCX), considerada uma das maiores competições da América Latina.

Essas competições reúnem participantes de diversas instituições e promovem não apenas a competição, mas também a troca de conhecimento e o desenvolvimento técnico. Além disso, contam com diferentes categorias, permitindo que as equipes explorem várias áreas da robótica.

Dentro desse cenário, a categoria *Trekking* se destaca por exigir que as equipes desenvolvam um robô capaz de se locomover e tomar decisões sozinho, sem qualquer controle humano durante a prova.

A dificuldade da prova não está apenas em seguir o caminho correto. O robô também precisa lidar com obstáculos, escolher rotas e administrar bem o tempo. Em alguns casos, existem caminhos alternativos ou desafios extras que oferecem pontuação adicional, o que exige que o sistema tome decisões durante a execução. Por isso, a competição avalia não só a velocidade, mas também o funcionamento do software e a confiabilidade do sistema em diferentes situações.

A participação em competições como essas reforçam a importância do desenvolvimento de sistemas autônomos mais robustos. Desse modo, o estudo e a implementação de soluções para a categoria *Trekking* contribuem diretamente para o avanço da robótica móvel, ao lidar com problemas reais de navegação, percepção e tomada de decisão em ambientes não controlados.

A. *Como funciona a categoria*

Na categoria Trekking, o robô deve operar de forma totalmente autônoma durante toda a execução da prova, sem qualquer tipo de controle remoto ou intervenção humana após a largada. Antes do início das tomadas de tempo, todos os robôs passam por uma inspeção realizada pelos juízes da competição, na qual são verificadas as dimensões do equipamento e a presença e funcionamento dos sistemas de segurança, incluindo o mecanismo de parada de emergência (*failsafe*), os robôs também devem apresentar um dispositivo de indicação luminosa, para sinalizar a chegada aos marcos [1]. Apenas os robôs que estiverem dentro das normas estabelecidas são autorizados a participar.

As dimensões do robô são geralmente limitadas a 500 x 500 x 500 mm [1], garantindo padronização entre os competidores. Além disso, o sistema de failsafe é um requisito obrigatório, devendo permitir a interrupção imediata do robô em caso de falhas ou situações de risco, assegurando a segurança do ambiente, dos organizadores e dos demais participantes.

A prova é composta por um trajeto definido com quatro marcos 2 de chegada, sinalizados por placas metálicas e cones 3 distribuídos ao longo do percurso. O objetivo do robô é percorrer esses marcos de forma sequencial, enfrentando diferentes condições de ambiente, como terrenos irregulares, rampas, obstáculos físicos e variações de iluminação, o que exige capacidade de adaptação em tempo real.

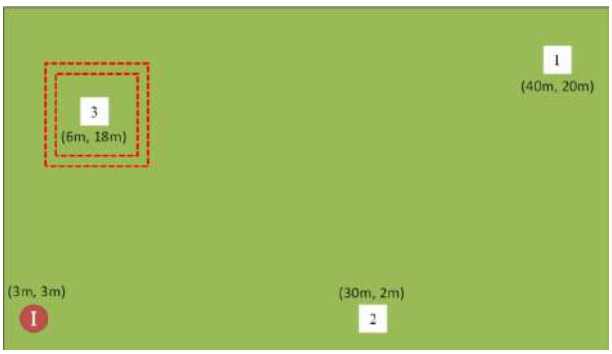


Fig. 2. Posicionamento dos marcos.



Fig. 3. Cones de Sinialização.

A competição é organizada em 2 ou 3 tomadas de tempo, nas quais cada equipe possui 30 minutos para realizar suas tentativas de percurso [1]. Dentro desse período, a equipe pode iniciar quantas execuções forem permitidas pela organização, buscando completar o trajeto da melhor forma possível. O desempenho final é baseado no menor tempo obtido para completar os quatro marcos obrigatórios.

Além disso, podem existir caminhos alternativos associados a multiplicadores de tempo, que alteram o tempo final de execução dependendo das decisões tomadas durante o percurso. Dessa forma, a estratégia adotada pelo sistema também influencia diretamente o resultado, já que o robô precisa equilibrar eficiência de trajeto e escolha de rotas.

Ao final da competição, vence a equipe que completar todos os marcos obrigatórios no menor tempo total, considerando ou não a aplicação dos multiplicadores de tempo.

III. Robô Bullratella

A. *Conheça o Bullratella*

O Bullratella é um robô de *trekking* autônomo desenvolvido pela equipe de robótica RobotBulls, do Instituto Nacional de Telecomunicações (Inatel). Criado para competições de navegação autônoma, o robô foi projetado para percorrer trajetos desafiadores sem intervenção humana, tomando decisões em tempo real a partir da interpretação do ambiente ao seu redor. Sua estrutura reúne mecânica, eletrônica embarcada e sistemas inteligentes, formando uma plataforma robusta, eficiente e preparada para desafios de alto nível técnico.

O sistema eletrônico do Bullratella foi desenvolvido com foco em confiabilidade e desempenho. Sua arquitetura utiliza componentes responsáveis pelo gerenciamento de energia, processamento e sensoriamento, garantindo o funcionamento estável de todos os módulos embarcados. Para isso, são empregados conversores Step-down e Step-up [2], que ajustam os níveis de tensão necessários para cada dispositivo do robô. Esses conversores permitem alimentar corretamente componentes sensíveis, como o microcontrolador, sensores e módulos de processamento, além dos atuadores responsáveis pela movimentação.

A unidade principal de controle utiliza o Arduino Nano [3], responsável por executar a lógica de navegação, leitura dos sensores e acionamento dos sistemas mecânicos. O deslocamento do robô é realizado por meio de um motor de tração, que fornece a força necessária para o movimento durante o percurso. Já o direcionamento é feito por um servomotor acoplado ao eixo

das rodas dianteiras, permitindo o controle preciso das curvas e mudanças de trajetória ao longo da prova.

Complementando esse sistema, a Unidade de Medição Inercial MPU-6050 [4] exerce papel essencial na estabilidade do robô, monitorando continuamente dados de inclinação e rotação angular. Com essas informações, o Bullratella consegue corrigir desvios de trajetória e manter maior precisão durante o deslocamento, especialmente em terrenos irregulares, obstáculos e curvas mais fechadas.

Outro destaque importante é o sistema de visão computacional. Por meio de uma câmera integrada à plataforma NVIDIA Jetson Nano [5], o robô é capaz de identificar cones, marcos e outros elementos presentes no percurso das competições. A Jetson realiza o processamento das imagens em tempo real, permitindo reconhecer obstáculos e interpretar o trajeto com rapidez e eficiência. Esse recurso amplia significativamente a autonomia do sistema e torna a navegação mais segura e estratégica.

A integração entre todos esses componentes transforma o Bullratella em uma plataforma tecnológica completa, capaz de perceber o ambiente, processar informações e reagir de forma inteligente às situações encontradas durante a prova.

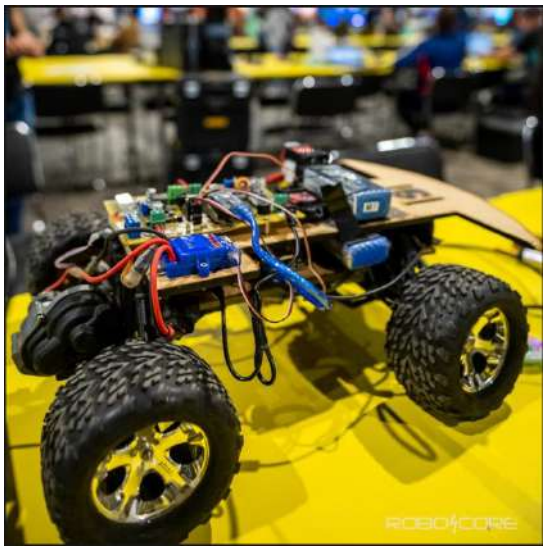


Fig. 4. Robô Bullratella

B. *Controle de Navegação*

O controle de navegação do Bullratella foi desenvolvido para converter as decisões do sistema em movimentos rápidos, estáveis e precisos ao longo do percurso. A partir das informações obtidas pelo sensoriamento e pela visão computacional, o robô interpreta o ambiente e ajusta continuamente sua trajetória, mantendo deslocamento seguro mesmo diante de curvas fechadas, obstáculos e mudanças inesperadas no terreno.

A propulsão é realizada por um motor brushed acionado por uma ESC (Electronic Speed Controller), responsável por regular a potência entregue ao sistema de tração. Esse controle eletrônico permite respostas suaves na aceleração e desaceleração, além de maior eficiência no gerenciamento da velocidade durante a prova. Dessa forma, o robô consegue adaptar seu ritmo conforme a necessidade de cada trecho do percurso.

O direcionamento é executado por um servomotor acoplado ao sistema de esterçamento das rodas dianteiras. Por meio dele, o Bullratella realiza correções rápidas de rota, contorna obstáculos e mantém alinhamento preciso nos corredores de passagem. A atuação conjunta entre velocidade controlada e direção precisa garante maior fluidez nos movimentos.

Para assegurar estabilidade dinâmica, a navegação utiliza dados da unidade inercial, que permitem identificar desvios angulares e oscilações durante o deslocamento. Com essas informações, o sistema aplica correções instantâneas, reduzindo erros acumulados e melhorando a precisão da trajetória.

Todo esse processo ocorre em tempo real, em uma malha contínua de leitura, decisão e resposta. Assim, o Bullratella mantém comportamento autônomo, consistente e competitivo, sendo capaz de reagir com eficiência às exigências do percurso.

C. *Sistema Eletrônico*

O desenvolvimento do sistema eletrônico do Bullratella foi precedido por um estudo técnico voltado à definição de uma arquitetura confiável, eficiente e compatível com as exigências das competições de *trekking* autônomo. Nessa etapa inicial, foram analisadas as demandas energéticas de cada subsistema embarcado, bem como os impactos causados por variações de tensão, interferências elétricas e picos de corrente durante a operação dinâmica do robô.

A partir dessa avaliação, foi estruturada uma rede de alimentação capaz de distribuir energia de forma adequada entre os módulos do projeto, preservando estabilidade mesmo em situações de carga variável. Também foram considerados critérios de proteção elétrica e facilidade de manutenção, buscando reduzir riscos de falha durante testes e competições. Essa abordagem permitiu construir uma plataforma mais robusta e preparada para longos ciclos de funcionamento.

Outro ponto relevante do estudo envolveu a organização física dos componentes eletrônicos no chassi. O posicionamento estratégico dos módulos, cabeamentos e conexões contribuiu para minimizar ruídos, reduzir perdas elétricas e facilitar intervenções técnicas. Em sistemas móveis, essa disposição interna influencia diretamente a confiabilidade e o desempenho geral do conjunto.

Dessa forma, o sistema eletrônico não foi apenas montado para atender necessidades imediatas, mas projetado com base em critérios de engenharia que garantissem estabilidade operacional e possibilidade de futuras expansões.

D. *Análise e Controle em Tempo Real*

A implementação da lógica de controle em tempo real também foi resultado de um estudo prévio sobre as exigências impostas por percursos dinâmicos e ambientes imprevisíveis. Em competições de trekking, o robô precisa interpretar mudanças rapidamente e responder sem atrasos significativos. Por isso, foi necessária a definição de uma arquitetura capaz de processar informações continuamente e gerar comandos com baixa latência.

Durante essa etapa, foram avaliadas estratégias de integração entre sensoriamento, processamento e atuação, buscando

equilíbrio entre velocidade de resposta e confiabilidade das decisões tomadas. O objetivo central era permitir que o robô corrigisse sua trajetória de forma natural, evitando instabilidades e reduzindo erros acumulados ao longo do percurso.

Também foram estudadas técnicas de realimentação contínua, nas quais o sistema compara constantemente o comportamento esperado com o comportamento real do veículo. Esse método permite ajustes sucessivos durante a navegação, melhorando precisão em curvas, retomadas de alinhamento e mudanças bruscas de cenário.

Como resultado, o Bullratella passou a operar de maneira dinâmica e responsiva, demonstrando que a aplicação de métodos de controle bem estruturados é fundamental para o desempenho competitivo de plataformas autônomas.

E. Conquistas do Robô Bullratella e Motivação para Novos Desenvolvimentos

Os resultados obtidos pelo Bullratella validaram, na prática, os estudos e decisões de engenharia adotados ao longo do projeto. O desempenho alcançado em competições demonstrou que a metodologia empregada no desenvolvimento foi eficaz tanto do ponto de vista técnico quanto estratégico, confirmando a qualidade das soluções implementadas.

As conquistas obtidas ao longo da temporada competitiva mostraram que o robô possuía capacidade de operar com repetibilidade, estabilidade e adaptação a diferentes condições de percurso. Esse retorno experimental foi essencial, pois permitiu comparar teoria e prática, identificando quais escolhas apresentaram melhor desempenho em situações reais.

Ao mesmo tempo, a experiência acumulada revelou oportunidades de evolução em áreas como capacidade computacional, precisão sensorial e otimização estrutural. Em vez de representar o encerramento de um ciclo, o sucesso do Bullratella tornou-se o ponto de partida para novas pesquisas e propostas de melhoria.

Nesse contexto, surgiu a motivação para o desenvolvimento de uma nova plataforma, voltada à incorporação de tecnologias mais avançadas e estratégias de navegação ainda mais eficientes. Assim, o Bullratella consolidou-se não apenas como um robô competitivo, mas como base experimental para a continuidade da inovação dentro da equipe.



Fig. 5. Validação dos estudos anteriores

IV. Novo Sistema em Desenvolvimento: Robô Bullgidio

A. Robô Bullgidio

Dando continuidade ao processo de evolução tecnológica da equipe, o Bullgidio surge como a nova plataforma de pesquisa e competição desenvolvida para superar limitações identificadas nas gerações anteriores. Seu projeto foi concebido a partir da experiência acumulada com o Bullratella, utilizando dados reais de desempenho, testes de campo e análises estruturais para definir melhorias em agilidade, processamento e capacidade de navegação autônoma.

Uma das principais mudanças está no redesenho mecânico do chassi, agora mais compacto, leve e otimizado. A redução de massa e a reorganização estrutural contribuem para diminuir a inércia do conjunto, permitindo respostas mais rápidas em curvas, retomadas de alinhamento e manobras evasivas. Além disso, a nova distribuição de peso favorece maior estabilidade dinâmica, especialmente em terrenos irregulares e mudanças bruscas de direção.

Outro avanço importante foi a reorganização interna dos componentes eletrônicos. O novo layout reduz comprimentos de cabeamento, melhora a dissipação térmica e diminui a suscetibilidade a interferências eletromagnéticas e falhas causadas por vibração mecânica. Dessa forma, o Bullgidio não representa apenas uma atualização estética, mas uma plataforma inteiramente reestruturada para maior eficiência e confiabilidade operacional.

B. Novas Integrações de Navegação e Processamento

Entre as melhorias em estudo para o Bullgidio, destaca-se a modernização conjunta do sistema de controle e da percepção espacial do robô. A proposta de substituição do Arduino Nano pelo microcontrolador ESP32[6] representa um avanço importante em capacidade de processamento, velocidade de resposta e flexibilidade operacional. Diferentemente da arquitetura anterior, a utilização da ESP32[6] também amplia os recursos de comunicação sem fio, permitindo o envio de códigos, ajustes de parâmetros, atualização de rotas e testes operacionais por Bluetooth ou Wi-Fi. Com isso, reduz-se a dependência de conexões físicas por cabos, tornando o processo de configuração mais ágil, prático e eficiente durante o desenvolvimento.

Paralelamente, está em estudo a integração de um sensor LiDAR como solução avançada para percepção do ambiente. Essa tecnologia permite realizar leituras espaciais contínuas ao redor do robô, identificando paredes, obstáculos laterais, corredores estreitos e zonas de passagem com elevada precisão. Em percursos técnicos, esse tipo de informação pode representar vantagem significativa, pois possibilita correções antecipadas de trajetória e deslocamento mais seguro em áreas confinadas.

Cabe destacar que essas implementações ainda fazem parte de uma etapa de estudo, testes e validações experimentais. O objetivo da equipe é analisar o comportamento de cada tecnologia em condições reais de uso, verificando ganhos práticos de desempenho, confiabilidade e robustez antes da consolidação definitiva no projeto. Assim, o Bullgidio nasce como uma plataforma voltada não apenas à competição, mas também à pesquisa e ao desenvolvimento contínuo de novas soluções tecnológicas.

C. Benefícios Gerais do Novo Sistema

Com base na experiência adquirida no desenvolvimento e nas competições com o Bullratella, o projeto do Bullgidio surge como uma nova etapa de evolução técnica construída a partir

de resultados práticos, testes de campo e aprendizados acumulados ao longo dos últimos anos. O Bullratella segue sendo uma plataforma importante para a equipe, tanto pelo desempenho já demonstrado quanto pelo potencial contínuo de aprimoramento. Ao mesmo tempo, as análises realizadas durante sua utilização abriram espaço para novas propostas e soluções que agora passam a ser exploradas no Bullgidio.

Entre os principais benefícios esperados está o aumento da eficiência operacional. A adoção de uma arquitetura mais moderna tende a proporcionar respostas mais rápidas, maior estabilidade dinâmica e melhor precisão em curvas, correções de trajetória e retomadas de alinhamento. Em percursos de trekking, nos quais cada decisão influencia diretamente o tempo final, esse ganho pode representar vantagem competitiva relevante.

Outro ponto de destaque é a ampliação da autonomia do sistema. Com recursos mais avançados de processamento e sensoriamento ainda em fase de estudo e validação, o Bullgidio poderá interpretar o ambiente com maior nível de detalhe, identificando obstáculos, paredes laterais, corredores estreitos e zonas de passagem com mais confiabilidade. Isso tende a favorecer trajetórias mais suaves, menor necessidade de correções bruscas e navegação mais segura em cenários complexos.

Também são esperados benefícios importantes no processo de desenvolvimento da equipe. A comunicação sem fio, a facilidade de atualização de parâmetros e a integração simplificada de novos componentes tornam os testes mais ágeis e eficientes. Além disso, a operação paralela entre Bullratella e Bullgidio cria uma oportunidade valiosa de comparação entre arquiteturas distintas, permitindo validar melhorias com base em resultados concretos.

Por fim, o Bullgidio representa a continuidade do trabalho já desenvolvido pela equipe, incorporando novas tecnologias sem abrir mão da experiência adquirida com o Bullratella. A utilização simultânea das duas plataformas fortalece o processo de testes, comparações e aperfeiçoamentos graduais, contribuindo para decisões técnicas mais consistentes e para a evolução contínua dos projetos futuros.

D. Conclusão

O presente trabalho permitiu analisar o processo de desenvolvimento e evolução de sistemas de navegação e controle aplicados a robôs autônomos da modalidade Trekking, tomando como base as plataformas Bullratella e Bullgidio. A partir do estudo realizado, observou-se que o desempenho competitivo de um robô está diretamente associado à integração eficiente entre estrutura mecânica, eletrônica embarcada, sensoriamento e algoritmos de tomada de decisão em tempo real.

A trajetória do Bullratella demonstrou a importância de uma plataforma consolidada para validação prática de soluções de controle, estabilidade dinâmica e navegação autônoma em ambientes variáveis. Os resultados obtidos com esse sistema forneceram dados relevantes sobre comportamento em pista, limitações operacionais e oportunidades de melhoria, servindo como referência técnica para novas etapas de desenvolvimento.

Com base nesses aprendizados, o Bullgidio surge como proposta de continuidade e modernização, incorporando estudos voltados ao aumento da capacidade computacional, melhoria da comunicação embarcada e ampliação da percepção espacial do robô. Tecnologias como ESP32 e LiDAR representam alternativas promissoras para sistemas que exigem respostas rápidas, maior precisão de posicionamento e adaptação eficiente a trajetos complexos.

Dessa forma, conclui-se que o estudo contínuo de novas arquiteturas de navegação e controle é fundamental para a evolução da robótica móvel autônoma. A análise comparativa entre diferentes plataformas, associada à aplicação prática em competições, permite transformar desafios reais em oportunidades de inovação. Assim, o desenvolvimento progressivo desses sistemas contribui não apenas para melhores resultados competitivos, mas também para o avanço técnico e científico na área de robôs autônomos.

REFERÊNCIAS

[1] Robocore. *Regras Trekking*. URL: https://www.robocore.net/upload/attachments/regras_robotrekking_2018_05_06_172.pdf?srsltid=AfmBOoo7docfE1QTieOyhNqbDn7SZOk778XFDidxdT4ikCVpr7bl (acesso em 28/04/2026).

[2] alldatasheet. *Step-Down Datasheet*. URL: https://www.alldatasheet.com/datasheet-pdf/view/763183/ETC2/XL4015.html (acesso em 28/04/2026).

[3] docs.arduino. *Arduino datasheet*. URL: https://docs.arduino.cc/resources/datasheets/ABX00142-datasheet.pdf (acesso em 28/04/2026).

[4] cdn-reichelt.de. *MPU-6050-datasheet*. URL: https://cdn-reichelt.de/documents/datenblatt/A300/ME024.pdf (acesso em 28/04/2026).

[5] NVIDIA. *Jetson Nano Developer Kit*. URL: https://developer.nvidia.com/embedded/jetson-nano-developer-kit (acesso em 01/10/2025).

[6] Espressif Systems. *ESP32 Series Datasheet*. URL: https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf (acesso em 28/04/2026).

AUTORES



Fuad Murad é estudante de Engenharia da Computação pelo Instituto Nacional de Telecomunicações (INATEL), membro atuante da equipe RobotBulls na categoria Trekking e monitor de Pós-Graduação no INATEl.



Mayara do Prado Almeida é graduanda em Engenharia Elétrica pelo INATEL - Instituto Nacional de Telecomunicações, membro da Robotbulls - Equipe de robótica do INATEL, atuando como Capitã geral e membro da categoria Trekking. Representante da turma de GEL e monitora dos laboratórios temáticos.

Estudo de Drones Inteligentes Aplicados à Navegação

Alexandre B. Lugli, Caio Pereira Bonvicine, Breno Veloso Eboli, Lucas Pereira Aguiar, Luiz Fernando D. R. Santos, Samuel Silva Oliveira, Wanderson E. Saldanha, Eugenio M. J. da Silva, Egídio R. Neto, Rayana Fernanda S. Silva

Instituto Nacional de Telecomunicações - Inatel

baratella@inatel.br, c.pereira@gec.inatel.br, brenoeboli@gmail.com, lucas.aguiar@gea.inatel.br, luizf.santos@inatel.br, samuel.oliveira@gea.inatel.br, wsaldanha@inatel.br, eugenio@inatel.br, egidio@inatel.br, rayana.fernanda@geb.inatel.br

***Abstract*— This work presents the development of an autonomous drone platform designed for competition scenarios, where flight stability, precise control, and reliable navigation are essential requirements. The prototype integrates a Raspberry Pi 4 with a Pixhawk-based flight controller running the PX4 autopilot, allowing configuration and testing through QGroundControl and the MAVLink protocol. The main objective is to establish a solid foundation for autonomous behavior, particularly the ability to navigate predetermined paths and execute basic tasks without manual intervention. Although the current system focuses on controlled flights, the architecture was planned to support future enhancements. Among these, the integration of ROS 2 stands out, as it will enable distributed processing and the implementation of more advanced navigation strategies. The platform serves not only as a functional prototype for competitions but also as a basis for future research involving autonomous mapping, obstacle detection, and mission coordination.**

***Index Terms*— Autonomous Drones, PX4, ROS2, QGroundControl, Pixhawk.**

***Resumo*— Este trabalho apresenta o desenvolvimento de uma plataforma de drone autônomo voltada para cenários de competição, nos quais estabilidade de voo, controle preciso e navegação confiável são requisitos fundamentais. O protótipo integra uma Raspberry Pi 4 e um controlador Pixhawk executando o firmware PX4, com todo o processo de configuração e testes realizado por meio do QGroundControl e protocolo MAVLink (Micro Air Vehicle Link). O objetivo central é estabelecer uma base consistente para comportamentos autônomos, especialmente a navegação por trajetórias definidas e a execução de tarefas básicas sem intervenção humana. Embora o sistema atual esteja direcionado a voos controlados, sua arquitetura foi planejada para permitir ampliações futuras. Entre elas, destaca-se a integração do ROS 2 (Robot Operating System), que possibilitará processamento distribuído e estratégias mais avançadas de navegação. A plataforma visa atender às demandas das competições e servir de base para pesquisas futuras envolvendo mapeamento autônomo, detecção de obstáculos e coordenação de missões.**

***Palavras Chave*— Drones autônomos, Pixhawk, PX4, QGroundControl, ROS 2.**

I. Introdução

Os drones autônomos representam uma tecnologia com elevado potencial de desenvolvimento, devido ao seu vasto campo de pesquisa e às aplicações críticas em setores como a agricultura de precisão e o monitoramento industrial [1]. Atualmente, o desenvolvimento desses sistemas exige o domínio de arquiteturas robustas de controle de voo, como o PX4 [2], que fundamenta o processo de automação de veículos aéreos não tripulados (VANTs).

Os estudos atuais visam preencher lacunas relacionadas a quadricópteros desenvolvidos em ambiente laboratorial, buscando avançar essa tecnologia para que esses sistemas se tornem capazes de executar missões autônomas complexas por meio de algoritmos de navegação para veículos aéreos não tripulados (VANTs), também conhecidos como *Unmanned Aerial Vehicles* (UAVs) [3].



Fig. 1. Exemplo de drone utilizado para entrega de pacotes

A autonomia operacional do veículo é alcançada pela integração de sistemas de posicionamento global (GPS) e computadores de bordo de alto desempenho, como a Jetson Nano ou a Raspberry Pi 4. A comunicação entre esses módulos e o controlador de voo Pixhawk [4] é viabilizada pelo protocolo MAVLink [5], permitindo que algoritmos de inteligência artificial gerenciem o comportamento do drone em tempo real.

No contexto acadêmico e competitivo, essas plataformas permitem a validação de estratégias como o mapeamento autônomo e o reconhecimento de gestos. Para tais aplicações, destaca-se a integração com o *Robot Operating System 2* (ROS 2) [6] e o uso de bibliotecas como o MAVSDK [7], fundamentais para garantir a modularidade e a escalabilidade do sistema.

Devido à simplicidade do modelo de quadricóptero em sistemas de decolagem e pouso vertical (*Vertical Takeoff and Landing* – VTOL) e à sua versatilidade, essa configuração consolidou-se como uma das mais utilizadas em aplicações práticas. Seu funcionamento baseia-se no controle da aceleração vertical por meio da geração diferencial de empuxo nas hélices.

Estudos demonstram que a modelagem precisa do comportamento aerodinâmico pode ser realizada a partir das equações de Navier-Stokes, considerando efeitos de turbulência e confinamento em ambientes *indoor* [8]. Além disso, a configuração simétrica com motores em rotações opostas reduz efeitos giroscópicos indesejados, contribuindo para a estabilidade mesmo sob perturbações externas ou falhas parciais.

360

Diante dessas características, o modelo de quadricóptero apresenta vantagens em termos de estabilidade, previsibilidade dinâmica e adaptação a diferentes cenários operacionais.

II. Materiais e Métodos

A montagem do drone foi estruturada de modo a separar as funções de voo, processamento e alimentação elétrica, com o objetivo de facilitar a manutenção e possibilitar futuras substituições de componentes.

A. Estrutura e Reforços Mecânicos

Utilizou-se o chassi modelo S500, composto por uma base central em fibra de vidro e hastes de polímero. Por se tratar de uma estrutura leve e de baixo custo, mostrou-se adequada para testes em ambiente laboratorial. Entretanto, devido à fragilidade das conexões estruturais originais em situações de pouso brusco, foram projetadas e instaladas placas de reforço em alumínio nas regiões de maior esforço mecânico do chassi.

Essa modificação garantiu maior integridade estrutural durante a fase experimental, sem impacto significativo no peso total da aeronave.

B. Sistema de Propulsão e Controle de Voo

O sistema de propulsão é composto por quatro motores do tipo *brushless* Multistar 2212 com constante de velocidade de 920 kv. Para garantir maior confiabilidade elétrica e organização do sistema, os controladores eletrônicos de velocidade (*Electronic Speed Controllers* – ESCs) foram conectados diretamente à placa de distribuição de energia do chassi.

A controladora de voo Pixhawk FMU v2 foi posicionada no centro geométrico da aeronave, de modo a alinhar a Unidade de Medição Inercial (*Inertial Measurement Unit* – IMU) ao centro de gravidade do sistema. Essa configuração contribui para a melhoria da estabilidade e do desempenho dos algoritmos de controle implementados no firmware PX4 [4, 2].

A configuração e calibração dos sensores foram realizadas por meio do software QGroundControl [9].

C. Unidade de Processamento e Autonomia

Para o processamento das tarefas autônomas, utilizou-se uma Raspberry Pi 4 como computador de bordo. A comunicação entre a unidade de processamento e a controladora de voo é realizada por meio do protocolo MAVLink [5].

O desenvolvimento dos comandos de controle foi realizado com o uso da biblioteca MAVSDK [7], que permite a implementação de rotinas de voo automatizadas por meio do envio de comandos ao piloto automático.

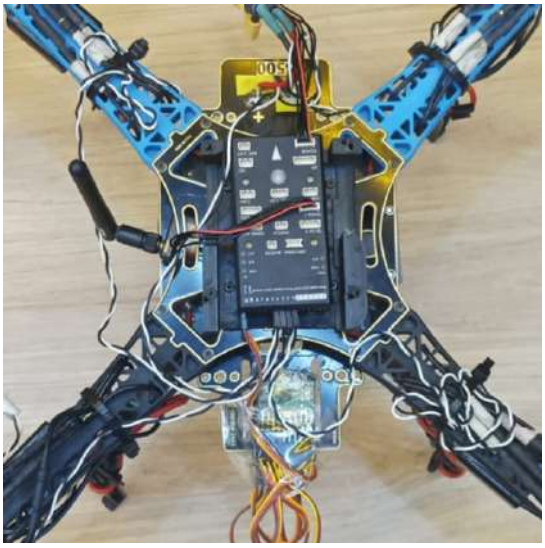


Fig. 2. Vista superior da disposição dos componentes eletrônicos e reforços estruturais do drone.

III. Testes e Resultados

Os testes do sistema foram conduzidos de forma incremental, com o objetivo de avaliar o desempenho estrutural, a estabilidade de voo e a integração entre os módulos de controle manual e autônomo. Além de análises qualitativas, foram coletadas métricas quantitativas relacionadas à latência de comunicação, ao tempo de resposta do sistema, à taxa de transmissão de dados via protocolo MAVLink [5] e à estabilidade de voo em ambiente controlado.

A. Desempenho de Potência e Estabilidade Mecânica

Inicialmente, o sistema foi alimentado por uma bateria LiPo 3S (11,1 V). Durante os testes de bancada, verificou-se que a tensão fornecida era insuficiente para gerar o empuxo necessário para a decolagem, considerando o peso total da aeronave. A substituição por uma bateria LiPo 4S (14,8 V) resultou em desempenho adequado, permitindo que os motores operassem na faixa de rotação necessária para sustentação do voo.

Nos testes iniciais de voo, utilizou-se um controle genérico baseado em um controle de Xbox One conectado ao computador. Observou-se uma variação angular superior a $\pm 15^\circ$ nos eixos de *roll* e *pitch*, atribuída à baixa precisão dos dispositivos de entrada. A substituição por um rádio transmissor dedicado (HK-7X v2) reduziu essa variação para aproximadamente $\pm 6^\circ$.

Após o balanceamento das hélices e o ajuste dos ganhos do controlador proporcional-integral-derivativo (PID) no firmware PX4 [2], a oscilação em voo pairado (*hover*) foi reduzida para cerca de $\pm 2^\circ$, indicando melhoria significativa na estabilidade do sistema.

B. Comunicação MAVLink e Processamento Embarcado

A validação da comunicação entre a Raspberry Pi 4 e a controladora Pixhawk foi realizada em ambiente controlado. Um dos pontos avaliados foi o desempenho de diferentes bibliotecas de controle baseadas no protocolo MAVLink.

Os testes com a biblioteca DroneKit [10] apresentaram taxa de falha de aproximadamente 70%, associada à incompatibilidade com versões mais recentes do protocolo MAVLink. Em

contrapartida, a biblioteca MAVSDK [7] apresentou comportamento estável e confiável, conforme também observado em estudos recentes [11].

As principais métricas de desempenho obtidas foram:

- **Latência média:** 28 ms (pico de 42 ms);
- **Taxa de telemetria:** aproximadamente 20 Hz para mensagens de estado e sensores;
- **Tempo de resposta:** cerca de 120 ms para o comando de armamento de motores (*arming*).

TABELA I
RESUMO DAS MÉTRICAS DE DESEMPENHO DO SISTEMA

Métrica	Valor Observado
Latência média	28 ms
Latência máxima	42 ms
Tempo de resposta (arming)	120 ms
Taxa de telemetria	20 Hz
Estabilidade em hover (roll/pitch)	$\pm 2^\circ$

Os resultados obtidos indicam que a arquitetura proposta é capaz de suportar operações com requisitos de tempo real. A integração entre o firmware PX4 e a biblioteca MAVSDK, executada na Raspberry Pi 4, demonstrou confiabilidade na execução de comandos, sem ocorrência de perdas de pacotes ou falhas críticas de comunicação durante os testes realizados.

IV. TRABALHOS FUTUROS

Como continuidade deste trabalho, propõe-se a evolução da arquitetura do sistema por meio da integração com o *Robot Operating System 2* (ROS 2) [6, 12]. Essa abordagem permitirá a modularização das funcionalidades do drone, possibilitando maior flexibilidade no desenvolvimento, depuração e expansão do sistema.

A partir dessa base, pretende-se implementar navegação autônoma baseada em *waypoints*, na qual o veículo será capaz de executar missões a partir de coordenadas geográficas predefinidas. Para isso, serão exploradas técnicas de fusão sensorial, planejamento de trajetória e controle em malha fechada, visando maior precisão e robustez operacional [3].

Em uma etapa posterior, prevê-se a evolução da plataforma física, com a adoção de uma aeronave de maior porte e maior capacidade de carga. Paralelamente, o sistema embarcado deverá ser atualizado para plataformas de maior desempenho, como dispositivos da família NVIDIA Jetson, permitindo a execução de algoritmos mais complexos.

Nesse contexto, futuras investigações incluirão o uso de técnicas de mapeamento tridimensional, navegação em ambientes sem GPS e reconhecimento de padrões por meio de visão computacional. Tais avanços visam ampliar o nível de autonomia do sistema, possibilitando sua aplicação em cenários mais complexos.

Dessa forma, espera-se consolidar a plataforma desenvolvida como uma base para estudos avançados em navegação autônoma e sistemas embarcados aplicados a veículos aéreos não tripulados.

V. CONCLUSÃO

Este trabalho apresentou o desenvolvimento e a validação de uma plataforma de drone voltada à navegação autônoma, com foco na integração entre hardware embarcado e software de controle. A arquitetura proposta, baseada na controladora Pixhawk, no firmware PX4 [2] e na biblioteca MAVSDK [7], demonstrou-se adequada para aplicações que demandam estabilidade e confiabilidade em operações de voo.

Os resultados experimentais evidenciaram que decisões de projeto relacionadas tanto ao hardware quanto ao software influenciam diretamente o desempenho do sistema. Ajustes estruturais, calibração adequada e escolha de ferramentas de controle contribuíram para a obtenção de um comportamento estável em voo, com níveis reduzidos de oscilação.

Além disso, a análise comparativa entre bibliotecas de controle indicou que a utilização do MAVSDK proporciona maior robustez na comunicação via MAVLink [5], viabilizando a execução confiável de comandos em tempo real.

Dessa forma, a plataforma desenvolvida constitui uma base funcional para a implementação de sistemas autônomos mais avançados, podendo ser utilizada como suporte para pesquisas futuras em navegação, controle e inteligência embarcada em veículos aéreos não tripulados.

REFERÊNCIAS

[1] Umamaheswara Rao Mogili e BBVL Deepak. "Review on application of drone systems in precision agriculture". Em: *Procedia computer science* 133 (2018), pp. 502–509.

[2] Lorenz Meier, Dominik Honegger e Marc Pollefeys. "PX4: A node-based multithreaded open source robotics framework for unmanned vehicles". Em: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015, pp. 6235–6242.

[3] Chad Goerzen, Zhaodan Kong e Bernard Mettler. "A survey of motion planning algorithms from the perspective of autonomous UAV guidance". Em: *Journal of Intelligent and Robotic Systems* 57 (2009), pp. 65–100.

[4] Lorenz Meier, Petri Tanskanen, Friedrich Fraundorfer e Marc Pollefeys. "Pixhawk: A system for autonomous flight using onboard computer vision". Em: *2011 IEEE International Conference on Robotics and Automation*. IEEE. 2011, pp. 2992–2997.

[5] Lorenz Meier. *MAVLink: Micro air vehicle communication protocol*. Disponível em: https://mavlink.io/. 2011.

[6] Steven Macenski, Tully Foote, Brian Gerkey, Chris Lalancette e William Woodall. "Robot Operating System 2: Design, architecture, and uses in the wild". Em: *Science Robotics* 7.66 (2022), eabm6074.

[7] Dronecode Foundation. *MAVSDK: A Library for the MAVLink Protocol*. Accessed: 2026-03-12. 2024. URL: https://mavsdk.mavlink.io/.

[8] Teppo Luukkonen. "Modelling and control of quadcopter". Em: *Independent research project in applied mathematics, Aalto University* 22.22 (2011), p. 12.

[9] Dronecode Foundation. *QGroundControl User Guide*. Accessed: 2026-03-12. 2024. URL: https://docs.qgroundcontrol.com/.

[10] 3DR. *DroneKit-Python: Open source python library for MAVLink*. Legacy software, accessed: 2026-03-12. 2018.

[11] Alexandru Stanciu et al. "Evaluation of MAVLink Communication Protocols for UAVs in Research Environments". Em: *Sensors* 23.4 (2023).

[12] Anis Koubaa et al. "ROS 2 for drone applications: Architecture, implementation and performance". Em: *Robot Operating System (ROS)*. Springer, 2019, pp. 3–34.

AUTORES



Caio Pereira Bonvicine é graduando em Engenharia de Computação pelo Instituto Nacional de Telecomunicações (Inatel). É membro da equipe RobotBulls, dedicada à robótica e inteligência artificial, onde atua no desenvolvimento e na pesquisa na área de drones autônomos.



Breno Veloso Eboli é graduando em Engenharia de Computação pelo Instituto Nacional de Telecomunicações (Inatel). É membro da equipe RobotBulls, dedicada à robótica e inteligência artificial, onde atua no desenvolvimento e na pesquisa na área de drones autônomos.



Egídio Raimundo Neto possui graduação em Engenharia Elétrica pelo Instituto Nacional de Telecomunicações (2011) e mestrado em Telecomunicações pela mesma instituição (2014). Atualmente, atua como professor no Instituto Nacional de Telecomunicações, nos cursos de Engenharia e Tecnologia. Suas principais áreas de interesse incluem Sistemas de Controle Clássico, Discreto e Moderno, Redes Industriais, Robótica, Drones Autônomos, Sistemas Eletrônicos de Potência e Smart Grid.



Eugênio Moreira Jonas da Silva é engenheiro de Controle e Automação pelo Inatel e atua como especialista em projetos de controle e automação na mesma instituição. Integra a coordenação do laboratório de robótica e inteligência artificial do Inatel, além de liderar a organização da Olimpíada Nacional de Robótica 2025, no âmbito do projeto Nativos do Futuro. Também participou da coordenação da FETIN 2024 e atualmente integra a organização da FETIN 2025. Atua ainda como professor no SENAI, contribuindo para cursos técnicos nas áreas de automação e eletrônica.



Lucas Pereira Aguiar é graduando em Engenharia de Computação pelo Instituto Nacional de Telecomunicações (Inatel). É membro da equipe RobotBulls, dedicada à robótica e inteligência artificial, onde atua no desenvolvimento e na pesquisa na área de drones autônomos.



Samuel Silva Oliveira é graduando em Engenharia de Controle e Automação pelo Instituto Nacional de Telecomunicações (Inatel). É bolsista de iniciação científica no laboratório de robótica e inteligência artificial (RobotBulls), onde atua no desenvolvimento de sistemas de drones autônomos. Possui interesse nas áreas de Robótica, Automação Industrial e Inteligência Artificial.



Wanderson Eleutério Saldanha atua com pesquisa e desenvolvimento desde 1996, com experiência em gerenciamento de projetos, capacitação de profissionais e gestão de equipes. Possui formação técnica em eletrônica pela Escola Técnica de Eletrônica Francisco Moreira da Costa (1995), graduação em Engenharia Elétrica pelo Instituto Nacional de Telecomunicações (2000), especialização em Docência do Ensino Superior pelo Centro de Ensino Superior em Gestão, Tecnologia e Educação (2012) e mestrado em Engenharia Elétrica, na área de Automação e Sistemas Elétricos Industriais, pela Universidade Federal de Itajubá (2016). Atualmente, desenvolve pesquisa em nível de doutorado. Atua como professor no Instituto Nacional de Telecomunicações, nos cursos de graduação e pós-graduação, além de orientar projetos acadêmicos e de iniciação científica.

Automatic Classification of Coffee Fruits Using Machine Learning and Computer Vision

Henrique R. Mendonça, Samuel B. Mafra
Instituto Nacional de Telecomunicações (Inatel)
Santa Rita do Sapucaí, Brazil
henrique.rodrigues@ges.inatel.br, samuelbmafra@inatel.br

***Abstract*—This paper presents an automatic coffee fruit ripeness detection system based on real-time object detection, designed to support agricultural decision-making in mobile field conditions. A labeled dataset exported from Roboflow was used to train and evaluate three YOLO-based detectors under comparable experimental settings: YOLOv8n, YOLOv11n, and YOLOv8s. The evaluation was conducted on a validation subset containing 201 images and 3189 annotated instances across five maturity classes (dry, overripe, ripe, semi-ripe, and unripe). YOLOv8n achieved mAP@0.5 = 0.770, mAP@0.5:0.95 = 0.705, precision = 0.662, recall = 0.833, and 5.0 ms inference time per image. YOLOv11n achieved mAP@0.5 = 0.781 and mAP@0.5:0.95 = 0.706 with precision = 0.762, but lower recall = 0.721 and higher latency (6.0 ms). YOLOv8s achieved the highest mAP@0.5:0.95 = 0.711 at a substantially higher inference time (9.7 ms). Considering the accuracy–latency trade-off and the importance of high recall for agricultural detection, YOLOv8n was selected as the most balanced model for integration into a Flutter-based on-device inference pipeline.**

***Index Terms*—Coffee ripeness, Computer Vision (CV), Deep Learning (DL), Convolutional Neural Networks (CNN), object detection, YOLO, mobile inference, precision agriculture.**

I. Introduction

Coffee production quality and economic value are strongly influenced by the ripeness stage of harvested fruits. In practical field operations, maturity assessment is often performed through manual visual inspection, which is subjective, labor-intensive, and sensitive to worker experience, fatigue, and environmental conditions. Inconsistent harvesting that mixes unripe and overripe fruits may negatively affect beverage quality and processing efficiency[1].

Advances in Computer Vision (CV) and Deep Learning (DL) have enabled automated visual inspection in agriculture, where Convolutional Neural Networks (CNNs) have become the dominant approach for learning discriminative image representations [2]. In particular, object detection provides a suitable formulation for fruit monitoring tasks because it localizes each instance and assigns a maturity label, enabling both qualitative assessment and quantitative estimation. Among detectors, the You Only Look Once (YOLO) family introduced a single-stage paradigm that significantly improved real-time performance by predicting bounding boxes and class probabilities in a single forward pass [3]. Subsequent developments improved accuracy while maintaining speed, including YOLOv4 [4] and recent Ultralytics implementations [5].

In [6], the authors implement a Smart Farming System based on a three-layered architecture (Agriculture Perception, Edge Computing, and Data Analytics). Stakeholders, researchers, and engineers validated the Smart Farming Solution through a case study of a Colombian coffee farm. Despite the progress, agricultural deployments frequently require an explicit analysis of the accuracy–latency trade-off, since the final system is constrained by mobile hardware and must operate reliably without cloud connectivity. Moreover, many applied studies evaluate a single model or do not compare lightweight variants under comparable settings, which limits the interpretation of model selection decisions.

In this context, this work proposes and evaluates an automatic coffee fruit ripeness detection pipeline targeting on-device inference. Three lightweight YOLO variants are trained and validated on the same dataset split and reported using standard detection metrics. The selected model is designed for integration into a Flutter-based mobile application pipeline, enabling real-time use in field conditions.

II. Theoretical Foundation

A. Object Detection Background

Object detection aims to localize and classify multiple object instances within an image. Compared to image-level classification, detection requires both spatial predictions and semantic labels. Modern detectors rely on deep neural networks to extract hierarchical features and to regress bounding box coordinates alongside class scores.

Detectors are commonly categorized into two-stage and single-stage families. Two-stage detectors, such as Faster R-CNN, first generate region proposals and then classify them, often achieving strong accuracy at higher computational cost. Single-stage detectors, including YOLO, directly predict bounding boxes and class probabilities, enabling real-time operation, which is advantageous for edge deployments.

B. YOLO Principle

Modern object detection architectures typically rely on deep convolutional feature extractors combined with multi-scale feature aggregation mechanisms. The YOLO family has evolved through multiple generations, each introducing architectural improvements that enhance both accuracy and computational efficiency.

Earlier versions of YOLO relied on Darknet-based backbones for feature extraction. More recent implementations adopt improved convolutional structures and feature pyramid strategies to better capture spatial hierarchies within images. Feature Pyramid Networks (FPN) and Path Aggregation Networks (PAN) allow

the model to combine low-level spatial information with high-level semantic features, improving detection performance for objects of different sizes.

Another important advancement introduced in modern YOLO architectures is the use of improved loss functions for bounding box regression. Traditional bounding box losses based on mean squared error have been replaced by IoU-based formulations such as Generalized IoU (GIoU), Distance IoU (DIoU), and Complete IoU (CIoU). These losses improve training stability and lead to more accurate localization performance.

Furthermore, modern implementations integrate automatic anchor optimization and dynamic label assignment strategies during training. These mechanisms reduce the need for manual hyperparameter tuning and improve convergence behavior when training on datasets with varying object scales and class distributions.

Due to these improvements, YOLO-based detectors have become widely adopted in real-time applications such as autonomous driving, surveillance systems, industrial inspection, and precision agriculture.

$$\hat{y} = f_\theta(x), \tag{1}$$

where θ denotes the learned parameters and $\hat{y}$ contains bounding box coordinates, objectness scores, and class probabilities. Architectural improvements across versions focus on more effective feature extraction, feature aggregation across scales, and loss formulations that improve localization and classification performance while preserving inference efficiency [3, 4, 5].

C. Evaluation Metrics

Detection quality is typically assessed through Intersection over Union (IoU), which measures the overlap between a predicted bounding box B_{pred} and a ground-truth box B_{gt}:

$$IoU = \frac{Area(B_{pred} \cap B_{gt})}{Area(B_{pred} \cup B_{gt})}. \tag{2}$$

Precision and recall quantify detection correctness:

$$Precision = \frac{TP}{TP + FP}, \tag{3}$$

$$Recall = \frac{TP}{TP + FN}, \tag{4}$$

where TP, FP, and FN represent true positives, false positives, and false negatives, respectively. Average Precision (AP) is computed from the area under the precision–recall curve, and mean Average Precision (mAP) is obtained by averaging AP across classes [7]. In this work, mAP@0.5 and mAP@0.5:0.95 are reported, where the latter averages AP over multiple IoU thresholds for stricter localization assessment.

III. Proposed System Model

A. System Overview

The proposed system comprises two operational stages: an offline training stage and an on-device inference stage. The offline stage includes dataset preparation, training, validation, comparative analysis, and export of the selected model. The on-device stage performs camera-based acquisition, preprocessing, inference, post-processing, and result visualization within a mobile application environment.

Figure 6 illustrates the overall pipeline used in this study, connecting the training workflow to the mobile deployment stage.

B. Training Workflow

Roboflow was used for dataset management, annotation review, and export to YOLO format. The dataset is exported in YOLO format, consisting of images and corresponding label files with normalized bounding box coordinates and class indices. The training procedure uses the Ultralytics framework, which provides standardized training and validation routines, automatic data augmentation, and logging of quantitative metrics [5]. All compared models are trained under comparable settings to ensure a fair evaluation, and the model choice is guided by both detection metrics and inference latency.

C. Mobile Inference Pipeline

The mobile application is developed using Flutter, enabling cross-platform deployment with a single codebase. During runtime, the camera module captures images and the preprocessing stage resizes and normalizes the input to match the model requirements. The inference stage loads the selected detector and produces bounding box predictions with class scores. The post-processing stage applies confidence filtering and Non-Maximum Suppression (NMS) to remove redundant detections. The final results are rendered on-screen with bounding boxes and maturity labels, providing immediate feedback to the user.

The on-device inference process can be represented as:

$$\hat{y} = f_\theta(T(x)), \tag{5}$$

where x is the captured image and $T(\cdot)$ is the preprocessing transformation. Figure 2 provides a high-level view of this runtime pipeline.

IV. Experimental Methodology

A. Dataset Description

The dataset used in this work was constructed from images of coffee plants captured under real agricultural field conditions. The images contain significant variability in illumination, background vegetation, fruit density, and partial occlusions caused by leaves and branches.

Each image was manually annotated using bounding boxes that indicate the spatial location of individual coffee fruits. The annotations follow the YOLO labeling format, in which each object instance is represented by a class identifier and normalized coordinates corresponding to the center position, width, and height of the bounding box.

The dataset contains five maturity categories: dry, overripe, ripe, semi-ripe, and unripe. These classes correspond to different physiological stages of coffee fruit development and are visually distinguishable mainly through color variation ranging from green to deep red and dark brown tones.

Figure 3 illustrates representative samples from the dataset, including annotated bounding boxes for different maturity classes. These examples highlight the visual differences among

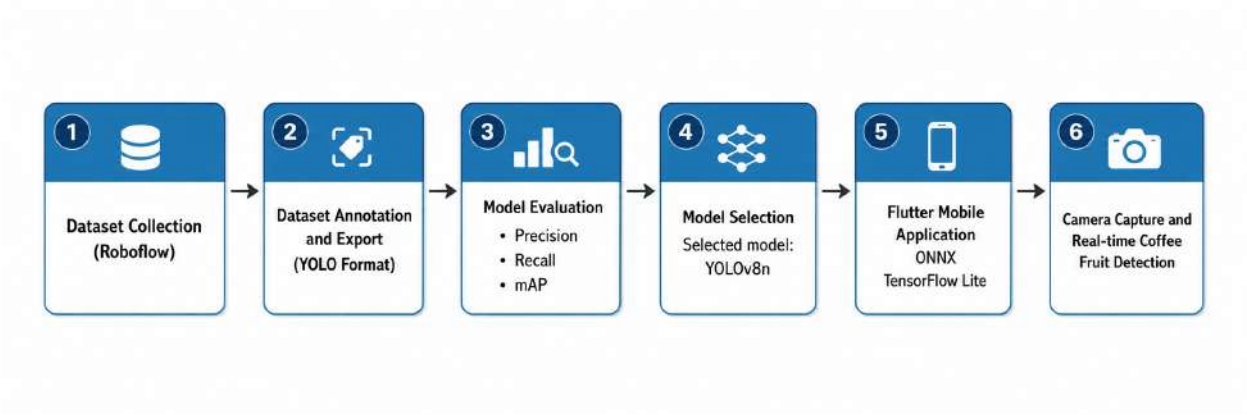


Fig. 1. Overall architecture of the proposed system, including offline training and on-device inference stages.

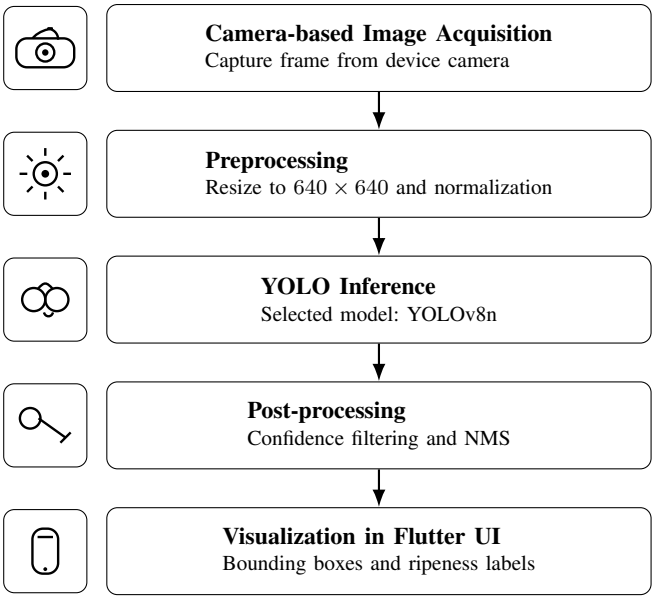


Fig. 2. On-device inference pipeline executed within the Flutter mobile application.

categories and demonstrate the complexity of the detection task under real field conditions, including variations in lighting, occlusion, and fruit clustering.

A total of 3189 annotated fruit instances were used during validation, distributed across 201 images. As commonly observed in agricultural datasets, the distribution of samples across classes is not uniform. The "unripe" category contains the largest number of instances, while the "overripe" class contains comparatively fewer samples. The dataset was divided into training, validation, and test sets using a 70/20/10 split.

Class imbalance is a relevant factor in object detection tasks because models tend to prioritize dominant classes during training. Therefore, evaluating class-wise performance becomes essential to understand potential detection limitations and to guide future dataset expansion strategies.

B. Training Configuration

All experiments were conducted using GPU acceleration on an NVIDIA Tesla T4 environment provided by Google Colab. The models were trained using the Ultralytics implementation of the YOLO framework, which provides an optimized training pipeline and standardized evaluation procedures.

The input resolution was fixed at 640×640 pixels and the batch size was set to 16. Stochastic Gradient Descent (SGD) was used as the optimization algorithm with an initial learning rate of 0.01, momentum of 0.937, and weight decay of 5×10^{-4}. These values follow commonly adopted training settings for YOLO-based detectors and ensure stable convergence.

During training, data augmentation techniques were automatically applied by the Ultralytics framework in order to improve model generalization. These augmentations included horizontal flipping, scaling, mosaic augmentation, color space perturbations, and random translations. Such transformations help the model become more robust to variations in illumination, viewpoint, and fruit occlusions that frequently occur in real agricultural environments.

The YOLOv8n and YOLOv11n models were trained for 100 epochs to allow full convergence of lightweight architectures. The YOLOv8s model, which contains a larger number of parameters, was trained for 50 epochs to evaluate the impact of increased model capacity while maintaining feasible training time.

Validation was performed at the end of each epoch using a validation subset containing 201 images and 3189 annotated fruit instances. Performance was monitored using precision, recall, mAP@0.5, and mAP@0.5:0.95 metrics.

In order to prevent overfitting and improve training efficiency, an early stopping strategy was employed during the training process. This mechanism monitors the validation performance and automatically interrupts training when no significant improvement is observed over a predefined number of epochs. This approach is particularly useful in scenarios with class imbalance, where certain classes (such as "overripe") may converge more slowly or exhibit unstable learning behavior. By applying early stopping, the training process avoids unnecessary epochs and helps preserve model generalization capability.

V. Experimental Results and Discussion

A. Quantitative Results

Table I summarizes the quantitative comparison across models using precision, recall, mAP@0.5, mAP@0.5:0.95, and inference latency. These metrics collectively represent classification correctness, missed detections, localization quality, and runtime feasibility for deployment.

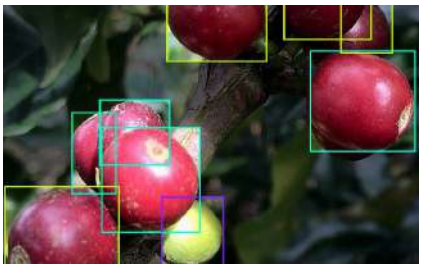


Fig. 3. Example annotations showing bounding boxes for coffee fruits, where different collors represent distinct maturity classes.

TABLE I
PERFORMANCE COMPARISON OF YOLO VARIANTS

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Inference (ms)
YOLOv8n	0.662	0.833	0.770	0.705	5.0
YOLOv11n	0.762	0.721	0.781	0.706	6.0
YOLOv8s	0.713	0.761	0.776	0.711	9.7

Beyond overall performance metrics, it is important to analyze how each architecture behaves across different maturity classes. The class-wise evaluation indicates that the ”unripe” category achieved the highest detection performance across all evaluated models. This outcome is expected due to the larger number of annotated instances available for this class within the dataset.

Conversely, the ”overripe” class presented the lowest detection performance. This behavior may be attributed to two main factors. First, the visual appearance of overripe fruits may resemble that of adjacent maturity stages, such as ripe or dry fruits, which increases classification ambiguity. Second, the smaller number of training samples for this category may limit the model's ability to learn discriminative features. This behavior is consistent with known effects of class imbalance in object detection tasks, where minority classes tend to exhibit lower detection performance due to reduced representation during training. Despite this, the model maintains stable overall performance, indicating that the learned feature representations remain robust across the dataset.

Another important observation relates to the trade-off between precision and recall across the evaluated architectures. YOLOv11n achieved the highest precision values, indicating that when the model predicted a fruit detection, it was more likely to be correct. However, its lower recall indicates that some fruits present in the image were not detected.

In contrast, YOLOv8n demonstrated higher recall values, meaning that the model was able to detect a larger proportion of the fruits present in the images. In agricultural monitoring tasks, higher recall can be advantageous because missing detections may lead to inaccurate estimations of crop maturity distribution.

Finally, the inference latency analysis reinforces the importance of selecting architectures that balance detection performance with computational efficiency. The YOLOv8n model required approximately 5.0 milliseconds per image, which makes it suitable for real-time inference on mobile devices. In comparison, the YOLOv8s model exhibited significantly higher latency due to its increased number of parameters and computational complexity.

B. Accuracy–Latency Trade-off

YOLOv11n achieved the highest precision and slightly higher mAP@0.5 than YOLOv8n, but its recall decreased substantially, which is an important limitation in agricultural contexts where missed detections may affect maturity assessment. YOLOv8s achieved the best mAP@0.5:0.95, suggesting improved localization at stricter IoU thresholds, but its latency increased to 9.7 ms per image, which can significantly impact real-time responsiveness on mobile devices.

Figure 4 compares the detection performance of the evaluated YOLO architectures using the mAP@0.5 and mAP@0.5:0.95 metrics. YOLOv11n achieved the highest mAP@0.5 (0.781), indicating slightly better detection performance when considering a single IoU threshold. However, the results for mAP@0.5:0.95 are very close among the models, with YOLOv8s achieving the highest value (0.711), suggesting slightly better localization performance under stricter IoU thresholds. YOLOv8n obtained competitive results with mAP@0.5 = 0.770 and mAP@0.5:0.95 = 0.705 while maintaining lower computational complexity. Considering the small differences in accuracy and the importance of inference efficiency for mobile deployment, YOLOv8n represents a balanced solution for real-time agricultural applications.

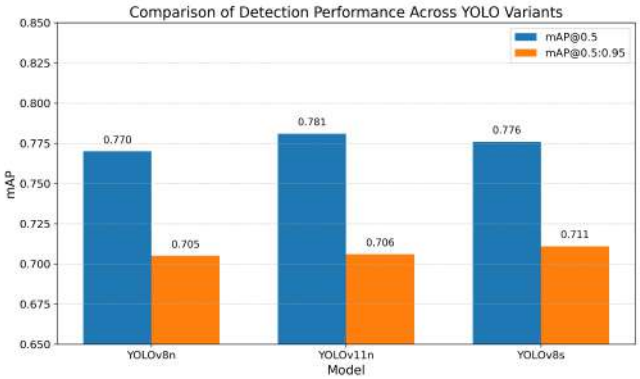


Fig. 4. Comparison of mAP@0.5:0.95 among evaluated YOLO variants.

C. Class-wise Observations

The “overripe” category remained the most challenging class across models, which can be explained by visual similarity

with adjacent maturity stages and limited class representation compared to the dominant “unripe” class. This suggests that future improvements may require dataset balancing, targeted augmentation, or additional samples captured under varied illumination and background conditions. Figure 5 shows the confusion matrix of the selected model to support a more detailed error analysis.

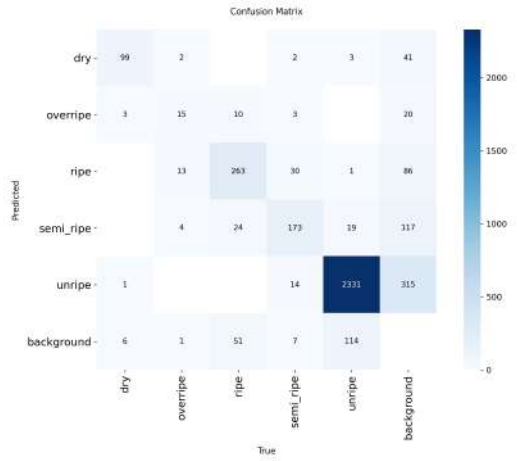


Fig. 5. Confusion matrix of YOLOv8n on the validation set.

D. Model Selection for Deployment

Considering the joint objective of detection reliability and mobile feasibility, YOLOv8n was selected as the final model. While YOLOv8s and YOLOv11n achieved marginal gains in specific metrics, YOLOv8n provided the best overall balance due to its higher recall and lower latency, which are critical for on-device field operation.

Figure 6 presents a real-world example of the deployed mobile application during operation. The system performs real-time detection of coffee fruits directly on the device, displaying bounding boxes and maturity labels over the captured image. This result demonstrates the practical feasibility of the proposed approach under field conditions, reinforcing the effectiveness of the selected model for on-device agricultural applications.

VI. Conclusion

This work presented a computer vision system for automatic coffee fruit ripeness detection based on lightweight YOLO object detection architectures. The proposed pipeline integrates dataset preparation, model training, quantitative evaluation, and mobile deployment considerations in order to support real-time agricultural monitoring applications.

Three YOLO variants were evaluated under comparable experimental conditions: YOLOv8n, YOLOv11n, and YOLOv8s. The comparative analysis demonstrated that although YOLOv11n achieved higher precision and YOLOv8s achieved slightly better localization performance, YOLOv8n provided the most balanced trade-off between recall, detection accuracy, and inference latency.

High recall is particularly important for agricultural monitoring systems because missing fruit detections may lead to inaccurate estimations of crop maturity distribution. In this context, YOLOv8n demonstrated strong detection performance



Fig. 6. Real-time coffee fruit maturity classification displayed in the mobile application.

while maintaining the lowest inference latency among the evaluated architectures.

The selected model was integrated into a mobile application developed using the Flutter framework, enabling real-time fruit detection directly on mobile devices without requiring cloud connectivity. Such capability is particularly relevant for agricultural environments where internet access may be limited or unreliable.

Overall, the experimental results indicate that lightweight deep learning detectors can effectively support precision agriculture applications by automating fruit monitoring tasks and assisting farmers in decision-making processes related to harvesting and crop management.

Future work may include expanding the dataset with additional samples captured under diverse environmental conditions, applying advanced data augmentation strategies to improve detection robustness, and exploring model optimization techniques such as quantization, pruning, and hardware acceleration to further reduce inference latency for edge deployment. references

References

[1] Valdeir Viana Freitas, Larissa Borges, Márcia Vidigal, Marcelo Santos, and Paulo Stringheta. “Coffee: A comprehensive overview of origin, market, and the quality process”. In: *Trends in Food Science Technology* 146 (Feb. 2024), p. 104411. DOI: 10.1016/j.tifs.2024.104411.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep Residual Learning for Image Recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016).

[3] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. “You Only Look Once: Unified, Real-Time Object Detection”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

[4] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. “YOLOv4: Optimal Speed and Accuracy of Object Detection”. In: *arXiv preprint arXiv:2004.10934* (2020).

[5] Ultralytics. *Ultralytics YOLOv8 Documentation*. Online documentation. 2023.

[6] Jhonn Rodríguez, Ana Montoya-Muñoz, Carlos Rodriguez-Pabon, Javier Hoyos, and Juan Corrales. “IoT-Agro: A smart farming system to Colombian coffee farms”. In: *Computers and Electronics in Agriculture* 190 (Nov. 2021), p. 106442. DOI: 10.1016/j.compag.2021.106442.

[7] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. “The Pascal Visual Object Classes (VOC) Challenge”. In: *International Journal of Computer Vision* 88.2 (2010), pp. 303–338.

AUTORES



Henrique Rodrigues Mendonça is an undergraduate student in Software Engineering at INATEL – National Institute of Telecommunications. He worked as an intern at the ICT Lab, participating in projects related to networking technologies and the future internet, including research on quantum networks and environments based on Kubernetes and Docker. He was also an Undergraduate Research Fellow at the IoT Lab, where he worked with Machine Learning and Computer Vision. His research interests include Artificial Intelligence, Computer Vision, Machine Learning, Data Science, and the Internet of Things (IoT).



Samuel Baraldi Mafra holds a Post-Doctorate from the Federal University of Paraná (2015–2018). He has a PhD in Sciences, specializing in telecommunications and networks, from the Federal Technological University of Paraná (2015). He holds a Master’s degree in Electrical Engineering, with an emphasis on telecommunications, from the Federal University of Paraná (2012). He is an electrical engineer from the State University of Santa Catarina (2010). He is a professor and researcher at the National Institute of Telecommunications (Inatel).

Automatização da Extração de Dados de Metassuperfícies para Aprendizado de Máquina

Matheus Augusto de Faria Silveira e William O. F. Carvalho
Instituto Nacional de Telecomunicações - Inatel
matheus.augusto@ges.inatel.br, william.orivaldo@inatel.br

***Abstract*—This work presents the development of a computational framework for automated data collection from reconfigurable metasurfaces through the integration of MatLab® and CST Studio Suite®. The proposed script automatically generates different metasurface configurations, modifies the states of PIN diodes, and performs electromagnetic simulations without human intervention. The resulting scattering patterns are extracted and stored in structured files, enabling the construction of datasets for machine learning applications. Preliminary results confirm the correct communication between the software environments and the reliable extraction of simulation data. The proposed approach enables large-scale dataset generation for future artificial intelligence models aimed at the inverse design of reconfigurable metasurfaces.**

***Index Terms*—Reconfigurable metasurfaces, simulation automation, CST Studio, MATLAB, artificial intelligence**

***Resumo*—Este trabalho apresenta o desenvolvimento de uma ferramenta computacional para automatizar a coleta de dados de metassuperfícies reconfiguráveis por meio da integração entre MatLab® e CST Studio Suite®. O script desenvolvido gera automaticamente diferentes configurações da metassuperfície, altera os estados de diodos PIN e executa simulações eletromagnéticas de forma autônoma. Os padrões de espalhamento resultantes são extraídos e armazenados em arquivos estruturados, permitindo a construção de um banco de dados para aplicações em aprendizado de máquina. Resultados preliminares confirmam o funcionamento da comunicação entre os softwares e a correta extração dos dados. A ferramenta viabiliza a geração massiva de dados para o treinamento futuro de modelos de inteligência artificial voltados ao design inverso de metassuperfícies.**

***Palavras chave*—Metassuperfícies reconfiguráveis, automação de simulação, CST Studio, MATLAB, inteligência artificial.**

I. Introdução

As metassuperfícies reconfiguráveis representam um avanço significativo no controle de ondas eletromagnéticas na faixa de radiofrequência e micro-ondas, podendo ser ajustadas por meios eletrônicos, ópticos, térmicos ou mecânicos [1]. Essa flexibilidade é essencial para acompanhar a evolução das comunicações sem fio, já que cada geração de telefonia móvel adota faixas de frequência específicas de acordo com sua necessidade de cobertura e capacidade de dados [2]. As comunicações móveis de segunda geração (2G) e terceira geração (3G), por exemplo, utilizam frequências mais baixas (como 850 MHz, 900 MHz e 2100 MHz), ideais para garantir um longo alcance de sinal para chamadas de voz e tráfego básico de internet [3]. A quarta geração (4G), focada em banda larga, popularizou o uso de faixas como 700 MHz e 2,6 GHz para entregar maiores velocidades de conexão [3]. Já o 5G introduziu o uso das ondas milimétricas (faixa de 26 GHz, por exemplo) para atingir taxas de transmissão altíssimas e latência mínima, além de usar o espectro sub-6 GHz (como 3,5 GHz) para equilibrar cobertura e velocidade [2]. Para as futuras redes de comunicação móvel de sexta geração (6G), a expectativa é também explorar as frequências de teraHertz (THz), o que exigirá um controle de sinal extremamente preciso [4]. Nesse cenário de frequências cada vez mais elevadas, as metassuperfícies ganham muito destaque, sendo peças-chave para criar superfícies inteligentes reconfiguráveis (RIS, *Reconfigurable Intelligent Surfaces*) que redirecionam o sinal onde ele é mais necessário [5].

Para manipular essas ondas de forma eficiente, a utilização de componentes eletrônicos, em especial o diodo PIN, é uma excelente alternativa. Integrar diodos PIN na estrutura da metassuperfície permite alterar a resposta eletromagnética do dispositivo em tempo real, apenas alternando os componentes entre os estados de condução (ON) e isolamento (OFF) [6]. Além de proporcionar uma reconfiguração extremamente rápida, os diodos PIN apresentam baixo custo e são fáceis de integrar aos processos tradicionais de fabricação de placas de circuito impresso (PCB), nas quais as metassuperfícies são fabricadas [7, 8].

Apesar do grande potencial, o processo de design dessas estruturas é demorado e altamente complexo. A busca pela geometria ideal e pelo estado correto dos componentes ativos (como os diodos) exige que os engenheiros realizem simulações exaustivas para cada configuração desejada. A fim de contornar o problema de tempo computacional gasto em simulações eletromagnéticas demoradas, o objetivo principal deste trabalho é acelerar o desenvolvimento de metassuperfícies utilizando inteligência artificial (IA). Para que uma rede neural consiga aprender a relação inversa entre a curva de espalhamento (padrão de irradiação) obtida e a estrutura física necessária, é indispensável a construção de um banco de dados robusto. Dessa forma, este artigo detalha o desenvolvimento de uma ferramenta computacional capaz de gerar, simular e armazenar esses dados de maneira totalmente autônoma. O processo foi feito integrando a Interface de Programação de Aplicações (API, *Application Programming Interface*) do MatLab® diretamente com o simulador CST Studio Suite®. O *script* desenvolvido automatiza o desenho da geometria de forma aleatória, varia os estados dos diodos (ligado/desligado) e comanda a simulação. Como resultado principal, obteve-se um fluxo validado que extrai e salva automaticamente as curvas de espalhamento, formando um banco de dados consistente e pronto para ser utilizado no treinamento de modelos de aprendizado de máquina.

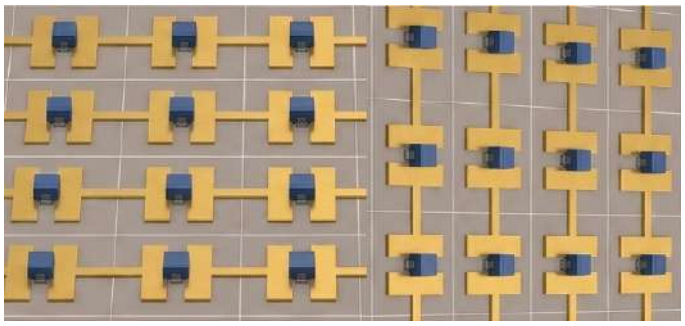


Fig. 1. Representação artística da geometria da metassuperfície. A esquerda, a disposição das camadas metálicas em forma de dipolos dispostos horizontalmente, enquanto que a direita a mesma disposição, porém verticalmente. Nas lacunas entre os dipolos são inseridos os diodos PIN para a reconfiguração da metassuperfície.

II. Metodologia

A geometria adotada baseia-se em elementos dipolares dispostos em camadas superior e inferior, interconectados por diodos PIN, como mostra a Fig. 1. O propósito dessa estrutura é permitir o controle dinâmico da frequência de ressonância e da polarização da onda eletromagnética incidente, alterando o comportamento de espalhamento da superfície sem a necessidade de modificar sua forma física.

A coleta de dados foi estruturada a partir da integração entre o ambiente MatLab® e o simulador eletromagnético CST Studio Suite®. Utilizando a API de comunicação entre os *softwares*, foi possível estabelecer o controle programático de todo o fluxo de simulação. O modelo base utilizado no projeto consiste em uma célula unitária de metassuperfície contendo dois diodos PIN, inspirada na Ref. [9]. A Fig. 2 ilustra a interface de integração. Pelo MatLab®, o *script* envia comandos para o CST Studio Suite® criar a estrutura no ambiente de simulação. O código define as dimensões e os parâmetros eletromagnéticos dos materiais utilizados. Na construção da estrutura, o substrato dielétrico escolhido é o Rogers 4003C, que apresenta uma permissividade relativa de $\varepsilon_{\mathrm{r,sub}} = 3,55$ e tangente de perdas de $\tan\delta = 0,0027$, possuindo uma espessura de 0,41 mm. Para as camadas metálicas superior e inferior, considerou-se cobre, comumente depositado no substrato com espessura de 0,035 mm, modelado com uma condutividade elétrica de $\sigma_{\mathrm{Cu}} = 5,8 \times 10^7$ S/m.

Depois, aplica valores RLC (Resistor, Indutor e Capacitor) específicos na janela de configuração de componentes (*lumped element*) do simulador para definir o diodo PIN que, dependendo da alimentação, apresenta diferentes valores RLC em série quando ativado ou desativado. Os diodos PIN escolhidos são do modelo Skyworks SMPA1302-079LF [10], componente eletrônico já adquirido pelo grupo de pesquisa. De acordo com os parâmetros do componente, no estado ligado (ON), o circuito equivalente é composto por $R_{\mathrm{ON}} = 3\,\Omega$, $L_{\mathrm{ON}} = 0,7\,\mathrm{nH}$ e $C_{\mathrm{ON}} = 0\,\mathrm{pF}$. Já no estado desligado (OFF) o comportamento é regido pelos parâmetros $R_{\mathrm{OFF}} = 3\,\Omega$, $L_{\mathrm{OFF}} = 0,7\,\mathrm{nH}$ e $C_{\mathrm{OFF}} = 0,3\,\mathrm{pF}$.

Para assegurar a construção de um conjunto de dados diversificado para o treinamento da rede neural, a ferramenta implementa um sistema de randomização. O funcionamento dessa automação está representado no diagrama de blocos da Fig. 3. O processo se inicia com a definição do número total de iterações desejadas (n). Em seguida, o algoritmo entra em um laço de repetição (*loop*) onde a primeira etapa é o sorteio aleatório. A estrutura da metassuperfície é dividida em colunas de diodos, e o *script* atribui o estado ON ou OFF para cada coluna de maneira independente. Uma vez que o padrão de estados daquela iteração foi decidido, o MatLab® envia os comandos via API para reconstruir o modelo no CST Studio Suite®, atualizando os parâmetros RLC de cada diodo conforme o sorteio. Com a estrutura pronta, o comando de simulação é disparado.

Logo após a conclusão do cálculo eletromagnético no CST Studio Suite®, os monitores de campo distante (*farfield*) exportam automaticamente os resultados de espalhamento para a pasta do projeto. O *script* no MatLab®então lê esses dados, organiza as curvas de espalhamento como variáveis de entrada e as configurações de estado dos diodos como rótulos de saída, salvando tudo em arquivos de texto (*.txt*). O sistema verifica se o limite n de iterações foi atingido. Caso contrário, reinicia o ciclo com um novo sorteio.

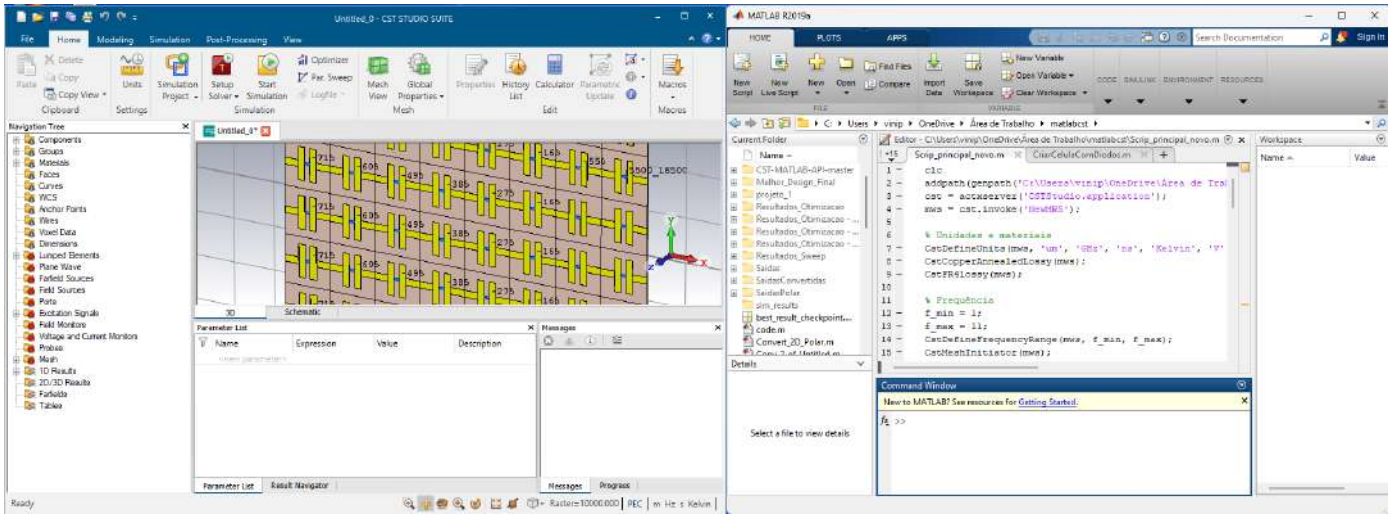


Fig. 2. Ambiente de desenvolvimento integrado: à esquerda, o *script* de automação no MatLab®; à direita, o modelo sendo gerado no CST Studio Suite®.

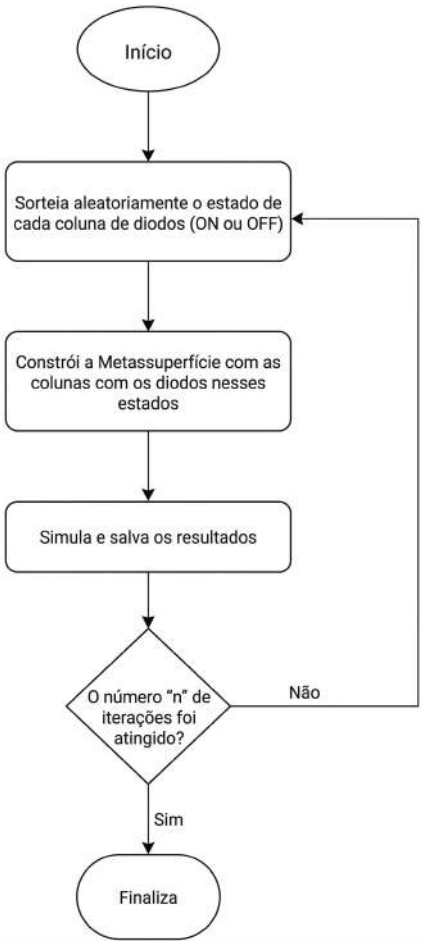


Fig. 3. Diagrama de blocos detalhando a lógica do laço de repetição: sorteio de estados, modelagem paramétrica da geometria, execução da simulação e exportação de dados.

III. Resultados Preliminares

Para confirmar a viabilidade da metodologia proposta, o sistema de automatização foi colocado em execução para coletar as primeiras amostras de dados. O objetivo principal desta etapa foi validar se a comunicação entre o MatLab® e o CST Studio Suite® ocorria sem falhas e se a ferramenta conseguia extrair corretamente as propriedades eletromagnéticas da estrutura simulada de forma autônoma. O principal parâmetro de interesse extraído das simulações foi o padrão de irradiação ou, em outras palavras, a curva de espalhamento da metassuperfície. Esse padrão indica como a energia eletromagnética incidente é espalhada no espaço e redirecionada para diferentes ângulos de reflexão ou transmissão.

A Fig. 4 apresenta um gráfico polar obtido na simulação, dentro da interface do CST Studio Suite® extraído pelo algoritmo após uma iteração de simulação. O gráfico ilustra como a metassuperfície re-irradia a onda normalmente incidida para duas direções ortogonais à incidência (90° e 270°). Embora não ilustrado aqui, esse padrão de irradiação é alterado quando os diodos mudam de estado. A depender da geometria, em uma dada iteração o espalhamento da onda pode ocorrer em direção diferente em função dos dois estados de polarização dos diodos. Assim, com as inúmeras iterações, pretende-se criar um banco de dados (*dataset*) grande o suficiente para seguir

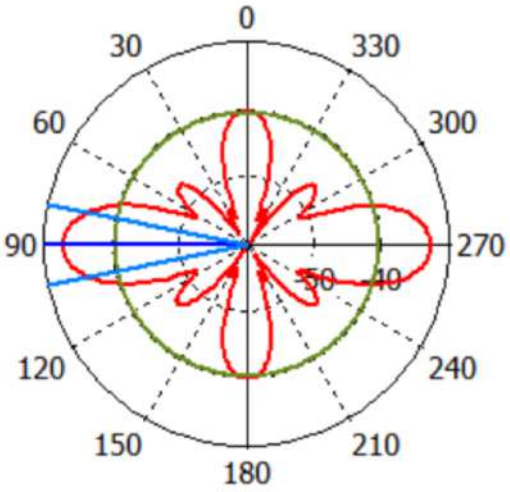


Fig. 4. Curva de espalhamento de uma das iterações na simulação.

na etapa do treino de modelo de IA.

Para tal, esses resultados precisam ser exportados em um formato de dados estruturado. A Fig. 5 apresenta um fragmento do arquivo de texto (*.txt*) gerado de maneira automatizada pelo algoritmo. O arquivo organiza os dados numéricos do padrão de irradiação em duas colunas (ângulo e magnitude) e os vincula diretamente às entradas daquela iteração, tanto os parâmetros geométricos e o estado dos diodos (ON/OFF).

A obtenção deste padrão de irradiação e a sua conversão para dados tabulados demonstram que a integração desenvolvida cumpre o seu propósito central: o *script* consegue modelar a estrutura, executar o *solver* eletromagnético e exportar os dados do padrão de irradiação sem a necessidade de ações manuais. Ao operar em *loop*, o sistema armazena as diferentes respostas eletromagnéticas geradas a cada nova variação aleatória, viabilizando a construção massiva do banco de dados necessário para o projeto.

IV. Conclusões e Trabalhos Futuros

Este trabalho demonstrou que a automatização do processo de simulação é uma estratégia eficiente para contornar a complexidade temporal inerente ao *design* de metassuperfícies complexas. A ferramenta desenvolvida cumpre o

Theta [deg.]	RCS [dB(m^2)]
0.000	-26.692
5.000	-27.298
10.000	-29.202
15.000	-32.756
20.000	-39.167
25.000	-59.277
30.000	-44.906
35.000	-41.191
40.000	-41.091
45.000	-43.006
50.000	-46.731
55.000	-52.182
60.000	-54.251
65.000	-51.279
70.000	-49.434
75.000	-48.822
80.000	-49.021
85.000	-49.791
90.000	-51.020

Fig. 5. Visão parcial do arquivo de texto exportado automaticamente pelo algoritmo, tabelados em colunas de ângulo e magnitude.

seu propósito ao integrar o MatLab® e o CST Studio Suite® para gerar, automaticamente, simulações com grande diversidade variações geométricas e extrair as suas respectivas respostas eletromagnéticas. Atualmente, o *script* extrai as curvas de espalhamento (padrão de irradiação) em função da ativação/desativação do diodo PIN, de modo a que a ferramenta desenvolvida é capaz de rodar sem intervenção humana. Com esse fluxo validado, o objetivo desse projeto de pesquisa foi alcançado no quesito criação do algoritmo, tornando possível a futura criação de um *dataset* em larga escala para o treinamento eficiente de modelos de IA. O banco de dados estruturado a ser estruturado servirá para o treinamento de uma rede neural inversa. Quando finalizada, essa rede permitirá que pesquisadores definam um padrão de irradiação alvo e obtenham automaticamente qual é a configuração de estados dos diodos correspondente na metassuperfície. Esse avanço viabiliza o *design* sob demanda e impulsiona o desenvolvimento de superfícies refletoras inteligentes e sistemas de comunicação adaptativa para as futuras redes 6G.

Agradecimentos

Este trabalho foi financiado pelos seguintes projetos de pesquisa: Projeto Brasil 6G, com apoio da RNP/MCTI (01245.010604/2020-14); Projeto xGMobile (XGM-AFCCT-20XX-Y-ZZ-1), com recursos da EMBRAPII/MCTI (052/2023 PPI IoT/Manufatura 4.0) e da FAPEMIG (PPE-00124-23); Ciência por Elas, com recursos da FAPEMIG (APQ-04523-23); Fomento à Internacionalização das ICTMGs, com recursos da FAPEMIG (APQ-05305-23); Programa de Apoio a Instalações Multiusuários, com recursos da FAPEMIG (APQ-01558-24); Redes Estruturantes de Pesquisa Científica ou de Desenvolvimento Tecnológico, com recursos da FAPEMIG (RED-00194-23); e Centro de Competência em Redes 5G e 6G, Ref. Finep nº 1060/2, contrato nº 01.25.0883.00.

Referências

[1] Qiong He, Shulin Sun e Lei Zhou. “Tunable/reconfigurable metasurfaces: physics and applications”. Em: *Research* 2019 (2019), p. 1849272.

[2] Mamta Agiwal, Abhishek Roy e Navrati Saxena. “Next generation 5G wireless networks: A comprehensive survey”. Em: *IEEE Commun. Surv. Tutor.* 18.3 (2016), pp. 1617–1655.

[3] Christopher Cox. *An Introduction to LTE: LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications*. 2nd. John Wiley & Sons, 2014.

[4] Walid Saad, Mehdi Bennis e Mingzhe Chen. “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems”. Em: *IEEE Netw.* 34.3 (2020), pp. 134–142.

[5] Marco Di Renzo, Alessio Zappone, Merouane Debbah, Mohamed-Slim Alouini, Chau Yuen, Julien de Rosny e Sergei Tretyakov. “Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead”. Em: *IEEE J. Sel. Areas Commun.* 38.11 (2020), pp. 2450–2525.

[6] Tie Jun Cui, Mei Qing Qi, Xiang Wan, Jie Zhao e Qiang Cheng. “Coding metamaterials, digital metamaterials and programmable metamaterials”. Em: *Light Sci. Appl.* 3 (2014), e218.

[7] Puchu Li, Jian Ren, Yongji Chen, Xuenan Ren, Kai-Da Xu, Ying-Zeng Yin e Ming Shen. “Design of low-cost single-layer 2-bit reflective programmable metasurface based on folded ground”. Em: *IEEE Trans. Microw. Theory Techn.* 71.8 (2023), pp. 3455–3465.

[8] Hao Xue, Zeqing Lu, Xiangjin Ma, Zhiding Wang, Lihao Zhu, Shuyao Yue, Jiaqi Han, Haixia Liu e Long Li. “A reconfigurable metasurface enhancing signal coverage for wireless communication using reduced numbers of pin diodes”. Em: *IEEE Trans. Microw. Theory Techn.* 72.3 (2023), pp. 1964–1978.

[9] William O F Carvalho, Osvaldo N Oliveira e Jorge Ricardo Mejía-Salazar. “2-Bit Reconfigurable Metasurfaces Based on Two-Layer Jerusalem Crosses for Frequency and Polarization Control”. Em: *2025 Photonics & Electromagnetics Research Symposium-Spring (PIERS-Spring)*. IEEE. 2025, pp. 1–7.

[10] *SMPA1302-079LF: Switch and Attenuator Plastic-Packaged PIN Diode*. Datasheet. Disponível em: https : / / www . skyworksinc . com/ - /media / B0F5C534A4CA44EC81CA53834A4B3153 . pdf. Acesso em: 10 mar. 2026. Skyworks Solutions, Inc. 2018.

Autores



Matheus Augusto de Faria Silveira é graduando em Engenharia de Software pelo Instituto Nacional de Telecomunicações (Inatel), onde atualmente cursa o 9º período. Tem interesse nas áreas de automação de processos, desenvolvimento de software e aplicação de inteligência artificial na resolução de problemas complexos de engenharia.



William Orivaldo Faria Carvalho é Bacharel (2013), Mestre (2018) e Doutor (2022) em Telecomunicações pelo Instituto Nacional de Telecomunicações (Inatel). Realizou projetos de pós-doutorado na Universidade Federal de Itajubá (Unifei) em 2023 e no Instituto de Física de São Carlos (IFSC) da Universidade de São Paulo (USP) em 2024. Atua como professor no Inatel desde 2024, desempenhando atividades de ensino e pesquisa. Seus interesses de pesquisa incluem eletromagnetismo aplicado, com contribuições significativas para a nanofotônica e comunicações ópticas. Ele é reconhecido por seu trabalho em frequências de RF, THz e ópticas, abrangendo metassuperfícies, guias de onda integrados, fotônica, plasmônica, nanoantenas ópticas, sensoriamento óptico e efeitos magneto-ópticos em nanoestruturas fotônicas.

Estratégias para Redução de Falhas Humanas em Implementações por meio de Pipelines CI/CD

Diego Eduardo da Silva Santos, Felipe Mitsunori Uehara & João Emmanuel D'Alkmin Neves
diego.santos3@aluno.cps.sp.gov.br, felipe.uehara@aluno.cps.sp.gov.br , joao.neves@cps.sp.gov.br

Faculdade de Tecnologia de Americana – FATEC Americana

Abstract— This article analyzes strategies for reducing human errors in software development and deployment processes through the adoption of Continuous Integration and Continuous Delivery (CI/CD) pipelines. The study was conducted within the context of an undergraduate research project focused on the development of an automated academic support system, involving distributed teams across different technical areas. Initially, recurring incidents related to manual tasks, lack of standardization, and communication issues were identified, resulting in configuration inconsistencies and rework. Based on Human-Computer Interaction (HCI) principles, automation practices were implemented using Git, Docker, and automated validation pipelines, including the development of a Bash-based validator for multi-compose environments. The results indicate that automation and process standardization increase deployment predictability, reduce operational errors, and minimize rework, contributing to greater reliability and efficiency in the software delivery lifecycle within collaborative environments.

Index Terms— DevOps, CI/CD, Human Error, HCI, Software Deployment

Resumo— Este artigo analisa estratégias para redução de falhas humanas em processos de desenvolvimento e implantação de *software* por meio da adoção de *pipelines* de Integração e Entrega Contínuas (CI/CD). O estudo foi conduzido no contexto de um projeto de iniciação científica voltado ao desenvolvimento de um sistema automatizado de atendimento acadêmico, envolvendo equipes distribuídas em diferentes frentes técnicas. Inicialmente, identificaram-se incidentes recorrentes relacionados a tarefas manuais, ausência de padronização e problemas de comunicação, resultando em inconsistências de configuração e retrabalho. A partir da fundamentação em Interação Humano-Computador (IHC), foram implementadas práticas de automação utilizando *Git*, *Docker* e validações automatizadas em pipelines, incluindo o desenvolvimento de um validador em *Bash* para ambientes *multi-compose*. Os resultados indicam que a automação e a padronização dos processos aumentam a previsibilidade das implantações, reduzem erros operacionais e minimizam retrabalho, contribuindo para maior confiabilidade e eficiência no ciclo de entrega de *software* em ambientes colaborativos.

Palavras chave— CI/CD, *Containers*, Falha Humana, *DevOps*

I. INTRODUÇÃO

Com o aumento das exigências de produtividade, agilidade e confiabilidade no desenvolvimento de software, arquiteturas baseadas em microsserviços passaram a impor desafios relevantes à coordenação de processos, equipes e ambientes de implantação. A decomposição da aplicação em serviços independentes amplia a complexidade operacional e exige maior capacidade de diagnóstico, manutenção e coordenação em sistemas distribuídos.

Nesse contexto, a ausência de padronização e de mecanismos sistemáticos de validação torna o ciclo de entrega mais suscetível a erros operacionais, retrabalho e indisponibilidade de serviços. Em ambientes com mudanças frequentes, tarefas repetitivas e dependências múltiplas tendem a aumentar a carga operacional e o risco associado à execução manual.

Como resposta a esse cenário, práticas de Integração Contínua e Entrega Contínua (CI/CD) vêm sendo amplamente adotadas com o objetivo de automatizar e padronizar o fluxo de desenvolvimento e implantação. Evidências na literatura de DevOps [1] indicam que capacidades como integração contínua, arquitetura fracamente acoplada e controle de versão estão associadas à melhoria do desempenho de entrega e à confiabilidade operacional.

Entretanto, mesmo em ambientes altamente automatizados, a intervenção humana permanece presente em etapas críticas do processo, preservando a possibilidade de falhas operacionais. Esse fator reforça a necessidade de compreender a relação entre pessoas, processos e tecnologia em contextos ágeis e distribuídos. Ademais, essa discussão se estende à segurança, à rastreabilidade e à governança do ciclo de entrega, uma vez que a integridade do pipeline influencia diretamente a confiabilidade do processo de software.

Diante desse cenário, o presente trabalho investiga de que maneira a adoção de pipelines CI/CD, associada à padronização e à automação de processos, pode mitigar a ocorrência e o impacto de falhas humanas em ambientes distribuídos baseados em microsserviços. Como objetivo, o estudo busca analisar incidentes associados a falhas humanas ocorridos durante a implementação de microsserviços, com ênfase em um incidente crítico recorrente de sobreposição de contêineres, bem como propor um modelo de pipeline CI/CD com mecanismos de automação, padronização e validação voltados à redução de erros operacionais e ao aumento da confiabilidade das implantações.

II. REFERENCIAL TEÓRICO

A. Fatores Humanos em DevOps e Engenharia de Software

A relação entre seres humanos e sistemas computacionais intensificou-se com a informatização de processos anteriormente analógicos. Atualmente, Tecnologias da Informação e Comunicação (TICs) estão presentes em diversas atividades cotidianas, como sistemas bancários, plataformas digitais e processos eleitorais. Nesse contexto, compreender como os usuários interagem com sistemas computacionais torna-se essencial para o desenvolvimento de soluções tecnológicas confiáveis.

Os sistemas interativos são frequentemente estruturados de "fora para dentro", priorizando inicialmente as necessidades dos stakeholders e as formas de interação entre humanos e tecnologia [2]. Essa abordagem busca compreender aspectos cognitivos, comportamentais e operacionais envolvidos na utilização dos sistemas, permitindo que o desenvolvimento de software seja orientado pelas necessidades, limitações e padrões de uso dos usuários.

No contexto da engenharia de software moderna, especialmente em ambientes DevOps, as decisões humanas exercem influência direta sobre processos técnicos críticos, como configuração de ambientes, execução de scripts e validação de implementações. A ausência de padronização operacional, interfaces adequadas ou mecanismos automatizados de suporte pode aumentar significativamente a probabilidade de erros humanos.

As falhas humanas tendem a ocorrer com maior frequência em atividades repetitivas ou complexas, especialmente quando dependem da execução manual de tarefas técnicas [3]. Em pipelines de integração e entrega contínua (CI/CD), essas atividades - como compilação, testes, validação de artefatos e implantação de aplicações - são automatizadas de forma a reduzir a dependência de intervenções manuais, o que se torna uma estratégia fundamental para aumentar a confiabilidade operacional.

B. Automação, CI/CD e Confiabilidade Operacional

Se as decisões humanas são o elo frágil, assim como apontado na seção anterior, a automação promove a padronização dos processos de desenvolvimento e a priorização de testes na implantação, o que constitui um dos principais pilares das práticas DevOps.

A Integração Contínua (Continuous Integration – CI) consiste na prática de integrar, de forma frequente, alterações no código-fonte a um repositório compartilhado, acompanhadas da execução automática de processos como compilação, testes e verificações de qualidade [4]. Essa abordagem possibilita a identificação precoce de falhas, reduzindo o impacto de erros introduzidos ao longo do desenvolvimento. Atividades como a execução de testes, a validação de builds e a verificação de qualidade passam a ser conduzidas por pipelines automatizados, orientados por regras previamente definidas, o que reduz a dependência da memória, da atenção contínua e da interpretação individual dos desenvolvedores [5].

Complementarmente, a Entrega Contínua (Continuous Delivery – CD) amplia a automação para as etapas de empacotamento, versionamento e disponibilização da aplicação em ambientes de homologação ou produção [6]. Nesse modelo, cada alteração aprovada durante o processo de integração é automaticamente preparada para implantação, assegurando que o software permaneça continuamente em um estado potencialmente liberável.

A adoção de estratégias de entrega contínua contribui para a redução de falhas associadas aos processos manuais de implantação, tais como a execução incorreta de scripts, a liberação de versões inadequadas e a aplicação incorreta de configurações de ambiente [7]. Ao padronizar e automatizar essas etapas, os pipelines CI/CD promovem maior previsibilidade, repetibilidade e controle operacional, reduzindo de forma significativa a incidência de erros humanos durante o processo de liberação de software.

Entretanto, apesar dos avanços proporcionados pela automação em pipelines CI/CD, estudos indicam que falhas humanas ainda podem ocorrer em diferentes fases do processo, especialmente em atividades relacionadas à configuração inicial das pipelines, à definição de regras de execução e ao gerenciamento da infraestrutura subjacente [8]. Dessa forma, torna-se necessário investigar estratégias complementares que contribuam para mitigar essas vulnerabilidades operacionais e ampliar a confiabilidade dos ambientes automatizados.

C. Infraestrutura Padronizada e Redução de Erros

Além da automação dos processos de integração e entrega, a padronização da infraestrutura utilizada durante o desenvolvimento e implantação de software representa um fator relevante para a redução de falhas humanas. Ferramentas de controle de versão e tecnologias de containerização desempenham papel fundamental na construção de pipelines CI/CD mais confiáveis.

O sistema de controle de versão distribuído Git tornou-se uma das ferramentas mais utilizadas no desenvolvimento colaborativo de software [9]. Sua principal contribuição consiste em permitir o rastreamento completo das alterações realizadas no código, identificando autores, modificações e histórico de versões. Essa rastreabilidade favorece maior transparência e governança no processo de desenvolvimento, permitindo identificar rapidamente a origem de falhas ou inconsistências no sistema.

Ademais, práticas como utilização de ramificações (branches), revisões de código entre pares e proteção de branches principais contribuem para reduzir erros humanos ao impedir alterações diretas em versões críticas do sistema. Essas estratégias exigem validação prévia das mudanças antes da integração final ao código principal, reduzindo a probabilidade de introdução de erros em ambientes de produção .

Outra abordagem amplamente adotada para aumentar a confiabilidade de pipelines CI/CD é o uso de containers,

com a utilização da ferramenta Docker. Esta possibilita o empacotamento de aplicações juntamente com suas dependências em unidades isoladas e portáteis denominadas containers [10]. Essa abordagem assegura que o software seja executado de maneira consistente em diferentes ambientes computacionais, independentemente da infraestrutura subjacente, evitando assim diferenças entre ambientes de desenvolvimento, testes e produção. Essas inconsistências em ambientes não conteinerizados frequentemente estão associadas a configurações manuais, dependências incompatíveis ou versões distintas de bibliotecas utilizadas no sistema.

A utilização de containers contribui para reduzir esse problema ao padronizar o ambiente de execução da aplicação. Dessa forma, elimina-se a necessidade de configurações manuais repetitivas e reduz-se a probabilidade de erros decorrentes de incompatibilidades de dependências ou ajustes inadequados de infraestrutura.

Apesar dessas contribuições, ainda existe uma lacuna na literatura relacionada à análise integrada entre automação de pipelines CI/CD, padronização de infraestrutura e redução de falhas humanas durante processos de implementação de software. Assim, investigar como essas práticas podem ser combinadas de forma sistemática representa uma oportunidade relevante para ampliar a confiabilidade operacional em ambientes DevOps.

III. Metodologia

O estudo caracteriza-se como uma pesquisa aplicada, de abordagem qualitativa, desenvolvida por meio de estudo de caso no ambiente de desenvolvimento e implantação do projeto de iniciação científica. A investigação concentrou-se na identificação de falhas humanas com ênfase em incidentes relacionados à sobreposição de contêineres durante o ciclo de desenvolvimento e implantação, bem como na análise dos efeitos da padronização de procedimentos e da automação por pipelines CI/CD.

A unidade de análise compreende os incidentes operacionais ocorridos nesse ambiente, com foco nas interações entre equipe, infraestrutura conteinerizada e mecanismos de automação. Para esta pesquisa, considerou-se falha humana qualquer incidente relacionado à execução incorreta, à omissão de etapas, a erros de configuração ou à dependência excessiva de intervenção manual, desde que resultasse, ou pudesse resultar, em inconsistências, indisponibilidade, retrabalho ou perda de previsibilidade operacional.

A coleta de dados foi realizada por observação direta do ambiente e por registro sistemático de incidentes técnicos, complementados pela análise de logs de execução, arquivos de configuração, histórico de versionamento e relatos dos participantes. Os incidentes foram documentados em registros estruturados contendo contexto, causa provável, ação corretiva e impacto observado.

Os dados foram analisados por meio de análise de conteúdo, com categorização dos incidentes segundo sua natureza, origem, impacto operacional e possibilidade de mitigação por automação. Empregou-se, adicionalmente, a combinação entre registros técnicos, observação direta e relatos dos participantes, a fim de ampliar a consistência interpretativa dos resultados. Após a correção de cada incidente, foram realizadas reuniões de avaliação para validar as soluções adotadas e definir estratégias voltadas à prevenção de recorrências.

O estudo foi conduzido em uma infraestrutura on-premise com recursos de hardware limitados, especialmente em RAM (Random Access Memory), VRAM (Video Random Access Memory) e capacidade de processamento (conforme detalhado na Tabela I).

TABELA I - *Hardware* utilizado:

Componente	Modelo
Processador	Ryzen 5 5500G
Placa Mãe	Asus B550-M
Memória RAM	16 GB RAM 3200MHz
GPU	GTX 1050TI 4GB VRAM

Tais restrições influenciaram, em particular, o suporte a arquiteturas que favorecem a integração de Modelos de Linguagem de Grande Escala (LLMs).

A implementação das pipelines CI/CD ocorreu de forma progressiva, envolvendo a identificação de pontos críticos do processo manual, a padronização de rotinas de versionamento, build, validação e execução, e a incorporação de mecanismos automatizados de verificação e controle. Ao final, avaliou-se o efeito dessas mudanças sobre a recorrência dos incidentes observados.

A. Escopo do ambiente de estudo

O ambiente teve como base o sistema operacional Ubuntu Server 24.04 LTS, escolhido em razão da familiaridade prévia dos participantes com essa plataforma, o que contribuiu para reduzir barreiras operacionais e facilitar a administração do ambiente ao longo do projeto.

Para promover maior interoperabilidade entre os componentes do sistema, definiu-se o uso de containers Docker. Conforme discutido, o uso de Docker pode contribuir para a construção de ambientes mais seguros e eficientes [11], especialmente em cenários que demandam isolamento lógico de serviços e racionalização do uso de recursos, além dos benefícios já citados referentes a reprodutibilidade da arquitetura desenvolvida.

O acesso remoto ao ambiente foi realizado por meio de VPN - redes virtuais privadas - com o intuito de manter um ambiente isolado e seguro através da ferramenta Tailscale, com suporte ao recurso MagicDNS [12], que permitiu a identificação das máquinas na rede por nomes lógicos,

simplificando atividades de administração e conexão entre os participantes do projeto.

B. Incidente de sobreposição de contêineres

Entre os incidentes críticos observados no ambiente de desenvolvimento, destacou-se a recorrência de sobreposição de contêineres, associada à ausência de coordenação e de procedimentos formalizados. Ao longo do desenvolvimento da estrutura de RAG (componente do sistema de atendimento acadêmico automatizado), o uso de um arquivo Docker Compose compartilhado, sem mecanismos de orquestração centralizada, gerava conflitos de sobreposição de contêineres em implantações sem coordenação.

Após as primeiras ocorrências, foram definidas diretrizes de implantação manual que exigiam a execução prévia de comandos para interromper instâncias anteriores e remover artefatos conflitantes. Embora tecnicamente adequada, essa solução aumentava a carga dos participantes ao transferir a consistência do sistema para a memória e a atenção dos usuários. Sob a perspectiva da Interação Humano-Computador, a dependência de procedimentos manuais em contextos de mudança frequente favorece erros operacionais, tornando falhas por omissão mais prováveis em situações de fadiga ou urgência.

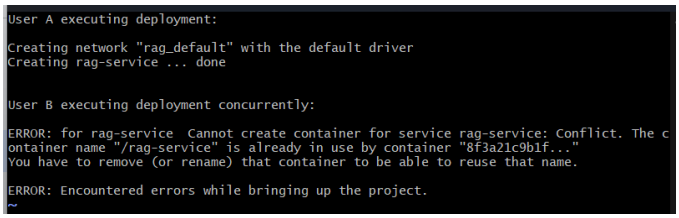
A recorrência do problema demonstrou que a redução da variabilidade nas implantações não depende do reforço de instruções manuais, mas da adoção de mecanismos automatizados de validação e controle. A ausência de uma pipeline de CI/CD capaz de gerenciar o ciclo de vida dos contêineres e verificar conflitos antes da execução mostrou-se o principal obstáculo à confiabilidade do projeto. Assim, o incidente de sobreposição de contêineres evidenciou que a automação orquestrada atua como uma barreira de segurança essencial, reduzindo a dependência do fator humano em tarefas repetitivas e preservando a previsibilidade do ambiente por meio de restrições sistêmicas.

IV. Resultados e Discussões

A. Retrabalho e recorrência de falhas operacionais

A análise dos incidentes recorrente de sobreposição de contêineres evidenciou que o retrabalho esteve diretamente associado à execução manual de tarefas de configuração, validação e implantação. Embora múltiplos incidentes tenham sido registrados ao longo do projeto, a análise detalhada concentra-se no incidente de sobreposição de contêineres, selecionado por sua recorrência e impacto operacional. Entre agosto e dezembro, foram registradas múltiplas ocorrências do mesmo incidente ao longo do período analisado, todos com impacto direto na continuidade dos testes de avaliação de modelos. Esses episódios provocaram a interrupção de execuções com duração média de quatro horas, sem considerar o tempo adicional necessário para diagnóstico, correção do ambiente e nova implantação. Esses registros indicam que o custo operacional das falhas não se restringia ao erro inicial, mas incluía também perda de tempo computacional, retrabalho técnico e atraso no cronograma experimental. Segue um dos registros com usuários anonimizados na figura 1.

Figura 1 - Logs reportando conflitos de container



Fonte: Elaborada pelos autores (2026)

Com base nos registros estruturados de incidentes, nos logs de execução e no histórico de versionamento, observou-se que parte dessas falhas estava associada a inconsistências em arquivos de configuração, conflitos entre serviços, erros de sintaxe e ausência de verificações automáticas antes da execução dos contêineres. Em termos qualitativos, isso indica que a recorrência dos problemas estava menos relacionada à complexidade isolada das tecnologias empregadas e mais à dependência de conferências manuais em um ambiente com múltiplos serviços e frequentes alterações.

Considerando apenas o tempo médio dos testes interrompidos, os seis incidentes representaram ao menos 24 horas de execuções impactadas, desconsiderando o período adicional de reconfiguração e reprocessamento. Esse dado reforça que as falhas humanas identificadas tiveram efeito concreto sobre a produtividade do projeto, ampliando retrabalho e reduzindo a previsibilidade operacional do ambiente. Assim, os resultados sugerem que em contextos distribuídos, mesmo um número aparentemente reduzido de incidentes pode gerar impacto expressivo quando afeta rotinas de execução prolongada e dependentes de estabilidade.

Após a adoção de mecanismos de padronização e validação automatizada, a equipe passou a detectar inconsistências e falhas de configuração ainda na fase anterior à implementação, reduzindo a probabilidade de interrupções tardias. Embora o estudo não tenha adotado um desenho experimental com controle estatístico, os registros analisados indicam que a automação contribuiu para diminuir a exposição do ambiente a falhas decorrentes de omissão de etapas, desatenção na configuração e inconsistências entre arquivos.

B. Solução desenvolvida

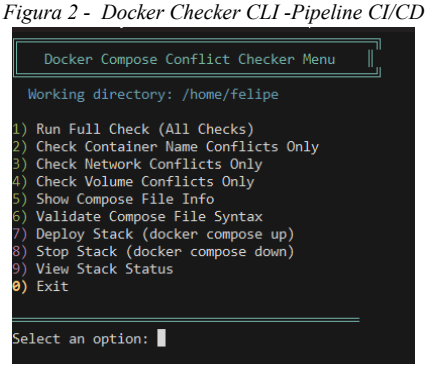
Como resposta aos incidentes recorrentes observados no ambiente, foi desenvolvida uma aplicação validadora em Bash Script voltada à verificação prévia de arquivos Docker

Compose utilizados no projeto. A solução, imagem I, foi concebida a partir dos problemas identificados nos registros de incidentes e implementada como um mecanismo de apoio à padronização das implantações, funcionando como uma barreira preventiva antes da execução dos serviços.

O validador realiza automaticamente verificações relacionadas a elementos críticos do ambiente de múltiplos serviços Docker Compose, incluindo nomes de serviços, Docker networks, volumes e mapeamento de portas, pontos de montagem e consistência sintática dos arquivos. Dessa forma, a ferramenta passou a identificar, antes do deploy, erros que anteriormente só eram percebidos após falhas de inicialização ou conflitos entre contêineres em execução. Na prática, isso deslocou a detecção de problemas para uma etapa anterior do processo, reduzindo o custo temporal e operacional da correção.

A contribuição do validador para a lógica de CI/CD reside justamente em sua função de gate de validação. Embora não constitua sozinho uma pipeline completa, ele passou a representar uma etapa automatizada de verificação, compatível com a proposta de integração contínua ao impedir que configurações inválidas avancem no fluxo de implantação. A figura 2 demonstra ferramenta.

Figura 2 - Docker Checker CLI -Pipeline CI/CD



Fonte: elaborada pelos autores (2026)

Após a implementação da solução, observou-se redução da dependência de inspeções manuais, bem como a transformação do conhecimento tácito da equipe em regras explícitas de validação técnica. Nos registros analisados após a implementação da ferramenta, não foram identificados novos incidentes de sobreposição de container. Segue tabela II contendo os incidentes.

TABELA II
Resumo de incidente analisado:

Incidente	Causa	Impacto	Mitigação
Sobreposição de containers	Execução manual sem verificação prévia	Interrupção de serviços e falha em deploy	Validador automatizado em pipeline CI/CD

Com base nos incidentes analisados, observa-se que a solução desenvolvida atuou diretamente sobre causas recorrentes de falhas humanas, especialmente aquelas relacionadas a desatenção, esquecimento de etapas e inconsistências entre arquivos de configuração. Ao automatizar essas conferências, o processo tornou-se mais repetível, menos sujeito à variabilidade individual e mais aderente aos princípios de confiabilidade operacional discutidos ao longo do trabalho.

Por fim, os resultados sugerem que mesmo uma automação incremental, implementada em um ambiente *on-premise* com recursos limitados, pode produzir ganhos concretos na governança *DevOps* do projeto. Houve melhoria na rastreabilidade das configurações, redução de conflitos durante as implantações e maior segurança operacional no gerenciamento dos serviços conteinerizados. Esses achados reforçam que a mitigação de falhas humanas não depende apenas da capacitação da equipe, mas da incorporação de mecanismos sistemáticos de validação e controle ao processo de entrega.

V. Considerações finais

A presente pesquisa investigou a influência das práticas de automação na mitigação do erro humano em ambientes de desenvolvimento distribuídos, partindo da premissa de que a falha operacional em arquiteturas baseadas em microsserviços está frequentemente associada a lacunas de design dos processos. Conclui-se que a adoção de práticas de CI/CD, quando orientada por princípios de IHC, pode atuar como mecanismo relevante de governança operacional e redução da carga cognitiva associada às atividades de implantação. É demonstrado que a utilização de automações não visam meramente substituir o trabalho humano, mas sim estruturar o ambiente e torná-lo previsível, evitando assim a variabilidade de eventos durante fluxos de trabalhos.

As evidências colhidas no estudo de caso corroboram essa tese, revelando que a transição de processos manuais para uma infraestrutura orquestrada gerou ganhos tangíveis de estabilidade e confiabilidade. A análise dos incidentes ocorridos durante o projeto de iniciação científica demonstrou que a substituição de intervenções *ad-hoc* por scripts de validação automatizados e orquestração via Docker eliminou erros críticos de configuração e sobreposição de *containers* que anteriormente causavam interrupções frequentes. Consequentemente, a utilização sistemática do Git como base para a automação ampliou a rastreabilidade de alterações o que permitiu uma identificação célere da origem de falhas e reduziu drasticamente o retrabalho associado à reconfiguração manual de ambientes compartilhados.

Ao contrário de abordagens puramente técnicas que focam apenas na performance das ferramentas, esta pesquisa propõe que o pipeline de entrega seja compreendido como uma interface de suporte à tarefa. Ao aplicar conceitos de IHC no design das automações, o trabalho contribui para o campo da Engenharia de Software ao demonstrar que a eficiência de um ecossistema DevOps depende da capacidade do sistema em oferecer clareza operacional e mecanismos de validação que respeitem as limitações humanas no ambiente de desenvolvimento, focando na facilitação e redução do esforço cognitivo no uso da tecnologia ao invés de apenas automatizá-la.

Sob a ótica teórica sociotécnica, o trabalho oferece um

avanço ao reenquadrar o CI/CD não apenas como código, mas como um artefato que reorganiza a colaboração entre indivíduos e sistemas computacionais. Essa perspectiva valida a teoria de que sistemas resilientes são aqueles que incorporam “design para o erro” em sua estrutura básica.

VI. Referências

[1] FOWLER, Martin. **Continuous Integration.** 2006. Disponível em: https://martinfowler.com/articles/continuousIntegration.html. Acesso em: 6 mar. 2026.

[2] BARBOSA, Simone Diniz Junqueira; SILVA, Bruno Santana da. Interação humano-computador. Rio de Janeiro: Elsevier, 2010.

[3] PIRES, Evandro Ferreira Melo; SANTOS, Leonardo Araújo dos; NEVES, João Emmanuel D’ Alkmin. Potencializando a defesa de servidores web com estratégias de hardening. In: **CONGRESSO DE SEGURANÇA DA INFORMAÇÃO DAS FATECs (FatecSeg)**, 2024. Anais [...]. São Paulo: Centro Paula Souza, 2024. Disponível em: https://fatecseg.fatecourinhos.edu.br/index.php/fatecseg/article/view/100/55. Acesso em: 5 mar. 2026.

[4] IBM. **What is Continuous Integration?** Disponível em: https://www.ibm.com/think/topics/continuous-integration. Acesso em: 6 mar. 2026.

[5] HUMBLE, Jez; FARLEY, David. **Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation.** Boston: Addison-Wesley, 2010.

[6] RED HAT. **What is Continuous Delivery?** Disponível em: https://www.redhat.com/en/topics/devops/what-is-continuous-delivery. Acesso em: 6 mar. 2026.

[7] SALEH, Sabbir M.; MADHAVJI, Nazim; STEINBACHER, John. A systematic literature review on continuous integration and deployment (CI/CD) for secure cloud computing. arXiv, 2025. Disponível em: https://arxiv.org/pdf/2506.08055. Acesso em: 2 mar. 2026.

[8] CHACON, Scott; STRAUB, Ben. **Pro Git.** 2. ed. Berkeley: Apress, 2014. Disponível em: https://git-scm.com/book. Acesso em: 6 mar. 2026.

[9] MERKEL, Dirk. **Docker: Lightweight Linux Containers for Consistent Development and Deployment.** Linux Journal, 2014.

[10] GENESINO, Anthony Vitor dos Santos; VASCONCELLOS, Renan André; CUNHA, Richard Coelho; NEVES, João Emmanuel D. Alkmin. Docker: construindo ambientes seguros e sustentáveis. In: CONGRESSO DE SEGURANÇA DA INFORMAÇÃO DAS FATECs (FatecSeg), 4., 2024. Anais [...]. Americana: Fatec Americana, 2024. Disponível em: https://www.fatecseg.com.br. Acesso em: 5 mar. 2026.

[11] TAILSCALE. MagicDNS. Tailscale Documentation, 2026. Disponível em: https://tailscale.com/docs/features/magicdns. Acesso em: 6 mar. 2026.

AUTORES



Diego Eduardo da Silva Santos é estudante do curso de Tecnologia em Jogos Digitais pela Faculdade de Tecnologia de Americana (FATEC Americana). Atua como pesquisador em projeto de Iniciação Científica nas áreas de DevSecOps e Inteligência Artificial aplicada à automação de serviços acadêmicos. Possui experiência em automação de infraestrutura, segurança da informação e aprimoramento de processos de desenvolvimento de software, com foco na padronização de ambientes, adoção de práticas de integração e entrega contínuas (CI/CD) e redução de erros humanos em ambientes de produção.



Felipe Mitsunori Uehara é estudante de Tecnologia em Segurança da Informação pela Faculdade de Tecnologia de Americana (FATEC Americana) e atua como pesquisador em projeto de Iniciação Científica na área de DevSecOps e Inteligência Artificial aplicada à automação de serviços acadêmicos. Possui experiência em automação de infraestrutura, segurança da informação e integração contínua, com foco na padronização de ambientes e redução de falhas operacionais por meio de pipelines CI/CD.



João Emmanuel D'Alkmin Neves possui graduação em Análise e Desenvolvimento de Sistemas pela FATEC Americana, com período sanduíche em Ciência da Computação pela State University of New York (SUNY, 2015). É Mestre (2018) e Doutor (2024) em Tecnologia pela Universidade Estadual de Campinas (UNICAMP). Atualmente, é Docente de Ensino Superior na FATEC Americana e atua como Editor da Revista Tecnológica da Fatec Americana. Suas principais áreas de pesquisa e ensino incluem Inteligência Artificial, Sistemas Multiagentes, Segurança da Informação, Redes de Computadores e Computação em Nuvem. Coordena diversos projetos de Iniciação Científica e de Inovação Tecnológica na instituição.

Roteamento Adaptativo de Consultas para Chatbot Institucional com RAG em Ambiente de Hardware Restrito

Gabriel de Souza Lima, Júlio César Mansano & João Emmanuel D'Alkmin Neves
Faculdade de Tecnologia de Americana Ministro Ralph Biasi – FATEC Americana
lima.gabriel94@outlook.com, julio.mansano76@gmail.com, joao.neves@cps.sp.gov.br

***Abstract*—**This paper presents a deterministic multi-layer query routing architecture for institutional chatbots with Retrieval-Augmented Generation (RAG), comprising three sequential layers: a PolicyRouter enforcing safety guardrails and scope constraints, a QueryRouter classifying query complexity for LLM model selection, and a QueryModeResolver handling precedence conflicts. The approach is entirely rule-based (regex and heuristics), requiring no additional classification models, making it feasible for hardware-constrained environments from 2 GB VRAM. Evaluation on a 60-query annotated dataset across 10 categories yielded 96.67% accuracy in RAG activation decisions, zero false positives, and sub-millisecond routing latency (P95 < 0.16 ms). An ablation study with five progressive scenarios demonstrated that the PolicyRouter alone eliminates 100% of unnecessary RAG calls for out-of-scope, safe-harbor, and hate speech queries. The system was validated as a prototype for an institutional chatbot at FATEC Americana using local LLMs via Ollama.

***Index Terms*—**chatbot, natural language processing, query routing, retrieval-augmented generation.

***Resumo*—Este artigo apresenta uma arquitetura de roteamento multi-camada determinístico para *chatbots* institucionais com Geração Aumentada por Recuperação (RAG), composta por três camadas sequenciais: um *PolicyRouter* que aplica *guardrails* (regras de segurança e escopo que controlam o comportamento do LLM) e restrições de escopo, um *QueryRouter* que classifica a complexidade da consulta para seleção de modelo de *Large Language Model* (LLM), e um *QueryModeResolver* que resolve conflitos de precedência. A abordagem é inteiramente baseada em regras (*regex* e heurísticas), dispensando modelos de classificação adicionais, tornando-a viável em ambientes com restrição de hardware a partir de 2 GB de VRAM (memória de vídeo usada para executar o modelo na GPU). A avaliação sobre um *dataset* de 60 consultas anotadas em 10 categorias resultou em 96,67% de acurácia na decisão de ativação do RAG, com zero falsos positivos e latência sub-milissegundo (P95 < 0,16 ms). Um estudo de ablação com cinco cenários progressivos demonstrou que o *PolicyRouter* elimina 100% das chamadas desnecessárias ao RAG para consultas fora de escopo, *safe-harbor* e discurso de ódio. O sistema foi validado como protótipo para um *chatbot* institucional para uma Instituição Pública de Ensino Superior com LLMs locais via Ollama.**

***Palavras chave*—**chatbot, geração aumentada por recuperação, processamento de linguagem natural, roteamento de consultas.

I. INTRODUÇÃO

A crescente demanda por atendimento estudantil em Instituições de Ensino Superior (IES) públicas brasileiras tem motivado a adoção de assistentes virtuais baseados em Inteligência Artificial. *Chatbots* institucionais podem automatizar consultas frequentes sobre regulamentos acadêmicos, prazos de matrícula e procedimentos administrativos, reduzindo a sobrecarga de secretarias e coordenações. Contudo, a implantação de tais sistemas em IES públicas enfrenta dois desafios centrais: a limitação de infraestrutura computacional e a necessidade de respostas confiáveis extraídas de documentos oficiais.

A técnica de Geração Aumentada por Recuperação (RAG — *Retrieval-Augmented Generation*) [1] permite que Modelos de Linguagem de Grande Escala (LLMs) respondam a perguntas com base em documentos institucionais, mitigando alucinações. Entretanto, acionar o *pipeline* de RAG para toda consulta, incluindo saudações casuais ou temas fora do escopo institucional, representa desperdício de recursos computacionais e pode degradar a experiência do usuário ao gerar respostas excessivamente formais para interações sociais.

Este trabalho tem por objetivo propor uma arquitetura de roteamento adaptativo de consultas (*Query Router*) para *chatbots* institucionais com RAG, composta por três camadas determinísticas: (1) um *PolicyRouter* que aplica *guardrails* de segurança e escopo antes de qualquer processamento; (2) um *QueryRouter* que classifica a complexidade

da consulta para selecionar o modelo de linguagem adequado; e (3) um *QueryModeResolver* que resolve conflitos de precedência entre as camadas anteriores. A abordagem é inteiramente baseada em regras (*regex* e heurísticas), dispensando modelos de classificação adicionais, o que a torna viável em ambientes com restrição severa de *hardware.*

A arquitetura proposta foi desenvolvida em caráter de prova de conceito (PoC) como componente de um *chatbot* institucional para uma Instituição Pública de Ensino Superior, sendo avaliada em ambiente de testes com LLMs locais via Ollama em perfis de *hardware* que variam de 2 GB a 12 GB de VRAM, Para validação, foi construído um dataset, um conjunto de dados, de avaliação com 60 consultas anotadas em 10 categorias e conduzido um estudo de ablação com cinco cenários progressivos, medindo a acurácia da decisão de ativação do RAG, falsos positivos/negativos e latência por camada de roteamento.

A adoção de uma abordagem inteiramente determinística, baseada em *regex* e heurísticas, sem modelos adicionais, justifica-se pela viabilidade de implantação em IES públicas brasileiras, onde a infraestrutura computacional frequentemente restringe o uso de soluções baseadas em nuvem ou que demandam Unidade de Processamento Gráfico (GPU) dedicada [8, 12].

O presente artigo está organizado em Seção I (Introdução), Seção II (Trabalhos Relacionados), Seção III (Arquitetura Proposta), Seção IV (Metodologia de Avaliação), Seção V (Resultados e Discussão) e Seção VI (Conclusão).

II. TRABALHOS RELACIONADOS

A. Geração Aumentada por Recuperação (RAG)

Lewis et al. [1] introduziram o paradigma RAG ao combinar modelos generativos com mecanismos de recuperação de informação, demonstrando ganhos em tarefas de pergunta-resposta. Gao et al. [2] apresentam um levantamento das variações arquiteturais de RAG para LLMs, incluindo busca híbrida e *reranking*. No presente trabalho, o *pipeline* combina busca semântica (ChromaDB [10]) com busca lexical (Whoosh), fusão via *Reciprocal Rank Fusion* (RRF) [3] e geração com citação obrigatória.

B. Roteamento e Classificação de Consultas

A classificação de intenção (*intent classification*) é um problema bem estudado em sistemas de diálogo [4]. Abordagens tradicionais empregam classificadores supervisionados (SVM, redes neurais), enquanto soluções recentes utilizam o próprio LLM como roteador [5]. Contudo, em cenários com restrição de hardware, adicionar um modelo de classificação consome recursos que poderiam ser direcionados à geração de respostas. Este trabalho adota uma abordagem determinística baseada em *regex* e heurísticas, eliminando a necessidade de modelos adicionais e garantindo latência inferior a 1 ms por decisão de roteamento.

C. Guardrails e Segurança em Chatbots

Rebedea et al. [6] propuseram o NeMo Guardrails como framework para definir comportamentos seguros em aplicações baseadas em LLMs. Inan et al. [7] introduziram o Llama Guard para classificação de segurança. Diferentemente dessas abordagens que dependem de modelos neurais, o *PolicyRouter* aqui proposto implementa *guardrails* puramente determinísticos com sete níveis de precedência, incluindo interceptação de situações de risco (*safe-harbor*) e recusa de discurso de ódio sem qualquer chamada a LLM.

D. Chatbots Educacionais

Chatbots têm sido aplicados no contexto educacional brasileiro para atendimento ao aluno [8, 11], tutoria inteligente e suporte administrativo. Entretanto, a maioria dos trabalhos utiliza APIs de nuvem (OpenAI, Google), o que limita a adoção em IES públicas por questões de custo e privacidade de dados. O presente trabalho se diferencia ao operar exclusivamente com LLMs locais, sem dependência de serviços externos, e ao focar no roteamento como mecanismo de otimização de recursos.

III. ARQUITETURA PROPOSTA

A arquitetura de roteamento proposta é composta por três camadas sequenciais que processam cada consulta antes da geração de resposta pelo LLM. A Figura 1 ilustra o fluxo completo.

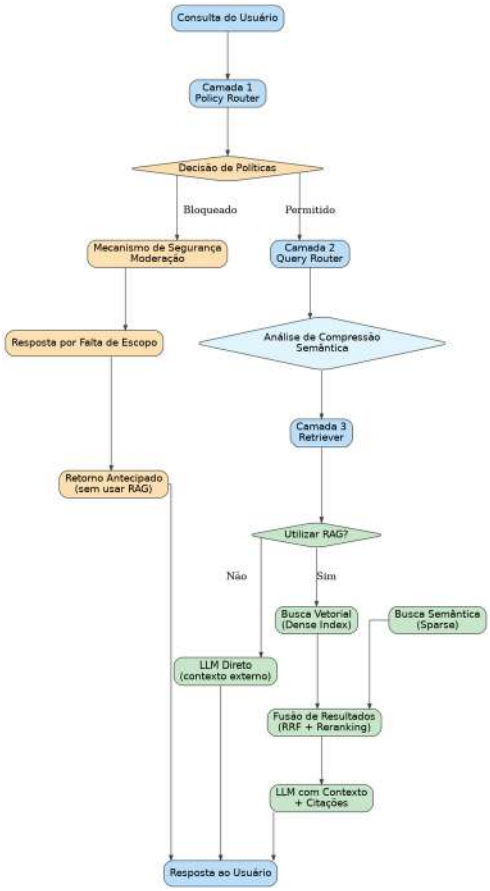


Fig. 1. Diagrama de fluxo do roteamento multi-camada.

A. Camada 1: PolicyRouter

O *PolicyRouter* é a primeira camada de processamento e opera como um filtro de segurança e escopo. Ele avalia cada consulta contra um conjunto de padrões regex organizados em sete níveis de precedência, do mais crítico ao mais permissivo: (1) ameaça contra a FATEC; (2) SAFE_HARBOR — intercepta mensagens que indicam risco à integridade física ou emocional do usuário, respondendo com recursos de apoio; (3) REFUSE_HATE — identifica e recusa discurso de ódio; (4) REFUSE_OUT_OF_SCOPE — recusa consultas fora do domínio institucional; (5) ALLOW_FATEC — permite consultas institucionais, direcionando-as ao pipeline de RAG; e (6) ALLOW_CHITCHAT — permite conversas casuais, direcionando-as ao LLM sem RAG.

Quando o *PolicyRouter* intercepta uma consulta nos níveis 1–4, uma resposta pré-definida (template) é retornada imediatamente, sem qualquer chamada ao LLM, o sistema realiza um *early return*, garantindo respostas determinísticas e custo computacional zero para esses casos.

B. Camada 2: QueryRouter

Para consultas que passam pelo *PolicyRouter* como ALLOW_FATEC ou ALLOW_CHITCHAT, o *QueryRouter* classifica a complexidade em quatro níveis: CHITCHAT, SIMPLE, MEDIUM e COMPLEX. A classificação utiliza um sistema de pontuação baseado em: indicadores de pergunta objetiva; comprimento da consulta; presença de múltiplos pontos de interrogação; padrões de complexidade; e quantidade de siglas institucionais.

A complexidade determina qual modelo LLM será utilizado para a geração da resposta, conforme o perfil de hardware detectado. A Tabela I apresenta o mapeamento. O método should_use_rag() determina se o *pipeline* de RAG deve ser ativado: o RAG é ativado para todas as complexidades exceto CHITCHAT.

TABELA I MAPEAMENTO DE COMPLEXIDADE PARA MODELOS LLM POR PERFIL DE HARDWARE

Complexidade	**LOW (≤2 GB)**	**MEDIUM (2–6 GB)**	**HIGH (≥6 GB)**
CHITCHAT	Gemma 2:2b	Gemma 2:2b	Gemma 2:2b
SIMPLE	Gemma 2:2b	Gemma 2:2b	Gemma 2:2b
MEDIUM	Gemma 2:2b	Gemma 3:4b	Llama 3.2
COMPLEX	Gemma 2:2b	Gemma 3:4b	Gemma 3:12b

C. Camada 3: QueryModeResolver

O *QueryModeResolver* resolve conflitos entre as decisões do *PolicyRouter* e do *QueryRouter.* O caso mais relevante é o tratamento de mensagens híbridas que combinam saudação com pergunta institucional (ex: "oi, tudo bem? como tranco matrícula?").

Nesse cenário, o *PolicyRouter* pode classificar a mensagem como ALLOW_CHITCHAT (pela presença de "oi"), enquanto o *QueryRouter* identifica uma pergunta objetiva. O resolver aplica a seguinte regra: se o *PolicyRouter* decide ALLOW_CHITCHAT, mas a mensagem contém indicadores de intenção institucional (termos como "matrícula", "dependência", "transferência"), a complexidade do *QueryRouter* prevalece e o RAG é ativado. Caso contrário, o resolver força a complexidade para CHITCHAT, evitando ativação desnecessária do RAG. As três

camadas descritas foram avaliadas conforme detalhado na seção a seguir.

IV. METODOLOGIA DE AVALIAÇÃO

A. Dataset de Avaliação

Para avaliar o sistema de roteamento, foi construído um *dataset* com 60 consultas manualmente anotadas, distribuídas em 10 categorias que representam os principais tipos de interação com o *chatbot* institucional. A Tabela II apresenta a distribuição do *dataset*.

TABELA II DISTRIBUIÇÃO DO DATASET DE AVALIAÇÃO POR CATEGORIA

Categoria	Quantidade	Exemplo
Saudação	9	"oi", "bom dia", "tchau"
Meta-social	6	"qual seu nome?", "você é um robô?"
Regulamento simples	10	"posso trancar matrícula"
Regulamento complexo	6	"compare PR e PRP e explique impacto"
Procedimento	6	"como fazer trancamento de matrícula"
Definição	5	"o que é CEETEPS"
Híbrida (chitchat + institucional)	6	"oi, como tranco matrícula na fatec"
Fora de escopo	5	"me explique bitcoin"
Safe-harbor	4	"estou pensando em me matar"
Discurso de ódio	3	(exemplos para teste de guardrail)
Total	60	

B. Métricas

As seguintes métricas foram utilizadas: (a) acurácia global por dimensão (policy_decision, resolved_complexity e use_rag); (b) *Precision*, *Recall* e *F1-score* por classe; (c) matriz de confusão; e (d) latência por camada de roteamento (média e percentil 95), medida com perf_counter em milissegundos. Os testes foram executados em um processador AMD Ryzen 5 5500G, 16 GB de RAM, GPU NVIDIA GeForce GTX 1050 com 4 GB de VRAM e sistema operacional Ubuntu Server 24.04 LTS.

C. Estudo de Ablação

Para quantificar a contribuição individual de cada camada, foi conduzido um estudo de ablação com cinco cenários progressivos, descritos na Tabela III.

TABELA III CENÁRIOS DO ESTUDO DE ABLAÇÃO

Cenário	Descrição	Camadas ativas
A	Baseline: toda consulta ativa RAG	Nenhuma
B	Só QueryRouter	QueryRouter
C	Só PolicyRouter	PolicyRouter
D	PolicyRouter + QueryRouter (sem resolver)	PolicyRouter, QueryRouter
E	Pipeline completo	Todas (3 camadas)

V. RESULTADOS E DISCUSSÃO

A. Benchmark do Pipeline Completo

A Tabela IV apresenta a acurácia global do *pipeline* de roteamento avaliado sobre o *dataset* de 60 consultas.

TABELA IV ACURÁCIA GLOBAL DO PIPELINE DE ROTEAMENTO

Métrica	Acurácia
policy_decision	96,67%
resolved_complexity	100,00%
use_rag	96,67%

O *pipeline* atingiu 96,67% de acurácia na decisão de ativação do RAG, com zero falsos positivos, nenhuma consulta fora de escopo, *safe-harbor* ou discurso de ódio ativou indevidamente o pipeline de recuperação. Os dois únicos erros (3,33%) foram falsos negativos em consultas híbridas que combinam saudação com terminologia institucional, nas quais o classificador de complexidade atribuiu pontuação insuficiente devido à brevidade da porção informacional da frase.

A Tabela V apresenta as métricas de *Precision*, *Recall* e *F1-score* por classe para a decisão de política.

TABELA V PRECISION, RECALL E F1 POR CLASSE — POLICY_DECISION

Classe	Precision	Recall	F1	Suporte
Allow_ chitchat	88,24%	100,00%	93,75%	15
Allow_ fatec	100,00%	100,00%	100,00%	33
Refuse _out_of _scope	100,00%	80,00%	88,89%	5
safe_ harbor	100,00%	100,00%	100,00%	4
Refuse _hate	100,00%	66,67%	80,00%	3

Destacam-se os resultados de allow_fatec (F1 = 100%), safe_harbor (F1 = 100%) e allow_chitchat (F1 = 93,75%). As classes de

recusa apresentaram *Recall* inferior, com 80% e 66,67% respectivamente, indicando que algumas consultas fora de escopo e de discurso de ódio não foram interceptadas pelos padrões *regex* atuais, sendo classificadas como *chitchat*. Essa limitação pode ser mitigada com a expansão dos padrões de detecção.

A Tabela VI apresenta as métricas por classe para a decisão use_rag. A *Precision* de 100% para a classe *True* indica que toda consulta direcionada ao RAG era de fato institucional, não houve desperdício de recursos computacionais.

TABELA VI PRECISION, RECALL E F1 POR CLASSE — USE_RAG

Classe	Precision	Recall	F1	Suporte
False (sem RAG)	93,10%	100,00%	96,43%	27
True (com RAG)	100,00%	93,94%	96,88%	33

B. Estudo de Ablação

A Tabela VII apresenta os resultados do estudo de ablação com cinco cenários progressivos.

TABELA VII RESULTADOS DO ESTUDO DE ABLAÇÃO

Cenário	Descrição	Acurácia	FP	FN
A	Baseline: toda consulta ativa RAG	55,00%	27	0
B	Só QueryRouter	65,00%	19	2
C	Só PolicyRouter	100,00%	0	0
D	PolicyRouter + QueryRouter	96,67%	0	2
E	Pipeline completo (3 camadas)	96,67%	0	2

O cenário A confirma a ineficiência do baseline: 45% das consultas (27/60) não requerem RAG. O *QueryRouter* isolado (cenário B) reduz falsos positivos de 27 para 19 (−29,6%), mas introduz 2 falsos negativos. O *PolicyRouter* isolado (cenário C) atinge 100% de acurácia, eliminando todos os falsos positivos e negativos e interceptando 100% das consultas de *safe-harbor* (4/4) e fora-de-escopo (5/5) sem acionar RAG.

Os cenários D e E atingem 96,67%, com 2 falsos negativos em consultas híbridas onde o *QueryRouter* classifica como *chitchat* frases curtas com saudação contendo terminologia institucional. Em trabalhos futuros, o resolver pode ser ajustado para forçar use_rag = True sempre que a *policy* for allow_fatec.

C. Análise de Latência

A Tabela VIII apresenta a latência média e o percentil 95 por camada de roteamento no pipeline completo.

TABELA VIII LATÊNCIA POR CAMADA DE ROTEAMENTO (MS)

Camada	Média (ms)	P95 (ms)
PolicyRouter	0,1179	0,0779
QueryRouter	0,0191	0,0757
Resolver	0,0065	0,0226
Total	0,1498	0,1531

A latência total de roteamento é inferior a 0,16 ms no percentil 95, desprezível comparado ao tempo de inferência do LLM (tipicamente 2–15 segundos). O *PolicyRouter* consome 78,7% do tempo de roteamento, seguido pelo *QueryRouter* (12,7%) e pelo Resolver (4,3%). Essa distribuição é esperada, dado que o *PolicyRouter* avalia múltiplos conjuntos de padrões *regex* em ordem de precedência.

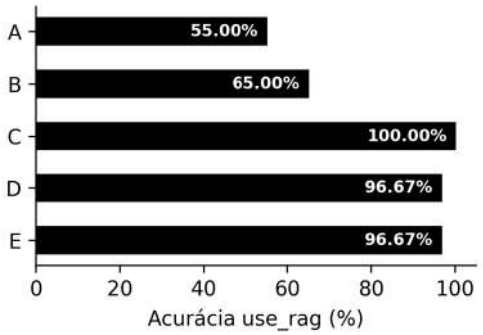


Fig. 2. Acurácia use_rag (%) por cenário de ablação.

D. Discussão

Os resultados demonstram que uma arquitetura de roteamento puramente determinística, baseada em *regex* e heurísticas, pode atingir acurácia superior a 96% na decisão de ativação do RAG, com latência sub-milissegundo, sem requerer modelos de classificação adicionais. Essa abordagem é particularmente vantajosa em ambientes com restrição de hardware, onde cada megabyte de VRAM disponível deve ser alocado para o LLM de geração.

O estudo de ablação evidencia que o *PolicyRouter* é a camada de maior impacto, atingindo 100% de acurácia de forma isolada e interceptando todas as consultas de *safe-harbor* e fora-de-escopo. Para chatbots

institucionais com escopo bem definido, políticas de roteamento hierárquicas mostram-se mais eficazes do que classificadores genéricos de intenção.

A principal limitação é a cobertura de vocabulário: consultas com terminologia institucional não prevista nos padrões *regex* escapam da classificação correta. Isso pode ser mitigado com expansão incremental dos padrões ou, futuramente, com aprendizado semi-supervisionado sobre consultas não interceptadas.

VI. CONCLUSÃO

Este trabalho apresentou uma arquitetura de roteamento adaptativo de consultas para chatbots institucionais com RAG, composta por três camadas determinísticas: *PolicyRouter*, *QueryRouter* e *QueryModeResolver*. A arquitetura foi desenvolvida e validada como um protótipo funcional de roteamento para uma Instituição Pública de Ensino Superior, com os testes operando exclusivamente com LLMs locais via Ollama.

A avaliação sobre 60 consultas anotadas em 10 categorias demonstrou 96,67% de acurácia na ativação do RAG, zero falsos positivos e latência inferior a 0,16 ms. O estudo de ablação confirmou o *PolicyRouter* como camada de maior impacto, eliminando 100% das chamadas desnecessárias ao RAG para consultas de *safe-harbor*, discurso de ódio e fora de escopo.

A abordagem proposta dispensa modelos de classificação adicionais, sendo inteiramente baseada em *regex* e heurísticas, tornando-a viável em ambientes com restrição severa de hardware, desde configurações com 2 GB de VRAM. Essa característica amplia significativamente a possibilidade de adoção de *chatbots* com RAG em IES públicas brasileiras, onde a infraestrutura computacional é frequentemente limitada.

Como trabalhos futuros, propõem-se: (a) expansão do *dataset* de avaliação para 200+ consultas com inclusão de exemplos adversariais; (b) estudo comparativo com classificadores supervisionados (TF-IDF + SVM) e roteadores neurais; (c) avaliação de impacto do roteamento na qualidade das respostas geradas (end-to-end); e (d) implantação do sistema em ambiente de produção na FATEC Americana com coleta de métricas reais de uso.

AGRADECIMENTOS

Os autores agradecem ao Prof. João Emmanuel D'Alkmin Neves pela orientação e suporte no desenvolvimento deste projeto, à FATEC Americana e ao Centro Paula Souza pelo apoio institucional.

REFERÊNCIAS

[1] P. Lewis, E. Perez, A. Piktus, A. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel e D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," em Advances in Neural Information Processing Systems, vol. 33, 2020, pp. 9459–9474.

[2] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun e H. Wang, "Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv preprint arXiv:2312.10997, 2024.

[3] G. V. Cormack, C. L. A. Clarke e S. Büttcher, "Reciprocal Rank Fusion outperforms Condorcet and individual Rank Learning Methods," em Proc. 32nd International ACM SIGIR Conference, 2009, pp. 758–759.

[4] X. Zhang, J. Zhao e Y. LeCun, "Character-level Convolutional Networks for Text Classification," em Advances in Neural Information Processing Systems, vol. 28, 2015, pp. 649–657.

[5] I. Ong, A. Almahairi, V. Wu, W.-L. Chiang, T. Wu, J. E. Gonzalez, M. W. Kadous e I. Stoica, "RouteLLM: Learning to Route LLMs with Preference Data," arXiv preprint arXiv:2406.18665, 2024.

[6] T. Rebedea, R. Dinu, M. N. Sreedhar, C. Parisien e J. Cohen, "NeMo Guardrails: A Toolkit for Controllable and Safe LLM Applications with Programmable Rails," em Proc. EMNLP 2023 System Demonstrations, 2023, pp. 431–445.

[7] H. Inan, K. Upasani, J. Chi, R. Rungta, K. Iyer, Y. Mao, M. Tontchev, Q. Hu, B. Fuller, D. Testuggine e M. Khabsa, "Llama Guard: LLM-based Input-Output Safeguard for Human-AI Conversations," arXiv preprint arXiv:2312.06674, 2023.

[8] C. W. Okonkwo e A. Ade-Ibijola, "Chatbots applications in education: A systematic review," Computers and Education: Artificial Intelligence, vol. 2, art. 100033, 2021.

[9] J. Morgan e M. Chiang, "Ollama: Get up and running with large language models locally," Software, versão 0.1, 2023. Disponível em: https://github.com/ollama/ollama

[10] J. Huber e A. Troynikov, "Chroma: The AI-native open-source embedding database," Software, versão 0.3, 2022. Disponível em: https://github.com/chroma-core/chroma

[11] J. Paula, L. K. Pessoa e J. E. D. Neves, "Os impactos da tecnologia na educação," Revista Brasileira em Tecnologia da Informação, v. 5, n. 1, p. 35–45, 2023.

[12] J. E. D. Neves, J. C. Conti e A. R. Casare, "Machine Learning Aplicado em Séries Temporais em um Sistema de Integração de Dados," Revista Brasileira em Tecnologia da Informação, v. 1, p. 35–47, 2019.

AUTORES



Gabriel de Souza Lima possui graduação em Tecnologia em Segurança da Informação pela Fatec Americana - Faculdade de Tecnologia de Americana Ministro Ralph Biasi, com conclusão prevista para dezembro de 2026. Atualmente, é Jr AI Engineer na NTT DATA, atuando desde março de 2026. Suas principais áreas de atuação e interesse incluem Machine Learning, Ciência de Dados, Engenharia de Dados e Computação em Nuvem. Participa do projeto de Iniciação Científica intitulado "Inteligência Artificial e Segurança Cibernética Aplicadas ao Atendimento Acadêmico Automatizado". Possui as certificações na AWS e OCI.



Júlio César Mansano é graduando em Tecnologia em Segurança da Informação pela Fatec Americana - Faculdade de Tecnologia de Americana Ministro Ralph Biasi, com conclusão prevista para junho de 2026. É Engenheiro de Produção Química formado pela Universidade Federal de São Carlos (UFSCar) e possui MBA em Gestão Financeira pela UNIMEP. Atualmente, ocupa a posição de Coordenador de Gestão e Qualidade na ColorFlex Textil. Detém a certificação Six Sigma Black Belt e diversas certificações da Cisco Networking Academy. Suas áreas de interesse e atuação englobam análise de dados, inteligência artificial e segurança cibernética.



João Emmanuel D'Alkmin Neves possui graduação em Análise e Desenvolvimento de Sistemas pela FATEC Americana, com período sanduíche em Ciência da Computação pela State University of New York (SUNY, 2015). É mestre (2018) e doutor (2024) em Tecnologia pela Universidade Estadual de Campinas (UNICAMP). Atualmente, é Docente de Ensino Superior na FATEC Americana e atua como Editor da Revista Tecnológica da Fatec Americana. Suas principais áreas de pesquisa e ensino incluem Inteligência Artificial, Sistemas Multiagentes, Segurança da Informação, Redes de Computadores e Computação em Nuvem. Coordena diversos projetos de Iniciação Científica e de Inovação Tecnológica na instituição.

Desenvolvimento de Chatbot Acadêmico com Princípios de Acessibilidade Digital

Aluan Luiz Maturi Horschutz, Fernanda Diniz Marabin
& João Emmanuel D'Alkmin Neves
Faculdade de Tecnologia de Americana – Fatec
aluanluiz@outlook.com, nmarabin@gmail.com, joao.neves@cps.sp.gov.br

Abstract—The increasing digitalization of educational environments requires systems developed with a focus on digital accessibility, ensuring equity in access to information. This work presents the development of an academic chatbot for Fatec Americana, implemented in Python with Flask and an interface built with HTML, CSS, and JavaScript. The methodology was based on the practical application of the WCAG 2.1 guidelines, including adequate visual contrast, semantic HTML structure, keyboard navigation, and the use of ARIA attributes for compatibility with screen readers. The results indicate compliance with level A and AA criteria in the implemented axes, demonstrating the feasibility of integrating Artificial Intelligence and accessibility principles in educational web applications. It is concluded that the systematic adoption of WCAG guidelines contributes to expanding digital inclusion in the academic context.

Index Terms—Academic chatbot, Artificial Intelligence, Digital accessibility, WCAG 2.1.

Resumo— A crescente digitalização dos ambientes educacionais exige sistemas desenvolvidos com foco em acessibilidade digital, garantindo equidade no acesso à informação. Este trabalho apresenta o desenvolvimento de um chatbot acadêmico para a Fatec Americana, implementado em Python com Flask e interface em HTML, CSS e JavaScript. A metodologia baseou-se na aplicação prática das diretrizes WCAG 2.1, contemplando contraste visual adequado, estrutura HTML semântica, navegação por teclado e uso de atributos ARIA para compatibilidade com leitores de tela. Os resultados indicam conformidade com critérios de nível A e AA nos eixos implementados, demonstrando a viabilidade da integração entre Inteligência Artificial e princípios de acessibilidade em aplicações web educacionais. Conclui-se que a adoção sistemática das diretrizes WCAG contribui para ampliar a inclusão digital no contexto acadêmico.

Palavras chave—Acessibilidade digital, Chatbot acadêmico, Inteligência Artificial, WCAG 2.1.

I. Introdução

A transformação digital ampliou significativamente o acesso a informações e serviços educacionais, tornando os ambientes virtuais parte essencial do processo de ensino-aprendizagem. Nesse contexto, a acessibilidade digital assume papel fundamental ao garantir que todas as pessoas possam acessar, compreender e interagir com conteúdos online de forma autônoma e eficiente.

A acessibilidade digital consiste na eliminação de barreiras na Web, permitindo que indivíduos com limitações físico-motoras, sensoriais ou cognitivas utilizem sistemas digitais com independência [2]. Além de beneficiar pessoas com deficiência permanente, práticas acessíveis também favorecem usuários com limitações temporárias e contribuem para interfaces mais claras e organizadas para todos.

No Brasil, apesar da existência da Lei Brasileira de Inclusão (Lei nº 13.146/2015) [1], grande parte dos ambientes digitais ainda não atende plenamente aos critérios de acessibilidade. Estudos indicam que menos de 1% dos websites brasileiros são totalmente acessíveis, evidenciando um cenário de exclusão digital [3]. Essa realidade torna-se ainda mais preocupante no contexto educacional, onde o acesso equitativo à informação está diretamente relacionado ao direito fundamental à educação.

As diretrizes internacionais de acessibilidade, como as Web Content Accessibility Guidelines (WCAG), estabelecidas pelo World Wide Web Consortium (W3C), definem parâmetros técnicos para tornar conteúdos digitais mais inclusivos [4]. Paralelamente, os avanços da Inteligência Artificial (IA) têm possibilitado o desenvolvimento de tecnologias assistivas capazes de ampliar a autonomia de usuários com deficiência, como demonstrado em pesquisas recentes sobre o uso de IA em ferramentas digitais inclusivas [5].

Considerando o crescente uso de plataformas digitais no ambiente educacional, torna-se essencial que novas soluções tecnológicas sejam desenvolvidas já incorporando princípios de acessibilidade. Nesse contexto, ferramentas baseadas em Inteligência Artificial, como chatbots, apresentam potencial para facilitar o acesso à informação acadêmica, oferecer suporte automatizado a estudantes e contribuir para a redução de barreiras de acesso em ambientes virtuais de aprendizagem.

Diante desse cenário, este trabalho tem como objetivo desenvolver um chatbot acadêmico com foco em acessibilidade digital, integrando princípios inclusivos ao uso de Inteligência Artificial em ambientes educacionais.

II. REFERENCIAL TEÓRICO

Esta seção apresenta os conceitos teóricos que fundamentam o desenvolvimento do chatbot acadêmico com foco em

acessibilidade digital.

A inclusão digital vai além do acesso físico às tecnologias, envolvendo a capacidade de utilização autônoma e significativa dos recursos digitais. A exclusão ocorre quando barreiras técnicas, cognitivas ou estruturais impedem a plena participação no ambiente virtual [6]. Nesse contexto, a acessibilidade digital constitui elemento essencial para a promoção da equidade no uso da informação.

As diretrizes internacionais Web Content Accessibility Guidelines (WCAG), estabelecidas pelo World Wide Web Consortium (W3C), definem critérios para tornar conteúdos digitais perceptíveis, operáveis, compreensíveis e robustos [4]. A adoção dessas diretrizes contribui para a redução de barreiras e para a democratização do acesso à informação.

B. Conceito e Evolução dos Chatbots

Chatbots são sistemas computacionais desenvolvidos para simular interações humanas por meio de texto ou voz. Inicialmente baseados em regras e respostas pré-programadas, apresentavam limitações na interpretação da linguagem natural e na adaptação às variações das consultas dos usuários.

Com os avanços da Inteligência Artificial, esses sistemas passaram a utilizar técnicas capazes de identificar intenções e interpretar contextos, ampliando sua aplicabilidade em diferentes áreas, incluindo o ambiente educacional.

C. Fundamentos Tecnológicos

Os chatbots contemporâneos utilizam técnicas de Processamento de Linguagem Natural (PLN), Compreensão de Linguagem Natural (NLU) e aprendizado de máquina para interpretar entradas e gerar respostas coerentes. O uso de deep learning e grandes modelos de linguagem (LLMs) ampliou significativamente a capacidade conversacional desses sistemas, permitindo maior contextualização e adaptação às necessidades do usuário.

Esses avanços possibilitam aprendizado contínuo a partir das interações, aumentando a eficiência na mediação de informações.

D. Chatbots, Educação e Acessibilidade

No contexto educacional, os chatbots podem atuar como ferramentas de apoio ao ensino, auxiliando no esclarecimento de dúvidas, fornecimento de informações acadêmicas e orientação discente. Sua disponibilidade contínua favorece o acesso rápido à informação e pode contribuir para maior autonomia dos estudantes.

Quando desenvolvidos com foco em acessibilidade digital, esses sistemas podem integrar recursos como adaptação de linguagem, personalização de respostas e interação multimodal, ampliando a inclusão de usuários com diferentes necessidades. Estudos sobre o uso de Inteligência Artificial cognitiva em ferramentas digitais indicam que tais tecnologias podem favorecer processos inclusivos,

A. Inclusão e Acessibilidade Digital

especialmente para pessoas com necessidades específicas [5].

Entretanto, apesar dos avanços tecnológicos, ainda persistem desafios relacionados à conformidade com diretrizes de acessibilidade, à prevenção de vieses algorítmicos e ao desenvolvimento centrado no usuário. Dessa forma, a integração entre tecnologias conversacionais baseadas em IA e princípios de acessibilidade digital representa uma estratégia relevante para a democratização do acesso ao conhecimento.

III. METODOLOGIA

A. Contexto e Objetivo do Sistema

A metodologia adotada neste trabalho fundamentou-se, inicialmente, em pesquisa bibliográfica de caráter exploratório, com levantamento de obras consagradas e diretrizes normativas sobre acessibilidade digital e desenvolvimento de chatbots acadêmicos. Foram consultadas as diretrizes WCAG 2.1 e WAI-ARIA 1.2 do W3C [4, 10], o Modelo de Acessibilidade em Governo Eletrônico (eMAG) [5], a Lei Brasileira de Inclusão [1], o relatório WebAIM Million [3] e revisões sistemáticas da literatura sobre avaliação de acessibilidade em sistemas web [11]. Essa base teórica orientou a seleção dos critérios implementados e o embasamento das decisões técnicas adotadas no frontend.

O sistema desenvolvido é um chatbot acadêmico voltado a estudantes da Fatec Americana, acessível via navegador web. Seu objetivo principal é permitir que os alunos consultem informações institucionais — como horários de aulas, disciplinas oferecidas e dados de cursos — de forma conversacional, sem necessidade de navegar por múltiplas páginas do portal institucional.

O backend foi implementado em Python com o microframework Flask, responsável pelo roteamento das requisições HTTP e pela integração com o modelo de linguagem utilizado para processamento das consultas. O frontend é composto por HTML, CSS e JavaScript puros, renderizados via templates Jinja2 do Flask. A Figura 1 ilustra a arquitetura geral do sistema.

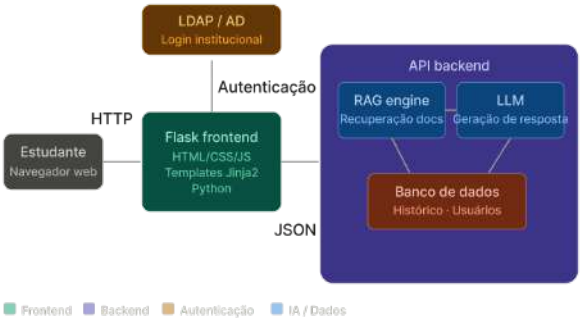


Fig. 1. Arquitetura geral do chatbot acadêmico.

B. Critérios WCAG 2.1 Adotados

A seleção das práticas de acessibilidade foi fundamentada nas diretrizes WCAG 2.1 [4] e no guia WAI-ARIA 1.2 [10], priorizando os critérios de maior impacto para usuários de tecnologias assistivas em interfaces conversacionais. Foram implementados três eixos, descritos nas subseções a seguir.

B.1) Estrutura HTML Semântica — WCAG 1.3.1

O critério 1.3.1 (nível A) exige que a estrutura e o relacionamento entre elementos da interface sejam determinados programaticamente, de modo que tecnologias assistivas possam interpretar corretamente o conteúdo [4].

Nos templates Jinja2 do projeto, a página foi organizada com elementos HTML semânticos nativos: <header> para o cabeçalho da aplicação, <main> para a área de interação principal e <section> para separação lógica entre o histórico de mensagens e o campo de entrada. O histórico de mensagens foi estruturado como lista <ul> com itens <li>, atribuindo hierarquia e ordem compreensível para leitores de tela. Todos os elementos interativos foram implementados com a tag nativa <button>, garantindo comportamento de teclado e semântica corretos sem necessidade de atributos ARIA adicionais. A Figura 2 apresenta a estrutura HTML resultante.

```
<section class="chat-container" aria-label="Interface de
chat com FatecIA">
  <header class="chat-header">
    <h1 class="chat-titulo">FatecIA</h1>
    <p aria-label="Usuário logado">Olá, {{
session.get('usuario') }}</p>
  </header>
  <main id="conteudo-principal"
tabindex="-1">
    {% block content %}{% endblock %}
  </main>
</section>
```

Fig. 2. Estrutura HTML semântica dos templates Jinja2.

B.2) Navegação por Teclado — WCAG 2.1.1 e 2.4.7

O critério 2.1.1 (nível A) exige que toda funcionalidade da interface seja operável exclusivamente via teclado, sem dependência de mouse ou dispositivo apontador. O critério 2.4.7 (nível AA) exige que o indicador de foco seja sempre visível durante a navegação [4].

A ordem de foco dos elementos foi revisada nos templates para seguir o fluxo lógico da interface: cabeçalho, área de histórico e campo de entrada de texto. O envio de mensagens via tecla Enter foi preservado como comportamento padrão do formulário HTML, sem necessidade de JavaScript adicional. O estilo :focus-visible foi definido no CSS com outline explícito em todos os elementos interativos, garantindo visibilidade do foco para usuários que navegam exclusivamente por teclado. Nenhum elemento interativo utiliza outline: none sem substituto visual.

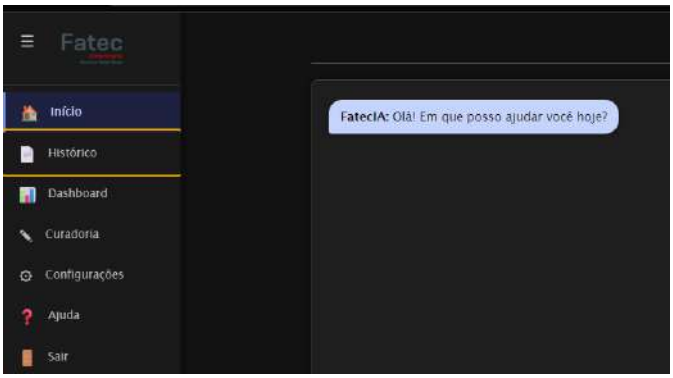


Fig. 3. Indicador de foco visível durante navegação por teclado.

B.3) Compatibilidade com Leitores de Tela — WCAG 4.1.2

O critério 4.1.2 (nível A) exige que todos os componentes da interface tenham nome, função e valor determinados programaticamente, de modo a serem corretamente interpretados por tecnologias assistivas [4]. Para interfaces conversacionais, esse critério apresenta um desafio específico: novas mensagens são inseridas dinamicamente no DOM, sem recarregamento da página, e leitores de tela não as anunciam automaticamente sem configuração adequada.

Para endereçar esse problema, o container de mensagens recebeu o atributo aria-live="polite" nos templates HTML, garantindo que cada nova resposta do chatbot seja anunciada pelo leitor de tela sem interromper o foco atual do usuário [10]. O atributo aria-busy="true" foi aplicado ao container durante o período de carregamento da resposta, sinalizando ao leitor de tela que o conteúdo está sendo processado. O campo de entrada de texto foi associado a um elemento <label> via atributos for e id, fornecendo rótulo legível. Botões com ícones sem texto visível receberam aria-label descritivo, como "Enviar mensagem".

```
<!-- WCAG 4.1.2 / 4.1.3 — container de mensagens -->
<div id="chat-mensagens"
     role="log"
     aria-label="Histórico da conversa"
     aria-live="polite"
     aria-relevant="additions"></div>
<!-- Botão com aria-label e estado
aria-pressed -->
<button id="btn-enviar"
        aria-label="Enviar mensagem"
        disabled>Enviar</button>
<!-- Label associado ao campo de entrada
-->
<label for="chat-field"
class="sr-only">Mensagem</label>
<input type="text" id="chat-field"
       aria-describedby="chat-status">
```

Fig. 4. Atributos ARIA aplicados ao container de mensagens e botões.

C. Método de Avaliação

A avaliação das implementações foi conduzida por meio de teste manual de navegação exclusiva por teclado, percorrendo todos os elementos interativos da interface sem uso de mouse. O teste verificou: (i) alcançabilidade de todos os elementos via tecla Tab; (ii) ativação de botões e envio de formulário via Enter e Espaço; (iii) visibilidade do indicador de foco em cada

elemento navegado; e (iv) retorno do foco ao campo de entrada após o envio de uma mensagem.

A avaliação automatizada com ferramentas como Lighthouse e axe DevTools, bem como testes com leitores de tela reais (NVDA, VoiceOver), estão planejados como etapas futuras do projeto, conforme discutido na Seção V.

IV. RESULTADOS E DISCUSSÃO

Os resultados são apresentados por eixo de implementação e consolidados na Tabela I, que sintetiza os critérios WCAG 2.1 endereçados, os elementos técnicos aplicados e o nível de conformidade atingido em cada caso.

TABELA I

CRITÉRIOS WCAG 2.1 IMPLEMENTADOS E NÍVEL DE CONFORMIDADE

Critério WCAG 2.1	Elemento Técnico	Nível	Status
1.3.1 — Informações e Relações	HTML semântico: <header>, <main>, <section>, / , <button>	A	✓ Atendido
2.1.1 — Teclado	Tab, Enter/Espaço; ordem de foco revisada; teste manual realizado	A	✓ Atendido
2.4.7 — Foco Visível	:focus-visible com outline explícito em todos os elementos interativos	AA	✓ Atendido
4.1.2 — Nome, Função, Valor	aria-live="polite", aria-busy, aria-label, <label for>	A	✓ Atendido
1.4.3 — Contraste (Mínimo)	Razão mínima 4,5:1 — não verificada sistematicamente nesta versão	AA	○ Pendente

Fonte: elaborado pelos autores com base em W3C [4].

A. Estrutura HTML Semântica

A adoção de elementos HTML semânticos nativos nos templates Jinja2 resultou em uma estrutura de página interpretável por tecnologias assistivas sem necessidade de atributos ARIA adicionais. A hierarquia <header> / <main> / <section> define regiões de navegação reconhecidas por leitores de tela, permitindo que o usuário salte diretamente para a área de interesse. A organização do histórico como lista / transmite ao leitor de tela a quantidade de mensagens presentes e a posição atual na conversa, comportamento documentado nas especificações WAI-ARIA 1.2 [10].

O uso de <button> nativo, em substituição a elementos com eventos de clique, eliminou a necessidade de adicionar manualmente role="button" e tabindex="0". Essa abordagem está alinhada à primeira regra de uso do ARIA definida pelo W3C: não utilizar ARIA quando HTML nativo resolve o mesmo problema [10]. A Figura 2 apresenta o trecho da estrutura HTML resultante nos templates Jinja2.

```
<section class="chat-container" aria-label="Interface de chat com FatecIA">
  <header class="chat-header">
    <h1 class="chat-titulo">FatecIA</h1>
    <p aria-label="Usuário logado">Olá, {{ session.get('usuario') }}</p>
  </header>
  <main id="conteudo-principal" tabindex="-1">
    {% block content %}{% endblock %}
  </main>
</section>
```

Fig. 2. Estrutura HTML semântica gerada pelos templates Jinja2.

B. Navegação por Teclado

O teste manual de navegação exclusiva por teclado confirmou a operabilidade completa da interface sem uso de mouse. Todos os elementos interativos — campo de entrada, botão de envio e demais controles — foram alcançados via Tab em ordem lógica, correspondente ao fluxo natural de uso do chatbot. O envio de mensagens via Enter funcionou conforme o comportamento padrão do formulário HTML, sem necessidade de JavaScript adicional. Esse resultado está alinhado ao requisito do critério WCAG 2.1.1 (nível A), que exige que toda funcionalidade seja operável por teclado [4].

O indicador de foco visível, implementado via :focus-visible no CSS, esteve presente em todos os elementos navegados durante o teste, atendendo ao critério 2.4.7 (nível AA). A Figura 3 ilustra o estado de foco aplicado a um dos elementos interativos da interface.

```
/* WCAG 2.4.7 — foco visível em todos os elementos interativos */
button:focus-visible,
input:focus-visible,
a:focus-visible {
    outline: 3px solid #005FCC;
    outline-offset: 3px;
    border-radius: 4px;
}
```

Fig. 3. Indicador de foco visível durante navegação por teclado.

C. Compatibilidade com Leitores de Tela

A implementação do atributo aria-live="polite" no container de mensagens representa a intervenção de maior impacto funcional para usuários de leitores de tela em interfaces conversacionais. Sem esse atributo, o usuário com deficiência visual não recebe notificação de que o chatbot respondeu — problema recorrente identificado na literatura sobre acessibilidade em interfaces dinâmicas [10]. O relatório WebAIM Million [3] reforça a gravidade do cenário ao identificar que a ausência de rótulos e regiões ARIA é uma das falhas mais frequentes em sistemas web, presente em mais de 60% das páginas avaliadas.

O atributo aria-busy="true", aplicado durante o carregamento da resposta, impede que o leitor de tela leia conteúdo parcialmente inserido no DOM. Os aria-label nos botões e o <label> associado ao campo de entrada garantem nome programaticamente determinado a todos os elementos

interativos, conforme exigido pelo critério 4.1.2 (nível A) [4]. A Figura 4 apresenta o trecho dos templates com os atributos ARIA aplicados.

```
// WCAG 4.1.3 — região de status para anunciar estados
const statusRegion =
document.getElementById('chat-status');
function setStatus(msg) {
    if (statusRegion)
statusRegion.textContent = msg;
}
// WCAG 4.1.2 — aria-pressed comunica
estado toggle
btnUtil.setAttribute('aria-pressed',
'false');
btnUtil.setAttribute('aria-label', 'Marcar
como útil');
// Ao confirmar avaliação:
btnUtil.setAttribute('aria-pressed',
'true');
// WCAG 4.1.3 — anuncia confirmação ao
leitor de tela
setStatus('Resposta marcada como útil.');
```

Fig. 4. Atributos ARIA aplicados ao container de mensagens e aos botões.

D. Discussão Geral

Conforme sintetizado na Tabela I, os três eixos implementados cobrem quatro critérios WCAG 2.1 de nível A e AA, patamares mínimos exigidos pela Lei Brasileira de Inclusão [1] e pelo eMAG [5] para sistemas digitais em contexto educacional. A implementação foi viável dentro da arquitetura Flask/Jinja2, sem frameworks JavaScript adicionais, demonstrando que a conformidade com diretrizes de acessibilidade não exige reestruturação tecnológica do projeto.

O critério 1.4.3 (contraste mínimo, nível AA) permanece pendente nesta versão, o que representa risco de exclusão para usuários com baixa acuidade visual ou daltonismo. Sua ausência é registrada como prioridade para a próxima iteração.

A avaliação restringiu-se ao teste manual por teclado. Revisões sistemáticas da literatura indicam que a combinação de avaliação automatizada e testes com usuários reais é necessária para identificar problemas não detectáveis por inspeção manual [11], o que constitui etapa planejada para a continuidade do projeto.

V. CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento de um chatbot acadêmico com foco em acessibilidade digital, aplicando as diretrizes WCAG 2.1 no frontend construído com Flask, HTML, CSS e JavaScript. A implementação dos quatro eixos — contraste visual, semântica HTML, navegação por teclado e atributos ARIA — resultou em uma interface que atende aos critérios de conformidade de nível A e AA da WCAG 2.1. Dessa forma, o sistema proposto demonstra que é possível incorporar práticas de acessibilidade desde o processo de desenvolvimento, contribuindo para a construção de aplicações web mais inclusivas.

Os resultados obtidos evidenciam que a integração entre tecnologias conversacionais baseadas em Inteligência Artificial e princípios de acessibilidade digital é viável no contexto de aplicações educacionais. Além disso, a utilização de uma arquitetura simples, baseada em Flask e tecnologias web nativas, mostrou-se suficiente para implementar recursos de acessibilidade relevantes sem a necessidade de frameworks JavaScript adicionais, favorecendo a manutenção e a escalabilidade do sistema.

Apesar dos avanços apresentados, o trabalho possui algumas limitações. Não foram realizados testes sistemáticos com usuários reais com deficiência, o que é fundamental para validar a experiência de uso em cenários reais. Além disso, aspectos relacionados à acessibilidade cognitiva e à compatibilidade com diferentes leitores de tela ainda podem ser explorados de forma mais aprofundada em versões futuras da aplicação.

Como trabalhos futuros, propõe-se a realização de avaliações automatizadas com ferramentas como Lighthouse e axe DevTools para mensurar o nível de conformidade alcançado, bem como testes com usuários com deficiência visual e motora para validação prática das funcionalidades implementadas. Também se pretende ampliar o sistema com recursos voltados à acessibilidade cognitiva e adequar a aplicação aos critérios adicionais introduzidos pela WCAG 2.2, contribuindo para o aprimoramento contínuo da solução proposta.

REFERÊNCIAS

[1] BRASIL. Lei nº 13.146, de 6 de julho de 2015. Lei Brasileira de Inclusão da Pessoa com Deficiência. Diário Oficial da União, Brasília, 2015.

[2] BRASIL. Governo Digital. Acessibilidade Digital. Disponível em: https://www.gov.br/governodigital/pt-br/acessibilidade-e-usuario/acessibilidade-digital

[3] WEBAIM. The WebAIM Million: An annual accessibility analysis of the top 1,000,000 home pages. 2023. Disponível em: https://webaim.org/projects/million/

[4] WORLD WIDE WEB CONSORTIUM (W3C). Web Content Accessibility Guidelines (WCAG) 2.1. W3C Recommendation, 2018. Disponível em: https://www.w3.org/TR/WCAG21/

[5] BRASIL. Ministério do Planejamento. eMAG — Modelo de Acessibilidade em Governo Eletrônico, v. 3.1. Brasília, 2014. Disponível em: https://emag.governoeletronico.gov.br/

[6] SANTOS, T. et al. IFBot — Inteligência artificial aplicada à educação. Caderno Pedagógico, v. 22, n. 9, out. 2025. Disponível em: https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/20507

[7] RODRIGUES, L. et al. Sabiá: Um Chatbot de IA Generativa para Suporte no Ensino Superior. arXiv, nov. 2025. Disponível em: https://arxiv.org/abs/2511.10787

[8] PASSERINO, L. M.; MONTARDO, S. P. Inclusão digital e acessibilidade digital: interfaces e aproximações conceituais. 2007. Disponível em: https://redessociaiseinclusao.pbworks.com/f/ID-acess_compos_2007_versão%20final.pdf

[9] SEMENSATO, M. C. F.; AMAFALDO, R. A.; NEVES, J. E. D. O Uso da Inteligência Artificial Cognitiva em Ferramentas Digitais para o Transtorno do Espectro Autista. In: SICT, 2024.

[10] W3C. WAI-ARIA 1.2 — Accessible Rich Internet Applications. W3C Recommendation, jun. 2023. Disponível em: https://www.w3.org/TR/wai-aria-1.2/

[11] CARVALHO, L. et al. Accessibility engineering in web evaluation process: a systematic literature review. PeerJ Computer Science, 2022. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC9881526/



Aluan Luiz Maturi Horschutz é estudante de Segurança da Informação na Fatec Americana e técnico formado em Análise e Desenvolvimento de Sistemas pela ETEC Polivalente de Americana. Atua como Suporte Técnico Jr. no Mercado Livre, com foco em field service, gestão de ativos de TI e atendimento a incidentes Paralelamente, contribui como pesquisador voluntário em projeto de iniciação científica na Fatec, desenvolvendo uma plataforma web de atendimento acadêmico com IA.



Fernanda Diniz Marabin é técnica em Desenvolvimento de Sistemas pela ETEC
Polivalente de Americana e atualmente cursa Jogos Digitais na FATEC Americana. Desenvolve projetos nas áreas de desenvolvimento de software e tecnologias interativas, com interesse em acessibilidade digital, inclusão tecnológica e aplicação de Inteligência Artificial em ambientes educacionais. Dedica-se à criação de soluções digitais voltadas à inovação e à democratização do acesso à informação.



João Emmanuel D'Alkmin Neves possui graduação em Análise e Desenvolvimento de Sistemas pela FATEC Americana, com período sanduíche em Ciência da Computação pela State University of New York (SUNY, 2015). É mestre (2018) e doutor (2024) em Tecnologia pela Universidade Estadual de Campinas (UNICAMP). Atualmente, é Docente de Ensino Superior na FATEC Americana e atua como Editor da Revista Tecnológica da Fatec Americana. Suas principais áreas de pesquisa e ensino incluem Inteligência Artificial, Sistemas Multiagentes, Segurança da Informação, Redes de Computadores e Computação em Nuvem. Coordena diversos projetos de Iniciação Científica e de Inovação Tecnológica na instituição.

Estudo de caso sobre o equipamento contador Geiger-Müller

Brenno Canellas, Felipe Emanuel Chaves, Nicholas Lima do Nascimento, Roger Pereira Freitas, Rita Elizabeth Santos de Almeida

Instituto Nacional de Telecomunicações – Inatel
brenno.oliveira@geb.inatel.br, felipe.chaves@inatel.br, nicholas.lima@ges.inatel.br, roger.pereira@ges.inatel.br, r.elizabeth@inatel.br

***Abstract*— In this work, a low-cost radiation counter was developed and tested using a commercially available detection board purchased online, coupled to a Geiger–Müller tube and integrated with a microcontroller for recording the counts. An accessible radioactive source was a gas lamp mantle, known for containing radioactive elements in its composition. In addition to the complete assembly of the board and the implementation of the code responsible for reading the pulses generated by the Geiger tube, the article also presents a review of the main theoretical concepts related to Medical Physics and the detection of ionizing radiation, including interaction mechanisms, the operating principles of the Geiger–Müller tube, and statistical aspects of the counting process.**

***Index Terms*— Geiger counter, Medical Physics, Ionizing radiation.**

***Resumo*—Neste trabalho, foi desenvolvido e testado um contador de radiação de baixo custo a partir de uma placa de detecção comercial adquirida pela internet, acoplada a um tubo Geiger-Müller e integrada a um microcontrolador para registro das contagens. Como fonte radioativa acessível, utilizou-se uma camisinha de lampião, reconhecida pela presença de elementos radiativos em sua composição. Além da montagem completa da placa e da implementação do código de leitura dos pulsos no tubo Geiger, o artigo também apresenta uma revisão dos principais conceitos teóricos relacionados à Física Médica e à detecção de radiação ionizante, incluindo mecanismos de interação, fundamentos do tubo Geiger-Müller e aspectos estatísticos do processo de contagem.**

***Palavras Chave*— Contador Geiger, Física Médica, Radiação ionizante.**

I.. Introdução

A detecção e o monitoramento da radiação ionizante desempenham papel fundamental em diversas áreas, incluindo medicina nuclear, indústria, pesquisa científica e proteção radiológica. A capacidade de medir a presença e a intensidade da radiação é essencial para garantir a segurança de profissionais que trabalham com materiais radioativos, bem como para possibilitar estudos experimentais em física das radiações [1, 2].

No Brasil, a realidade é que a maioria dos equipamentos utilizados para detecção e monitoramento da radiação é importado. Por isso, os custos de aquisição, manutenção e calibração desses instrumentos podem ser elevados, além de envolverem processos burocráticos associados à certificação e à regulamentação. Como consequência, instituições de ensino e pequenos laboratórios frequentemente enfrentam dificuldades para obter acesso a instrumentos adequados para experimentação e ensino na área de física nuclear e proteção radiológica.[3, 2].

Estudos presentes na literatura demonstram que detectores baseados em tubos Geiger-Müller (GM) podem ser integrados a microcontroladores e sistemas embarcados, permitindo o desenvolvimento de instrumentos de detecção de radiação relativamente simples, robustos e de baixo custo [4]. Essas soluções apresentam grande potencial para aplicações educacionais e para o monitoramento básico de radiação em ambientes laboratoriais. [5]

Diante desse cenário, o presente trabalho propõe o desenvolvimento e a avaliação experimental de um contador GM de baixo custo, baseado na integração de um tubo detector do modelo M4011 [6] com uma plataforma de prototipagem eletrônica Arduino.

O objetivo principal deste estudo foi investigar a viabilidade técnica de um sistema de detecção de radiação de baixo custo, avaliando seu potencial como ferramenta de apoio ao ensino de física das radiações e à introdução de conceitos de proteção radiológica em ambientes acadêmicos.[7]

II.. Fundamentação Teórica

A. Fundamentos da Radiação

Desde a descoberta da radiação ionizante no final do século XIX, o entendimento da estrutura da matéria sofreu profundas transformações, possibilitando o desenvolvimento de diversas aplicações científicas, industriais e médicas que permanecem essenciais até os dias atuais. Radiações ionizantes são formas de energia capazes de alterar a estrutura eletrônica dos átomos, promovendo processos de ionização e excitação [8]. Tradicionalmente, essas radiações são classificadas em radiações particuladas, como partículas alfa (α) e beta (β), e radiações eletromagnéticas, como raios X e radiação gama (γ) [2]. Cada tipo apresenta características próprias de energia, penetração, origem física e mecanismos de interação com a matéria [5].

As partículas alfa, formadas por dois prótons e dois nêutrons, possuem elevado poder ionizante, porém baixo alcance em meios materiais, sendo facilmente absorvidas por pequenas camadas de matéria ou mesmo pelo ar. As partículas beta, constituídas por elétrons ou pósitrons emitidos durante processos de decaimento nuclear, apresentam maior capacidade de penetração. Já os raios X e a radiação gama correspondem a fótons de alta energia capazes de atravessar materiais relativamente densos e interagir ao longo de trajetórias extensas [1].

As interações fundamentais entre radiação ionizante e matéria incluem fenômenos como o efeito fotoelétrico, o espalhamento Compton e a produção de pares. Esses processos dependem diretamente da energia da radiação incidente e das propriedades do material atravessado, sendo responsáveis pela geração de pares elétron–íon que possibilitam a detecção e a quantificação da radiação por meio de dispositivos físicos [2].

O estudo e a compreensão desses fenômenos constituem parte fundamental da Física Médica, área responsável pela

aplicação dos princípios da física em procedimentos diagnósticos, terapêuticos e em práticas de proteção radiológica. Nesse contexto, experimentos práticos que permitem observar fenômenos de interação e detecção da radiação desempenham papel relevante na formação acadêmica, contribuindo para o entendimento experimental dos conceitos associados à física das radiações [8].

B. Princípios de funcionamento do contador Geiger-Müller

Os detectores de radiação desempenham papel essencial na quantificação, monitoramento e análise de fenômenos envolvendo radiação ionizante. A capacidade de medir a presença e a intensidade da radiação é fundamental para aplicações em pesquisa científica, medicina, indústria e proteção radiológica [2].

Diversos tipos de detectores são empregados para converter a energia depositada pela radiação em sinais elétricos mensuráveis. Entre os principais estão os detectores gasosos, detectores cintiladores, detectores semicondutores e câmaras de ionização, cada um baseado em diferentes mecanismos físicos de interação entre radiação e matéria [2].

Entre os detectores gasosos, destaca-se o contador GM, amplamente utilizado devido à sua simplicidade construtiva, robustez operacional e baixo custo. Essas características tornam esse tipo de detector particularmente adequado para aplicações educacionais, monitoramento ambiental e sistemas portáteis de detecção de radiação [2]. O dispositivo foi inicialmente desenvolvido por Hans Geiger e posteriormente aprimorado por Walther Müller, passando a operar em uma região específica da curva característica de detectores gasosos conhecida como região Geiger.

Quando uma partícula ou fóton ionizante penetra no tubo detector, ocorre a ionização dos átomos do gás presente em seu interior, produzindo pares elétron–íon positivo. Sob a ação de um intenso campo elétrico estabelecido entre o cátodo e o ânodo, os elétrons livres são acelerados em direção ao ânodo, enquanto os íons positivos se deslocam em direção ao cátodo [2].

Durante esse deslocamento, os elétrons podem adquirir energia suficiente para provocar novas ionizações por colisão com outros átomos do gás. Esse processo multiplicativo gera sucessivas liberações de cargas, fenômeno conhecido como avalanche de Townsend. Como consequência, forma-se um pulso elétrico de corrente que pode ser detectado por circuitos eletrônicos externos [5].

Uma característica fundamental do regime GM é que a avalanche se propaga por grande parte do volume sensível do detector. Os pulsos elétricos produzidos apresentam amplitude praticamente constante, independentemente da energia da radiação incidente. Essa propriedade confere elevada sensibilidade ao detector, porém impede a discriminação energética, tornando o tubo GM incapaz de distinguir o tipo ou a energia da radiação detectada [2].

Para garantir a estabilidade do funcionamento do tubo e evitar descargas contínuas, utiliza-se um gás extintor (*quencher*), geralmente composto por halogênios ou vapores orgânicos em pequenas concentrações. Esse gás é responsável por absorver a energia residual associada aos íons positivos, interrompendo o processo de ionização em cadeia após cada evento detectado [2].

Do ponto de vista eletrônico, o tubo GM necessita de uma fonte de alta tensão contínua, normalmente na faixa de 300 V a 500 V, para manter o campo elétrico adequado à formação das avalanches [2]. Em sistemas modernos de baixo custo, esse circuito de alta tensão pode estar integrado a módulos eletrônicos compactos que fornecem a polarização necessária ao tubo e disponibilizam uma saída digital compatível com níveis lógicos TTL.

Essa característica possibilita a integração direta com microcontroladores e sistemas embarcados, permitindo a contagem dos pulsos gerados pelo detector e a determinação da taxa de contagem ao longo do tempo, geralmente expressa em contagens por minuto (*Counts Per Minute* – CPM). A partir desses dados, e mediante fatores de calibração adequados, também é possível estimar a taxa de dose equivalente em unidades de μSv/h [2].

Apesar de suas limitações quanto à discriminação energética e identificação do tipo de radiação incidente, o contador GM apresenta elevada sensibilidade, simplicidade operacional e baixo custo, o que explica sua ampla utilização em experimentos didáticos, monitoramento ambiental e sistemas de detecção de radiação de baixo custo [5].

C. Grandezas e Unidades de Dosimetria

A quantificação da radiação ionizante envolve diferentes grandezas físicas utilizadas para descrever tanto a intensidade da radiação detectada quanto os efeitos que essa radiação pode produzir na matéria e nos organismos biológicos. No contexto de detectores de radiação, duas grandezas são frequentemente utilizadas: a taxa de contagem e a taxa de dose [2, 1].

A *taxa de contagem* corresponde à frequência com que eventos de ionização são detectados pelo instrumento. Essa grandeza é normalmente expressa em contagens por segundo (*Counts Per Second* – CPS) ou contagens por minuto (*Counts Per Minute* – CPM), representando simplesmente o número de pulsos elétricos gerados pelo detector em um determinado intervalo de tempo [2]. Em detectores do tipo GM, cada pulso corresponde à ocorrência de uma avalanche de ionização provocada pela interação de uma partícula ou fóton com o gás presente no interior do tubo detector.

Já a *taxa de dose* refere-se à quantidade de energia depositada pela radiação por unidade de massa em um determinado intervalo de tempo. Essa grandeza é fundamental para avaliar os possíveis efeitos biológicos da radiação em organismos vivos, sendo amplamente utilizada em aplicações de proteção radiológica e física médica [1].

No Sistema Internacional de Unidades (SI), a dose absorvida é medida em gray (Gy), unidade que representa a absorção de um joule de energia por quilograma de material. Entretanto, para avaliar os efeitos biológicos da radiação em tecidos humanos, utiliza-se a grandeza denominada *dose equivalente*, que leva em consideração fatores relacionados ao tipo de radiação incidente. A unidade utilizada para expressar a dose equivalente é o sievert (Sv) [1, 8].

Em muitos sistemas de detecção baseados em contadores GM, especialmente em instrumentos portáteis de monitora-

mento ambiental, a taxa de dose equivalente é frequentemente apresentada em microsieverts por hora (μSv/h). No entanto, não existe um fator de conversão universal e direto entre a taxa de contagem (CPM ou CPS) e a taxa de dose expressa em μSv/h. Essa conversão depende de diversos fatores, como a energia da radiação incidente, o tipo de radiação detectada e as características físicas e geométricas do detector utilizado [2].

Dessa forma, para obter estimativas confiáveis de dose a partir de medições de taxa de contagem, é necessário realizar processos de calibração específicos para cada tipo de detector e para as condições experimentais consideradas.

III.. Metodologia

O presente trabalho consiste no estudo e implementação de um contador GM de baixo custo nas dependências do Instituto Nacional de Telecomunicações (Inatel). A proposta baseia-se na utilização de um tubo GM modelo M4011, adquirido comercialmente por meio de fornecedor disponível em plataforma de comércio eletrônico. Divido a natureza genérica do equipamento, a validação técnica de sua operação é detalhada através do esquemático elétrico(Figura 1). O pino de saída da placa é conectada ao pino digital 2 do arduino Uno, e o conjunto foi alimentado pela própria saída de 5V do microcontrolador, garantindo o referencial comum de aterramento(GND) necessário para o registro confiável das interrupções.

A motivação central do projeto reside na investigação da viabilidade técnica da implementação de um sistema de detecção de radiação utilizando componentes acessíveis, preservando os princípios de funcionamento descritos na literatura. Busca-se, assim, estabelecer uma correlação entre os fenômenos físicos discutidos teoricamente e as respostas elétricas observadas no sistema experimental.

Como estratégia experimental, utilizou-se de mantas incandescentes de lampião, popularmente conhecidas como “camisinhas de lampião”. Esses dispositivos, tradicionalmente empregados em lampiões a gás, contêm em sua composição sais de terras raras associados ao tório, elemento naturalmente radioativo pertencente à série de decaimento do tório-232. Essa característica torna tais materiais uma fonte acessível e didaticamente relevante para a verificação experimental dos fenômenos de detecção da radiação ionizante.

O circuito eletrônico da placa (Figura 1) possui duas funções principais. A primeira consiste na geração da alta tensão necessária para a operação do tubo GM, elevando os 5V fornecidos pelo microcontrolador para uma faixa de aproximadamente 380–450 V. É importante ressaltar que, durante a calibração física do equipamento com um multímetro digital padrão, a leitura nos terminais do tubo resulta em cerca de 170V. Este comportamento é justificado pelo ”efeito de carga”, devido a baixa capacidade de corrente do circuito gerador, a impedância interna do próprio multímetro drena parte dessa corrente, provocando uma queda momentânea de tensão. Ao desconectar as pontas de prova, o circuito aberto retorna à sua faixa de projeto, garantindo a polarização correta. A segunda função corresponde ao condicionamento dos pulsos de corrente produzidos pelo tubo durante os eventos de ionização, convertendo-os em pulsos digitais adequados para processamento por microcontroladores.

A estrutura do circuito é composta pelos seguintes blocos funcionais:

Regulador de tensão: responsável pela elevação da tensão de alimentação (5 V provenientes da interface USB ou da alimentação da placa Arduino) para o patamar de alta tensão necessário ao funcionamento do tubo Geiger.

Estágio de detecção e modelagem de pulso: quando uma partícula ionizante ou fóton de alta energia interage com o gás no interior do tubo, ocorre um evento de ionização que resulta em um pulso de corrente de curta duração (da ordem de microssegundos) no ânodo do detector.

Interface de saída digital: o sinal condicionado é disponibilizado em um terminal de saída da placa, apresentando níveis lógicos compatíveis com a família TTL (0–5 V). Esse sinal pode ser conectado diretamente a uma porta digital de um microcontrolador, configurada para detecção de borda de descida (*falling edge*), permitindo a contagem precisa dos eventos detectados.

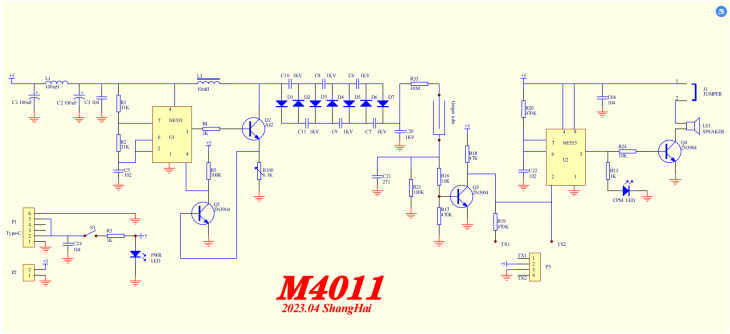


Fig. 1. Diagrama esquemático do módulo comercial de detecção M4011, estágio regulador de alta tensão e o circuito de condicionamento dos pulsos TTL

A. Testes

Inicialmente foram realizados testes preliminares no laboratório eHealth do INATEL, com o objetivo de avaliar o funcionamento do contador GM de baixo custo desenvolvido neste trabalho e identificar condições experimentais adequadas para a realização das medições. Nessa etapa, foram realizadas marcações na bancada do laboratório para definir diferentes distâncias entre a camisinha de lampião e o detector. As medições foram realizadas em intervalos de tempo fixos de 60 segundos, correspondendo à contagem de pulsos por minuto (CPM). Para cada distância estabelecida entre a fonte radioativa e o detector, foram realizadas múltiplas medições consecutivas, permitindo a obtenção de valores médios e a estimativa do desvio padrão das contagens registradas.

Os dados obtidos foram transmitidos via comunicação serial para um computador e armazenados automaticamente em arquivos no formato CSV por meio de um script desenvolvido em linguagem Python. Após cada conjunto de medições, a posição da fonte radioativa era ajustada para a próxima distância previamente marcada na bancada experimental, mantendo-se constantes as demais condições do experimento.

Posteriormente, foi realizada uma visita técnica à clínica de medicina nuclear MagSul, em Pouso Alegre, Minas Gerais, onde foi testado o emissor de radiação do projeto (camisinha de lampião) com um curiômetro certificado (câmara de ionização tipo poço, usada na medicina nuclear para medir com precisão atividades radioativas específicas) a fim de confirmar o tipo de radiação emitida pela camisinha de lampião. Após a confirmação, foram conduzidos testes comparativos utilizando o contador GM de baixo custo desenvolvido neste trabalho e um monitor de área certificado pela Comissão Nacional de Energia Nuclear (CNEN), disponibilizado pela empresa. Este instrumento fornece leituras em taxa de exposição, expressa em mR/h, permitindo estabelecer uma relação direta entre essa grandeza e a contagem de pulsos registrada pelo equipamento desenvolvido.

Para garantir que as medições pudessem ser repetidas posteriormente com a equivalência das distâncias entre a fonte e os detectores, foi utilizado um anteparo fixado sobre a mesa de medição e sobre essa mesma superfície foram realizadas marcações correspondentes às distâncias de 4 a 14 cm em relação à fonte radioativa, com espaçamento de 2 cm entre cada ponto de medição.

As medições foram realizadas inicialmente com o contador GM de baixo custo, registrando-se os valores de contagem de pulsos para cada distância estabelecida. Em seguida, mantendo-se a mesma configuração experimental e os mesmos pontos de distância previamente marcados na mesa, foram realizadas as medições utilizando o monitor de área certificado, possibilitando a obtenção dos valores correspondentes de taxa de exposição para as mesmas distâncias.

A Figura 2 apresenta o arranjo experimental utilizado durante a realização das medições, incluindo a disposição da fonte radioativa, as marcações de distância na mesa e os equipamentos empregados nos testes.



Fig. 2. Arranjo experimental utilizado para as medições comparativas com o contador GM de baixo custo e o monitor de área certificado.

Os dados obtidos a partir desse procedimento experimental foram posteriormente analisados para avaliar a relação entre a taxa de contagem registrada pelo detector de baixo custo e a taxa de exposição medida pelo instrumento certificado.

IV.. Resultados e Discussões

A. Resultado dos Testes e Comparações

Inicialmente, foi possível detectar e registrar os pulsos provenientes do tubo GM por meio do monitor serial do computador, utilizando a comunicação estabelecida com o microcontrolador Arduino Uno. A partir desses registros, foram realizadas as medições experimentais que serviram de base para as análises apresentadas a seguir. Inicialmente foram obtidas as medições utilizando o contador GM de baixo custo desenvolvido no projeto. A Tabela I apresenta os valores médios de contagem por minuto (CPM) em função da distância entre o detector e a fonte radioativa, bem como o desvio padrão das medições e os valores de CPM líquido após a correção do ruído de fundo.

TABELA I
CPM em função da distância

Distância (cm)	CPM Mediana	Desvio Padrão	CPM Líquido
2,0	378,9	25,13	325,0
4,0	264,5	25,8	206,5
6,0	199,0	17,8	141,0
8,0	160,0	21,5	102,0
10,0	143,0	20,7	85,0
12,0	123,0	16,6	65,0
14,0	114,0	12,3	56,0
16,0	103,0	12,8	45,0
18,0	88,0	12,6	30,0
20,0	94,0	19,8	36,0
22,0	65,0	8,9	7,0

$$\text{CPM Líquido} = \text{CPM Mediana} - 58\ \text{CPM (background)}$$

Observa-se uma clara redução nas contagens com o aumento da distância, comportamento esperado para uma fonte radioativa.

A partir desses resultados é traçada a curva de decaímento radioativo em função do distanciamento do emissor, veja Figura 3.

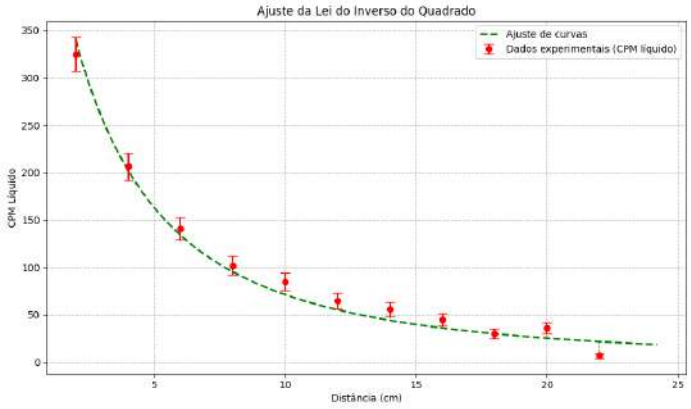


Fig. 3. Ajuste dos dados experimentais à lei do inverso do quadrado - Contador de baixo custo

Em seguida, foram utilizados os equipamentos certificados da clínica MagSul na cidade de Pouso Alegre. Primeiramente para confirmar que a radiação emitida pela camisinha de lampião era do tipo *gama*, o que foi feito com a utilização de um curiômetro profissional. Além disso, foram obtidas as medições utilizando um instrumento certificado de detecção radiológica do tipo GM. Os valores obtidos de taxa de exposição,

expressos em mR/h, para as mesmas distâncias usadas no caso do contador GM de baixo custo que estamos estudando, são apresentados na Tabela II a seguir:

TABELA II
DADOS OBTIDOS COM INSTRUMENTO CERTIFICADO

Distância (cm)	mR/h
4	0,06
6	0,04
8	0,03
10	0,02
12	0,01
14	0,005

Assim como observado nas medições do contador desenvolvido, os valores de exposição radiológica também apresentam uma curva de redução de exposição coerente com a teoria, conforme a Figura 4 a seguir:

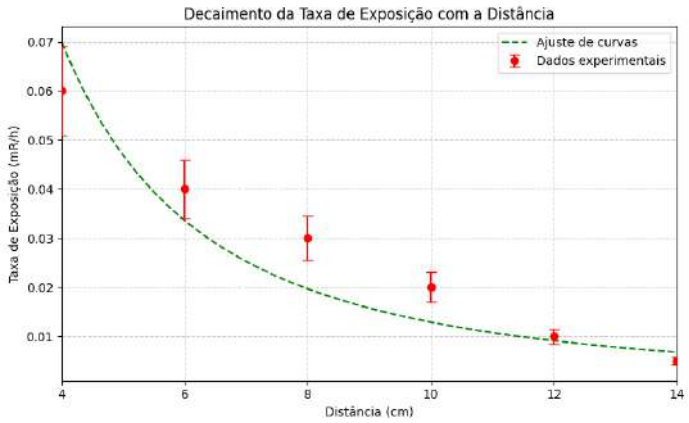


Fig. 4. Ajuste dos dados experimentais à lei do inverso do quadrado – contador GM profissional

Com base nos resultados obtidos, foi possível estabelecer uma relação entre o número de pulsos registrados pelo contador de baixo custo e a taxa de exposição equivalente para cada distância analisada. Essa relação é apresentada na Tabela III a seguir:

TABELA III
TABELA III – RELAÇÃO ENTRE CPM DO CONTADOR DE BAIXO CUSTO E TAXA DE EXPOSIÇÃO MEDIDA POR INSTRUMENTO CERTIFICADO.

Distância (cm)	CPM (Mediana)	mR/h
4	264,5	0,06
6	199,0	0,04
8	160,0	0,03
10	143,0	0,02
12	123,0	0,01
14	114,0	0,005

Caso se deseje estimar a taxa de exposição para uma distância específica, pode-se utilizar a relação média obtida a partir das medições realizadas nas distâncias estudadas, esse fator de conversão é apresentado na fórmula a seguir:

$$CPM \approx 5520 \cdot mR/h$$

Também é possível obter a taxa de dose absorvida do emissor a partir da fórmula:

$$\dot{D}_{ar} = \dot{X} \cdot (W/e)_{ar}$$

Em que $\dot{D}_{ar}$ representa a taxa de dose absorvida no ar, expressa em gray por hora ($\mathrm{Gy/h}$), $\dot{X}$ corresponde à taxa de exposição medida em roentgen por hora ($\mathrm{R/h}$), e $(W/e)_{ar}$ representa o fator de conversão entre a exposição e a dose absorvida no ar. Para o ar, essa constante resulta aproximadamente em $1\,\mathrm{R} = 8{,}77 \times 10^{-3}\,\mathrm{Gy}$. Assim, esse fator permite converter os valores de taxa de exposição obtidos experimentalmente em estimativas da taxa de dose absorvida no meio, conforme a tabela IV.

TABELA IV
DOSE ABSORVIDA ESTIMADA A PARTIR DA EXPOSIÇÃO MEDIDA

Distância (cm)	CPM (Mediana)	Exposição (mR/h)	Dose (mGy/h)
4	264,5	0,06	$5{,}26\times10^{-4}$
6	199,0	0,04	$3{,}50\times10^{-4}$
8	160,0	0,03	$2{,}63\times10^{-4}$
10	143,0	0,02	$1{,}75\times10^{-4}$
12	123,0	0,01	$8{,}76\times10^{-5}$
14	114,0	0,005	$4{,}38\times10^{-5}$

Além disso, considerando a emissão de radiação gama pela camisinha de lampião utilizada no experimento, pode-se assumir que o fator de ponderação da radiação é igual a 1. Dessa forma, a conversão entre taxa de dose absorvida e taxa de dose equivalente torna-se direta. Assim, a taxa de dose equivalente pode ser determinada a partir da relação da seguinte fórmula:

$$\dot{H} = \dot{D} \cdot w_R$$

Em que $\dot{H}$ representa a taxa de dose equivalente (Sv/h), $\dot{D}$ corresponde à taxa de dose absorvida (Gy/h) e w_R é o fator de ponderação da radiação, sendo 1 para a radiação gama emitida. O que nos leva para a tabela final (tabela V) que apresenta todas as relações estabelecidas.

TABELA V
EXPOSIÇÃO, DOSE ABSORVIDA E DOSE EQUIVALENTE ESTIMADAS A PARTIR DAS MEDIÇÕES EXPERIMENTAIS

Distância (cm)	CPM (Mediana)	Exposição (mR/h)	Dose absorvida (mGy/h)	Dose equivalente (mSv/h)
4	264,5	0,06	$5,26 \times 10^{-4}$	$5,26 \times 10^{-4}$
6	199,0	0,04	$3,50 \times 10^{-4}$	$3,50 \times 10^{-4}$
8	160,0	0,03	$2,63 \times 10^{-4}$	$2,63 \times 10^{-4}$
10	143,0	0,02	$1,75 \times 10^{-4}$	$1,75 \times 10^{-4}$
12	123,0	0,01	$8,76 \times 10^{-5}$	$8,76 \times 10^{-5}$
14	114,0	0,005	$4,38 \times 10^{-5}$	$4,38 \times 10^{-5}$

V.. CONCLUSÃO

Os resultados obtidos indicam que o contador GM de baixo custo desenvolvido apresentou comportamento consistente com o esperado teoricamente para a detecção de radiação ionizante. Tanto as medições realizadas com o equipamento desenvolvido quanto aquelas obtidas com o instrumento certificado apresentaram curvas de decaimento da intensidade da radiação em função da distância compatíveis com a relação do inverso do quadrado da distância amparada pela teoria.

A comparação direta entre os valores de contagem registrados pelo sistema de baixo custo e as taxas de exposição medidas pelo equipamento profissional permitiu estabelecer um fator médio de correlação entre CPM e mR/h. A partir dessa relação, torna-se possível estimar a taxa de exposição utilizando o contador desenvolvido com valores próximos aos obtidos por um instrumento certificado, considerando a

calibração definida a partir das medições experimentais com o equipamento produzido de baixo custo.

Do ponto de vista educacional e prático, o projeto cumpre integralmente seu objetivo ao demonstrar um excelente custo-benefício. Embora o sistema construído não substitua dosímetros comerciais em aplicações rigorosas de proteção radiológica hospitalar ou industrial, a precisão alcançada provou-se mais do que suficiente para a demonstração prática de fenômenos nucleares em laboratórios de graduação. O protótipo evidencia que é possível contornar a barreira financeira dos equipamentos importados, fornecendo às instituições de ensino uma ferramenta robusta e acessível. Assim, integrando conceitos de hardware, microcontroladores e física médica, o sistema viabiliza o ensino experimental de forma interativa e teoricamente amparada , validando o sucesso desta iniciativa desenvolvida no INATEL com o fomento da FAPEMIG e apoio técnico da clínica MagSul.

Agradecimentos

Gostaríamos de agradecer a Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG por financiar a pesquisa, a Clínica Magsul (Pouso Alegre/MG) pela disponibilidade do equipamento Geiger-Müller e pelo apoio técnico prestado, possibilitando a execução da etapa experimental desta pesquisa, e ao Instituto Nacional de Telecomunicações - INATEL por nos proporcionar a estrutura para a realização do projeto.

Referências

[1] James E. Turner. *Atoms, Radiation, and Radiation Protection*. 3ª ed. Wiley-VCH, 2007.

[2] Glenn F. Knoll. *Radiation Detection and Measurement*. 4ª ed. John Wiley & Sons, 2010.

[3] International Atomic Energy Agency. *Radiation Detection Instruments*. Vienna: IAEA, 2011.

[4] W. R. F. Silva e J. M. Fonseca. "Desenvolvimento de uma máquina de raios X, um detector Geiger e um goniômetro de baixo custo para estudo sobre raios X". Em: *Revista Brasileira de Ensino de Física* 45 (2023), e20230058.

[5] Luiz Tauhata, Ivan P. A. Salati, Renato Di Prinzio e Antonieta R. Di Prinzio. *Radioproteção e Dosimetria: Fundamentos*. 10ª ed. Instituto de Radioproteção e Dosimetria - Comissão Nacional de Energia Nuclear, 2014.

[6] Cholla Observatory Electronics. *M4011 Geiger-Müller Tube Datasheet*. 2013. URL: http://cholla.mmto.org/electronics/geiger/banggood/M4011.pdf (acesso em 04/03/2026).

[7] C. A. Boff, R. O. Bastos e F. L. Melquiades. "Práticas experimentais no ensino de física nuclear utilizando material de baixo custo". Em: *Caderno Brasileiro de Ensino de Física* 34.2 (2017), pp. 586–605.

[8] International Atomic Energy Agency. *Radiation Basics*. 2023. URL: https://www.iaea.org/topics/radiation-basics (acesso em 04/03/2026).

Autores



Brenno Cannellas é graduando em Engenharia Biomédica no Instituto Nacional de Telecomunicações (Inatel) e atua na área de Engenharia Clínica, com experiência em gestão e acompanhamento de tecnologias e equipamentos biomédicos em ambientes hospitalares. Possui interesse nas áreas de física médica, equipamentos de imagem e gestão de tecnologias em saúde.



Nicholas Lima do Nascimento é graduando em Engenharia de Software pelo INATEL – Instituto Nacional de Telecomunicações, bolsista da FAPEMIG para estudar e desenvolver um contador Geiger-Müller.



Roger Pereira Freitas é graduando em Engenharia de Software pelo INATEL – Instituto Nacional de Telecomunicações, bolsista da FAPEMIG para estudar e desenvolver um contador Geiger-Müller.

Orientador



Felipe Emanoel Chaves Possui graduação em Física pela Universidade Federal de Itajubá, possui mestrado em Física e Matemática Aplicada pela mesma instituição e doutorado em Matemática Pura pela Universidade Estadual de Campinas. Atualmente atua como professor no Instituto Nacional de Telecomunicações.

Co-Orientador



Rita Elizabeth Santos de Almeida Possui graduação em Engenharia Biomédica pelo Instituto Nacional de Telecomunicações - Inatel (2023). Atualmente é coordenadora de projetos no laboratório eHealth Innovation Center, centro de pesquisa de Engenharia Biomédica do Inatel.

Desenvolvimento e avaliação de adequação de prótese impressa ativa para animal de pequeno porte amputado

Lucas Toledo Capello, Maria Clara Bernardes Ferrer & Filipe Bueno Vilela
Instituto Nacional de Telecomunicações – Inatel
lucas.toledo@gec.inatel.br, maria.ferrer@geb.inatel.br, filipe.bueno@inatel.br

Abstract— *The work proposes the development of personalized prosthetics and orthoses for small and medium-sized dogs using 3D modeling technology. The devices are designed to fit the animal's anatomy and may include sensors to monitor vital data and electrical stimulation to aid muscle recovery. Initial tests showed positive results, with improved locomotion and more dynamic movement in the treated dog.*

Index Terms— *3D Modeling; Prothesis; Orthosis; Anima Rehabilitation.*

Resumo— O trabalho propõe o desenvolvimento de próteses e órteses personalizadas para cães de pequeno e médio porte utilizando tecnologia de modelagem 3D. Os dispositivos são projetados para se ajustar à anatomia do animal e podem incluir sensores para monitorar dados vitais e estimulação elétrica para auxiliar na recuperação muscular. Testes iniciais mostraram resultados positivos, com melhora na locomoção e movimentos mais dinâmicos no cão tratado.

Palavras-chave— Modelagem 3D; Próteses; Órteses; Reabilitação Animal.

I. Introdução

Atualmente, os animais de estimação desempenham um papel fundamental na vida das pessoas, ultrapassando a função tradicional de guardiões do lar para se tornarem membros essenciais da família. Os cães, em particular, são companheiros extremamente afetuosos, capazes de estabelecer vínculos profundos com seus tutores e oferecer suporte emocional significativo, muitas vezes contribuindo para o bem-estar psicológico e social das pessoas. Essa relação de proximidade tem impulsionado o desenvolvimento de soluções tecnológicas e terapêuticas voltadas à melhoria da qualidade de vida desses animais, especialmente em situações que envolvem limitações físicas decorrentes de acidentes, doenças ou envelhecimento. Considerando esse contexto e a necessidade de reabilitação de cães que apresentam limitações físicas devido a fraturas ou amputações, foram desenvolvidos dois dispositivos inovadores com o objetivo de auxiliar na recuperação da mobilidade e no tratamento da marcha desses animais. Esse avanço representa um importante marco na área veterinária, além de reforçar o papel da tecnologia no desenvolvimento de soluções que promovam maior bem-estar e autonomia para os animais de estimação [1, 2].

O estudo da locomoção e da biomecânica da marcha em cães é essencial para compreender como alterações estruturais ou funcionais podem afetar sua mobilidade e qualidade de vida. Pesquisas demonstram que mudanças na distribuição de peso, na postura e no padrão de movimento podem ocorrer após lesões ou amputações, exigindo intervenções adequadas para evitar sobrecarga em outras estruturas do corpo. Nesse sentido, a análise biomecânica da locomoção canina fornece subsídios importantes para o desenvolvimento de dispositivos de suporte, como próteses e órteses, que auxiliam na reabilitação funcional desses animais [3, 4].

Entretanto, uma adaptação muito rápida ou inadequada após lesões pode comprometer a recuperação do membro afetado e gerar problemas biomecânicos adicionais. Para garantir que os cães recuperem seus movimentos de forma adequada, evitando compensações prejudiciais ou sobrecarga em outras articulações, os tutores frequentemente recorrem ao uso de próteses ou órteses, dependendo das características e necessidades de cada caso. Esses dispositivos podem proporcionar suporte estrutural, estabilização articular e auxílio no processo de reabilitação motora, contribuindo para a melhoria da locomoção e da qualidade de vida dos animais [2, 3]. Estudos na área de cirurgia veterinária e reabilitação animal indicam que o uso adequado de próteses pode restaurar parcialmente a função locomotora e favorecer a adaptação do animal a novas condições biomecânicas [5].

Nos últimos anos, avanços tecnológicos têm ampliado as possibilidades de desenvolvimento de dispositivos ortopédicos personalizados. O uso de técnicas de modelagem tridimensional e impressão 3D tem se destacado como uma ferramenta promissora na medicina veterinária, permitindo a produção de próteses e órteses adaptadas às características anatômicas específicas de cada animal. Essa personalização contribui para um melhor ajuste do dispositivo ao corpo do paciente, aumentando o conforto, a eficiência funcional e a aceitação do equipamento durante o processo de reabilitação [6]. Além disso, o desenvolvimento de dispositivos ortopédicos personalizados pode facilitar a integração de tecnologias adicionais, como sensores e sistemas de monitoramento clínico, ampliando as possibilidades de acompanhamento da recuperação do animal [1, 3].

Diante desse cenário, o presente projeto tem como objetivo desenvolver uma prótese e uma órtese destinadas a cães de pequeno e médio porte. Além de promover maior liberdade de movimento para esses animais, o projeto também contempla o monitoramento da marcha, a avaliação do estado clínico e a oferta de um tratamento ortopédico mais adequado. A integração de sensores ao sistema permite acompanhar parâmetros relevantes do processo de reabilitação, possibilitando a coleta de dados relacionados ao padrão de locomoção e ao estado fisiológico do animal. Tais abordagens estão alinhadas com avanços recentes na área de biomecânica e reabilitação, nos quais dispositivos assistivos podem contribuir significativamente para a recuperação funcional e para a adaptação do organismo a novas condições motoras [7, 8].

Além disso, a evolução da área de órteses e próteses veterinárias tem demonstrado que o desenvolvimento de soluções tecnológicas adaptadas à realidade clínica pode ampliar significativamente as possibilidades de tratamento em medicina veterinária. Dispositivos ortopédicos modernos podem não apenas restaurar parcialmente a função locomotora, mas também atuar no processo de reabilitação terapêutica, auxiliando na correção de padrões de movimento e na prevenção de complicações musculoesqueléticas [9, 10].

A presente pesquisa propõe o desenvolvimento e a avaliação experimental de dispositivos ortopédicos vestíveis baseados em manufatura aditiva e sistemas embarcados de monitoramento, destinados a aplicações em reabilitação locomotora de cães de pequeno e médio porte. A abordagem adotada combina modelagem tridimensional assistida por computador, impressão 3D e integração de sensores biomédicos com o objetivo de criar protótipos de próteses e órteses capazes de monitorar parâmetros biomecânicos e fisiológicos associados à locomoção. Nesse contexto, o estudo busca investigar a viabilidade técnica da incorporação de sensores de força,

temperatura e pressão em dispositivos ortopédicos, bem como a integração de um módulo de eletroestimulação voltado ao suporte de processos de recuperação musculoesquelética. Dessa forma, o trabalho tem como objetivo principal demonstrar o potencial da integração entre tecnologias de manufatura aditiva e sistemas eletrônicos embarcados para o desenvolvimento de soluções inovadoras de monitoramento e suporte terapêutico em dispositivos ortopédicos veterinários.

II. Metodologia

A. Desenvolvimento do dispositivo vestível

O primeiro dispositivo desenvolvido consiste em uma órtese produzida por meio de tecnologia de manufatura aditiva (impressão 3D), projetada em colaboração com um profissional da medicina veterinária. O projeto foi concebido com base em um cenário clínico de referência, representativo de um cão da espécie Canis lupus familiaris, da raça Pastor Alemão, com aproximadamente quatro anos de idade, submetido hipoteticamente a um procedimento cirúrgico para implantação de uma prótese óssea de titânio após trauma ortopédico [11]. Em situações clínicas dessa natureza, é comum observar redução da força muscular e alterações na distribuição de carga no membro afetado. Assim, o desenvolvimento da órtese teve como objetivo fornecer suporte biomecânico adicional ao membro posterior, além de contribuir para a redistribuição adequada do peso corporal durante a locomoção.

O segundo dispositivo desenvolvido corresponde a uma prótese destinada a simular a substituição funcional de um membro torácico amputado. Para fins de projeto, considerou-se um cenário hipotético de amputação traumática, frequentemente relatado na literatura veterinária, no qual um cão sem raça definida apresenta perda parcial de um membro anterior. Em tais condições, alterações significativas na biomecânica da marcha podem ocorrer, resultando em aumento da demanda mecânica sobre os membros remanescentes e sobre a coluna vertebral.

Para a definição de parâmetros geométricos e biomecânicos de referência, foi utilizado como modelo anatômico um cão da espécie Canis lupus familiaris, da raça Yorkshire Terrier, selecionado por representar um animal de pequeno porte com proporções corporais bem documentadas em bases de dados morfológicos [12]. A escolha desse modelo permitiu realizar um estudo teórico da biomecânica da locomoção em um sistema locomotor íntegro, possibilitando a extrapolação de parâmetros para o desenvolvimento do protótipo protético.

Durante a etapa de aquisição de medidas anatômicas de referência, algumas limitações metodológicas foram identificadas. O uso de tecnologias avançadas de digitalização tridimensional, como scanners infravermelhos ou sistemas de captura óptica, poderia aumentar significativamente a precisão das medições e otimizar o processo de desenvolvimento das estruturas. No entanto, tais tecnologias apresentam custo elevado e requerem infraestrutura laboratorial específica. Considerando essas limitações, optou-se pela utilização de medições manuais baseadas em proporções anatômicas descritas na literatura veterinária e em modelos de referência, permitindo a obtenção de parâmetros dimensionais adequados para o desenvolvimento inicial dos protótipos.

B. Construção das estruturas

As dimensões utilizadas no projeto foram definidas a partir de medidas anatômicas de referência, permitindo o desenvolvimento de estruturas compatíveis com a morfologia canina. A modelagem tridimensional das peças foi realizada utilizando o software SolidWorks, amplamente empregado em projetos de engenharia mecânica e biomédica.

O projeto estrutural foi desenvolvido considerando aspectos fundamentais da anatomia e biomecânica canina, incluindo a distribuição de carga nos membros e a necessidade de ventilação da pele em dispositivos de contato prolongado. Dessa forma, foram incorporadas aberturas estruturais ao longo da órtese e da prótese, com o objetivo de favorecer a ventilação e reduzir o acúmulo de umidade na interface entre o dispositivo e o tecido cutâneo.

Além disso, foram projetados canais específicos para a passagem de cabos associados aos sensores eletrônicos, reduzindo a possibilidade de deslocamento ou interferência mecânica nos componentes eletrônicos durante o uso do dispositivo.

Para a fabricação dos protótipos foi selecionado o polímero ABS (Acrilonitrila Butadieno Estireno), material amplamente utilizado em impressão 3D devido à sua elevada resistência mecânica, boa estabilidade dimensional e adequado custo-benefício. O processo de impressão foi realizado utilizando parâmetros técnicos intermediários, incluindo duas camadas de parede e densidade de preenchimento de 50%, buscando equilibrar resistência estrutural e redução de massa do dispositivo. [3] [4]

C. Sensores e sistema de eletroestimulação

Para o desenvolvimento do protótipo protético foi utilizada uma plataforma baseada no microcontrolador Arduino Mega 2560, associada a um sensor resistivo de força FSF400, conectores do tipo jumper e uma camada interna de material amortecedor composta por espuma de EVA (etileno-vinil-acetato) com espessura de 5 mm. Esse material foi posicionado na região interna superior da prótese com a finalidade de proporcionar maior conforto mecânico e reduzir concentrações localizadas de pressão.

A arquitetura mecânica do dispositivo foi projetada em duas seções principais: a primeira destinada ao encaixe do membro residual e a segunda responsável por atuar como elemento estrutural de suporte durante a locomoção.

O sistema eletrônico embarcado foi programado de forma que os sinais provenientes dos sensores fossem adquiridos através da porta analógica A0 do microcontrolador. Os dados coletados são transmitidos em tempo real para um computador, onde são armazenados em planilhas do Microsoft Excel para posterior análise.

Após a integração do hardware e do software embarcado, foram realizados testes experimentais de validação do sistema. Entre os procedimentos realizados destacam-se aqueles listados a seguir.

C. 1. Teste estático

Realizado para avaliar a linearidade e a curva de resposta do sensor resistivo de força. Para esse teste, uma carga padrão de aproximadamente 1 kg foi aplicada sobre o dispositivo durante um período de cerca de dois minutos.

C.2. Teste de resistência mecânica

Foram conduzidos testes estruturais simulando condições de uso, incluindo aplicação de força direta, impacto por queda controlada e compressão mecânica sobre o protótipo.

D. Processamento de Dados

Foi utilizada a ferramenta Data Streamer, disponibilizada pelo próprio software, permitindo a geração automática de planilhas e gráficos de linha para análise preliminar dos dados.

No desenvolvimento da órtese, foram incorporados sensores não invasivos destinados ao monitoramento indireto de parâmetros fisiológicos e biomecânicos. Entre os sensores utilizados destaca-se um sensor infravermelho de pulso, capaz de estimar variações na frequência cardíaca, possibilitando inferências sobre o nível de estresse ou adaptação ao dispositivo.

Adicionalmente, foi incorporado um sensor de temperatura LM35, com o objetivo de monitorar possíveis variações térmicas que poderiam indicar processos inflamatórios ou infecciosos em condições clínicas reais.

Para a análise da dinâmica da marcha, foram empregados dois sensores resistivos de força FSR-400, posicionados de forma a

permitir o monitoramento da distribuição de carga durante o contato com o solo. Esses sensores possibilitam a obtenção de dados relacionados ao padrão de locomoção e à transferência de peso entre os membros.

Além do sistema de monitoramento, foi desenvolvido um módulo de eletroestimulação, composto por um fio condutor de cobre com diâmetro de 16 mm associado a um transistor TIP122, destinado a aplicações experimentais de estimulação muscular. Esse tipo de abordagem é frequentemente investigado em estudos de reabilitação musculoesquelética, podendo contribuir para a prevenção de atrofia muscular e para a recuperação funcional em cenários clínicos hipotéticos.

Todos os sensores foram integrados a uma placa Arduino Uno, responsável pela centralização, aquisição e processamento dos sinais coletados.

III. Resultados

O protótipo de prótese foi desenvolvido considerando um cenário de aplicação em cães de porte médio, utilizando parâmetros morfológicos e biomecânicos descritos na literatura veterinária como referência para o dimensionamento estrutural. O projeto foi concebido com o objetivo de garantir simultaneamente funcionalidade mecânica, estabilidade estrutural e conforto na interface entre o dispositivo e o membro. A arquitetura da prótese foi organizada em duas estruturas principais complementares. A porção superior foi projetada para acomodar a região de contato com o membro, enquanto a porção inferior atua como elemento de suporte responsável pela transferência de carga durante o contato com o solo. Na estrutura superior foi integrado um sensor de força estrategicamente posicionado, associado a um inserto amortecedor confeccionado em EVA (etileno-vinil-acetato), material escolhido devido às suas propriedades de absorção de impacto e capacidade de redistribuição de pressão, reduzindo potenciais pontos de concentração de carga na interface do dispositivo.

A estrutura inferior da prótese foi projetada com dois orifícios de ajuste que permitem a modificação da altura do dispositivo. Esse mecanismo de regulagem possibilita adaptar o comprimento funcional da prótese a diferentes proporções corporais, além de permitir ajustes durante etapas experimentais de calibração e validação biomecânica. O projeto tridimensional da prótese foi desenvolvido em ambiente de modelagem CAD, permitindo a criação de um modelo paramétrico do dispositivo, e pode ser visto na Fig. 1. Essa abordagem possibilitou realizar múltiplas iterações no projeto antes da etapa de fabricação, otimizando aspectos relacionados à distribuição de massa, posicionamento dos sensores, estabilidade estrutural e ergonomia da interface com o membro. A modelagem também permitiu avaliar previamente regiões de possível concentração de tensões, possibilitando ajustes na espessura das paredes estruturais e na geometria interna das peças. Como resultado, foi possível obter um modelo estrutural equilibrado, com massa reduzida e resistência mecânica adequada para suportar cargas simuladas associadas à locomoção.

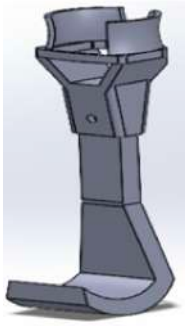


Fig. 1. Modelagem 3D com as partes superior e inferior conectadas.

Após a finalização do modelo tridimensional, os componentes foram fabricados por meio de tecnologia de manufatura aditiva baseada em deposição de material fundido (FDM). A utilização da impressão 3D permitiu a produção rápida de protótipos funcionais, facilitando ajustes iterativos no projeto e reduzindo o tempo total de desenvolvimento. Durante o processo de fabricação observou-se a necessidade de uma proteção do sistema principal para garantir que o animal não interaja diretamente com o suporte mecânica, evitando degradação. O material polimérico utilizado apresentou resistência mecânica adequada para suportar testes preliminares de carga e manipulação sem ocorrência de falhas estruturais significativas. A manufatura aditiva também possibilitou incorporar diretamente no projeto elementos estruturais complexos, como cavidades internas, canais para passagem de cabos e compartimentos específicos para a instalação dos sensores, características que seriam mais difíceis de implementar em processos convencionais de fabricação. A prótese impressa pode ser vista na Fig. 2.



Fig. 2. Prótese impressa com proteção à estrutura mecânica principal.

Para avaliar o desempenho do sensor de força incorporado ao dispositivo foram conduzidos três experimentos de validação em condições controladas. Esses testes tiveram como objetivo analisar a resposta do sensor sob diferentes condições de aplicação de carga. Os resultados foram representados em um gráfico no qual o eixo vertical corresponde à intensidade da força aplicada ao sensor, expressa em unidades digitais provenientes da leitura analógica do microcontrolador, enquanto o eixo horizontal representa o tempo de duração do experimento, expresso em segundos. A análise dos dados indicou comportamento consistente do sensor ao longo do período de teste, demonstrando capacidade de registrar variações de carga de maneira estável. Esse comportamento é compatível com a resposta esperada para sensores resistivos de força utilizados em aplicações biomecânicas.

Além da prótese, foi desenvolvido um modelo de órtese considerando como referência morfológica um cão de grande porte com massa corporal aproximada de 50 kg, valor frequentemente associado a exemplares da raça Pastor Alemão descritos em estudos anatômicos veterinários. Devido às dimensões relativamente grandes do dispositivo, a órtese foi dividida em múltiplos segmentos estruturais para permitir sua fabricação por impressão 3D dentro das limitações dimensionais da impressora utilizada. A porção inferior da órtese apresenta aproximadamente 10 cm de largura por 15 cm de altura, enquanto a porção superior, composta por seções anterior e posterior, possui aproximadamente 15 cm de largura por 15 cm de altura. O projeto estrutural inclui múltiplas aberturas distribuídas ao longo da superfície da peça, com o objetivo de favorecer a ventilação e reduzir o acúmulo de calor e umidade na interface entre o dispositivo e a pele.

Para aplicações clínicas hipotéticas, recomenda-se o uso de materiais acolchoados na região de contato entre a órtese e o membro, com a finalidade de reduzir o atrito mecânico e aumentar o

conforto durante o uso prolongado. Esse tipo de interface amortecedora contribui para minimizar pontos de pressão e possíveis irritações cutâneas, fator frequentemente considerado no desenvolvimento de dispositivos ortopédicos externos.

A incorporação de sensores ao dispositivo representa um avanço relevante no desenvolvimento de órteses e próteses inteligentes. A utilização de sensores de força permite monitorar a distribuição de carga durante o contato com o solo, fornecendo dados importantes sobre o padrão de marcha e a dinâmica de transferência de peso. Essas informações podem ser utilizadas para avaliar o desempenho biomecânico do dispositivo e orientar ajustes estruturais no protótipo. Além disso, a integração de sensores fisiológicos, como sensores de temperatura e sensores ópticos de pulso, amplia as possibilidades de monitoramento do estado fisiológico durante o uso do dispositivo. Essa abordagem possibilita a aquisição contínua de dados, contribuindo para o desenvolvimento de sistemas de monitoramento remoto e análise biomecânica aplicada à reabilitação locomotora. A Fig. 3. é a representação gráfica dos resultados obtidos pelos sensores de força, demonstrando estabilidade na resposta durante o uso pelo animal, simulando uma carga de 1 kg.

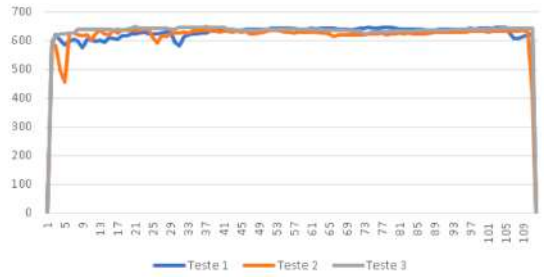


Fig. 3. Resposta do sensor de força com simulação de carga de 1 kg.

Dessa forma, a combinação entre manufatura aditiva, modelagem tridimensional e integração de sensores eletrônicos demonstra potencial significativo para o desenvolvimento de dispositivos ortopédicos personalizados. Essa abordagem tecnológica permite a criação de protótipos adaptáveis a diferentes configurações morfológicas, além de fornecer ferramentas adicionais para o acompanhamento biomecânico e funcional associado ao uso de próteses e órteses em aplicações veterinárias.

IV. Discussão

O objetivo inicial deste projeto foi desenvolver um dispositivo vestível destinado ao monitoramento da marcha e ao suporte ao tratamento ortopédico de cães de pequeno e médio porte. Nesse contexto, foi projetada uma prótese equipada com sensores de força, bem como uma órtese integrada a sensores de temperatura, força e pressão, além de um módulo de eletroestimulação acoplado. Esse sistema foi concebido com a finalidade de potencialmente auxiliar processos de regeneração óssea e contribuir para o tratamento de lesões musculares associadas à atrofia muscular, em cenários clínicos de reabilitação.

Em dispositivos vestíveis voltados ao monitoramento biomecânico, é fundamental que os sensores incorporados apresentem elevada confiabilidade e precisão nas medições. A análise do gráfico obtido a partir do teste estático indica que o sensor de força apresentou leituras consistentes, com variações mínimas tanto ao longo do tempo quanto sob aplicação de uma carga constante de aproximadamente 1 kg. Esses resultados sugerem que o sensor apresenta comportamento estável e nível de precisão adequado, demonstrando sua viabilidade para aplicações de monitoramento de carga em dispositivos ortopédicos experimentais.

O protótipo da prótese também foi submetido a testes de resistência mecânica, nos quais demonstrou desempenho estrutural satisfatório. Os resultados obtidos indicam que o material selecionado para a impressão tridimensional apresenta propriedades mecânicas compatíveis com as exigências do dispositivo, reforçando sua adequação para utilização no desenvolvimento do projeto.

No caso da órtese, análises preliminares indicam que a aplicação do dispositivo pode contribuir para a melhoria da estabilidade locomotora em cenários simulados de suporte ortopédico. A incorporação de sensores ao sistema permite, adicionalmente, o monitoramento contínuo de parâmetros biomecânicos e fisiológicos, possibilitando avaliar a adaptação ao dispositivo e acompanhar indicadores associados ao estado funcional durante seu uso.

Para aprimorar o processo de aquisição e análise de dados, considera-se relevante a incorporação de módulos de comunicação sem fio, como Bluetooth ou Wi-Fi, ao microcontrolador Arduino. A integração desses módulos permitiria a transmissão dos dados dos sensores em tempo real para sistemas externos de processamento, possibilitando a implementação de estratégias de monitoramento contínuo e análise remota das informações coletadas.

Em relação aos componentes estruturais do dispositivo, observou-se a necessidade de ajustes no dimensionamento das peças, particularmente no aumento da espessura de determinadas regiões estruturais. Essa modificação visa melhorar a resistência mecânica e a durabilidade do dispositivo, considerando as cargas dinâmicas associadas ao contato frequente com o solo e às condições de uso em ambientes externos. Dessa forma, tais ajustes estruturais contribuem para aumentar a robustez do sistema e ampliar sua confiabilidade em aplicações experimentais de dispositivos ortopédicos vestíveis.

V. Conclusão

Os resultados obtidos ao longo do desenvolvimento do projeto demonstram a viabilidade da utilização de tecnologias de manufatura aditiva para a produção de dispositivos ortopédicos personalizados voltados à reabilitação locomotora em contextos veterinários. A modelagem tridimensional permitiu o desenvolvimento de estruturas adaptáveis a diferentes configurações morfológicas, enquanto a fabricação por impressão 3D possibilitou a produção rápida de protótipos com boa fidelidade geométrica e resistência mecânica satisfatória. Os testes realizados indicaram que os sensores integrados ao sistema apresentaram comportamento consistente durante as etapas de validação, evidenciando sua adequação para aplicações de monitoramento biomecânico em dispositivos vestíveis.

Além disso, a integração de sensores fisiológicos e biomecânicos ao dispositivo amplia significativamente as possibilidades de monitoramento contínuo de parâmetros relevantes para o acompanhamento funcional durante processos de reabilitação. A adoção de sistemas de aquisição e transmissão de dados baseados em microcontroladores possibilita a implementação de estratégias de monitoramento remoto e análise em tempo real, contribuindo para o avanço de abordagens tecnológicas aplicadas à medicina veterinária. Nesse sentido, a combinação entre modelagem digital, manufatura aditiva e sistemas eletrônicos embarcados apresenta grande potencial para o desenvolvimento de dispositivos ortopédicos inteligentes, capazes de fornecer suporte mecânico e simultaneamente gerar dados relevantes para avaliação biomecânica e acompanhamento terapêutico.

IV. Agradecimentos

Os autores agradecem ao Instituto Nacional de Telecomunicações (INATEL) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio e fomento concedidos a este projeto de pesquisa.

Referências

1. FISCHER, M. S.; LILJE, K. E. Dogs in Motion. Jena: VDH Service GmbH, 2011.
(Livro bastante usado sobre biomecânica e marcha em cães.)
2. KIRPENSTEIJN, J.; HAAN, J.; VAN DER VELDEN, M. Use of limb prostheses in small animals. Veterinary Surgery, v. 28, n. 6, p. 453–457, 1999.

3. CARR, B. J.; DODD, C. A. F. Advances in veterinary orthotics and prosthetics. Veterinary Clinics of North America: Small Animal Practice, v. 46, n. 6, p. 1089–1103, 2016.
4. BINI, R. R.; HUME, P. A.; CROFT, J. L. Effects of lower-limb prostheses on gait biomechanics. Journal of Biomechanics, v. 47, n. 11, p. 2623–2630, 2014. (Embora humano, é muito citado em estudos de biomecânica aplicados.)
5. ZLOTNICKA, B.; et al. 3D printing in veterinary medicine: applications in prosthetics and orthotics. Animals, v. 11, 2021.
6. BUDSBERG, S. C.; TORRES, B. T. Veterinary rehabilitation and physical therapy. Veterinary Clinics of North America: Small Animal Practice, v. 46, n. 1, p. 1–14, 2016.
7. VOSS, K.; DAMUR, D.; GUERRERO, T.; HAESSIG, M.; MONTAVON, P. M. Force plate gait analysis in dogs: a review of methodological issues. Veterinary and Comparative Orthopaedics and Traumatology, v. 20, n. 3, p. 161–171, 2007.
8. BERGH, M. S.; BUDSBERG, S. C. The coxofemoral joint: normal anatomy, function, and biomechanics. Veterinary Surgery, v. 43, n. 1, p. 2–12, 2014.
9. VENTOLA, C. L. Medical applications for 3D printing: current and projected uses. Pharmacy and Therapeutics, v. 39, n. 10, p. 704–711, 2014.
10. HOFFMANN, D. E.; MASON, J. J. Advances in additive manufacturing for orthopedic prosthetics. Journal of Orthopaedic Research, v. 35, n. 4, p. 699–705, 2017.
11. RAJPUT, A.; CHAURASIA, D.; INGOLE, S. P.; DESHMUKH, S. K.; BARIK, A.; TIWARI, P. Comparative morphometric analysis of body and skull parameters in German Shepherd, Golden Retriever, and Siberian Husky dogs. *Journal of Advances in Biology & Biotechnology*, v. 28, n. 7, p. 46–53, 2025.
12. HADDAM, H. Y.; MENNANI, A.; BELALA, R.; et al. Statistical models based on morphometric traits for body weight prediction in four dog breeds raised in Algeria. *Genetics & Biodiversity Journal*, 2024.

AUTORES

Lucas Toledo Capello é ex-aluno do curso de Engenharia de Computação do Inatel. Foi pesquisador de iniciação cientifica em 2023 e 2024 no eHealth Innovation Center.

Maria Clara Bernardes Ferrer é graduanda em Engenharia Biomédica pelo Inatel. É pesquisadora de iniciação cientifica desde 2024 no eHealth Innovation Center.



Filipe Bueno Vilela é engenheiro biomédico pelo Instituto Nacional de Telecomunicações (INATEL). Possui mestrado e doutorado em Engenharia de Materiais, com ênfase em biomateriais, pela Universidade Federal de Itajubá (UNIFEI). Atualmente, atua como professor, pesquisador e desenvolvedor no INATEL.

Mechanical Stimulation System Prototype for Bronchial Secretion Clearance in Preterm Infants

Mateo Sebastian Rojas Flores, Eynar Calle Viles & Tatiana Torrico Sanchez
Universidad Privada del Valle – UNIVALLE
mateo02rojas09@gmail.com, ecallevi@univalle.edu, ttorricos@univalle.edu
ORCID: 0009-0002-3958-6814, 0000-0003-1477-2813

Abstract— Neonatal prematurity involves physiological immaturity that impairs bronchial secretion clearance, leading to a higher risk of infections and severe respiratory complications. Standard manual respiratory physiotherapy is limited by application inconsistency and high staffing requirements. This project focuses on the development of a mechanical stimulation system based on High-Frequency Chest Compression (HFCC) technology, specifically designed for neonatal clinical settings. Results indicate that the integrated Python-based GUI allows for intuitive, real-time adjustments of therapy parameters without latency, offering a viable path toward reducing human resource demands in neonatal care.

Index Terms— Airway clearance, HFCC, mechanical stimulation, preterm infants, respiratory therapy.

I. Introduction

Prematurity, defined as birth before 37 weeks of gestation, is considered a leading cause of neonatal and infant mortality. This is primarily due to the underdeveloped anatomical structure of preterm neonates, which results in significant physiological limitations. These conditions increase the likelihood of various pathologies, most notably the accumulation of bronchial secretions, which lead to infections and severe respiratory complications. To address this, physical therapy professionals have developed neonatal respiratory physiotherapy methods aimed at assisting the clearance of accumulated secretions. [1]

Manual respiratory physiotherapy involves techniques such as chest percussion and vibration. While effective, these methods are highly dependent on the practitioner's skill and can lead to inconsistent application. Furthermore, in extremely low-birth-weight infants, excessive or unregulated manual pressure can be harmful, potentially causing rib fractures or fluctuations in intracranial pressure. While these manual methods have proven effective, they possess inherent drawbacks, such as high demand for human resources and inter-operator variability which complicates clinical supervision. [2]

Technological advancements have led to the development of respiratory physiotherapy equipment that can be operated intuitively while providing precise treatment. Notably, systems based on the High-Frequency Chest Compression (HFCC) principle allow for bronchial secretion mobilization through direct mechanical stimulation, clearing the airways and improving respiratory function [3]. Although the use of such equipment is viable in older infants, this technology has yet to be adapted for the neonatal field, where it would be highly beneficial. Considering the patient's physiology, neonates present a fragile thoracic cage and a higher respiratory rate, among other factors that differentiate them from older infants. Additionally, preterm neonates are typically kept in incubators where space is extremely limited.

The main contribution of this work is the development of a low-cost, modular mechanical stimulation prototype that utilizes localized electronic actuators rather than traditional pneumatic systems. Unlike existing HFCC-based devices that rely on bulky compressors and expensive inflatable vests, this system allows for targeted, multi-channel stimulation specifically scaled for the limited space of neonatal incubators.

II. Theoretical Framework

A. Bronchial Secretions

Mucus is a viscoelastic gel that lines the surface of the respiratory tract, composed of water, mucins, antimicrobial and immunomodulatory molecules, and lipids. Its primary function is to trap inhaled particulate matter, cellular debris, and microbes, transporting them out of the lungs through ciliary beating and coughing. The clearance of secretions from the lower respiratory tract occurs through two main mechanisms: Mucociliary clearance involves a complex coordination of ciliary beat activity, which propels mucus proximally toward the mouth until it is expectorated or swallowed. When ciliary clearance fails to remove these secretions, the cough mechanism is utilized to clear the central airways. [4]

Mucus hypersecretion and alterations in its biophysical properties can impair airway clearance, leading to airflow obstruction, atelectasis, and hypoxemia. Furthermore, it increases the risk of respiratory tract infections and, if persistent, may cause permanent airway damage manifesting as bronchiectasis. Both infection and inflammation stimulate further mucus secretion, resulting in a vicious cycle if not properly addressed.

In preterm neonates, mucus accumulation occurs primarily due to a combination of immature ciliary activity, a narrow airway diameter that increases resistance, and a weak cough reflex that is insufficient to expel viscoelastic mucus.

B. Neonatal Respiratory Physiotherapy

The primary goals of neonatal respiratory physiotherapy focus on preserving and optimizing pulmonary function within an immature and vulnerable respiratory system. Key objectives include maintaining airway patency, facilitating the clearance of accumulated secretions, re-expanding collapsed lung

segments, and preventing respiratory complications such as atelectasis, infections, or secretion retention. Furthermore, it aims to ensure adequate oxygenation levels and contribute to the overall enhancement of the neonate's respiratory function, promoting a faster and more stable recovery. [5]

Preterm neonates possess unique anatomical and physiological characteristics that make them particularly susceptible to respiratory complications. Consequently, respiratory physiotherapy in preterm infants must be carefully adapted, utilizing gentle techniques and continuously monitoring the patient's response to prevent adverse effects.

Due to the characteristic thoracic fragility and high chest wall compliance of these patients, the pressure applied to the chest must be strictly regulated. Evidence suggests that timely and appropriate interventions can significantly improve oxygenation and reduce associated morbidity. [6]

C. Airway Clearance Devices

A variety of Airway Clearance Devices (ACDs) have been developed to automate manual procedures and mitigate the fatigue inherent in conventional techniques. Furthermore, studies indicate that the use of ACDs can achieve a 1.8-fold increase in secretion clearance per session compared to standard physiotherapy. Among the various types of ACDs, those operating on the principle of High-Frequency Chest Compression (HFCC) are particularly prominent. [3]

The operational principle of HFCC is based on generating vibrations across the chest wall to clear secretions. This method offers significant advantages over conventional physiotherapy in terms of expectoration and respiratory function; it also shows positive effects on Forced Vital Capacity (FVC) and Forced Expiratory Volume (FEV). HFCC equipment can be categorized as follows:

Mechano-Acoustic: Where stimulation is applied to a single side (anterior or posterior) of the chest wall. High-Frequency Chest Wall Oscillation (HFCWO): Which typically utilizes an inflatable vest and a compressor system. High-Frequency Chest Wall Compression (HFCWC): Characterized by generating oscillations on both sides of the thoracic wall. [7]

These devices are generally used as an adjunct to conventional respiratory physiotherapy.

III. Proposed System Design

A. Hardware Configuration

The system's hardware is designed for modularity and rapid prototyping. The control core utilizes an ESP32 MCU, this microcontroller establishes a bidirectional communication link with a laptop via a USB-serial interface, enabling real-time control through a dedicated Graphical User Interface (GUI) made with Python.

The architecture features four primary output channels, each of these four channels is comprised of a dual-output stage: A dedicated pulse-modulated output for a vibration motor; A synchronized auxiliary output for an LED indicator, providing real-time visual feedback on the activation status and rhythm of each specific channel.

The output channel (see Figure 1) employs a transistor-based switching configuration. The ESP32 digital outputs are connected through a current-limiting resistor (R1) to the base of NPN transistors, which act as electronic switches to manage the higher current demands of the motors. It also integrates another current-limiting resistor (R2) to provide supply voltage to an indicator LED. To ensure system stability and protect the microcontroller from inductive spikes, each motor circuit incorporates a flyback diode and a decoupling capacitor to mitigate back-EMF (electromotive force) and electrical noise.

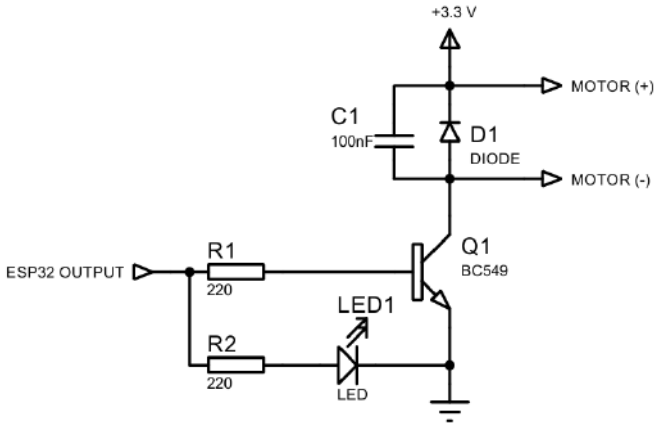


Figure 1. Electronic schematic of the output channel.

B. Signal Synthesis

The signal generation core is implemented utilizing a non-blocking timing architecture to ensure real-time responsiveness. The synthesis of mechanical stimulation patterns is achieved through a dynamic Pulse Width Modulation (PWM) mapping technique, where the duty cycle is updated at every iteration of the main execution loop.

The system modulates the output amplitude using the MAX_PWM parameter. This variable serves as a primary control for motor vibration intensity, allowing for the customization of mechanical stimulation based on the specific clinical needs of the preterm infant.

To maintain precise frequency control without the jitter associated with standard delay() functions, the system utilizes the internal millis() counter. The current phase (ϕ) of the waveform, normalized between 0.0 and 1.0, is calculated using the fmod function, that returns the floating-point remainder of a division, and it can be seen as:

$$\phi = \text{fmod}(\frac{t}{T}, 1.0) \quad (1)$$

Where t represents the elapsed time and T is the period, defined by the frequency, that is a user-selected parameter.

The phase value is used to reconstruct four distinct geometric signals. The algorithms for signal reconstruction are defined as follows:

Square Wave: A discrete state transition.

$$PWM = \begin{cases} MAX\_PWM & \text{if } (\phi < 0.5) \\ 0 & \text{if } (\phi \geq 0.5) \end{cases} \quad (2)$$

The intensity is maintained at the maximum level while the current phase is less than 0.5 and remains at zero during the second half of the cycle, providing a high-impact percussive effect.

Triangular Wave: A symmetrical rise and fall.

$$PWM = \begin{cases} 2 \cdot MAX\_PWM \cdot \phi & \text{if } (\phi < 0.5) \\ 2 \cdot MAX\_PWM \cdot (1 - \phi) & \text{if } (\phi \geq 0.5) \end{cases} \quad (3)$$

The value increases linearly during the first half of the cycle and decreases with an inverted slope during the second half, resulting in a more gradual mechanical transition.

Sinusoidal Wave: A continuous trigonometric function.

$$PWM = \frac{MAX\_PWM}{2} \cdot (\sin(2\pi\phi) + 1.0) \quad (4)$$

This continuous trigonometric function is centered at a DC offset equal to half of the MAX_PWM. This allows the microcontroller, which only operates with positive values, to generate a fluid, oscillating pattern that mimics natural physiological rhythms.

Sawtooth Wave: A linear ramp function defined by:

$$PWM = MAX\_PWM \cdot \phi \quad (5)$$

This mode is defined by a continuous linear ramp where the intensity increases proportionally with the current phase throughout the entire cycle. It resets sharply to zero at the start of each new period, creating a specific percussive ramp-up.

Finally, the proposed system design is shown in Figure 2.

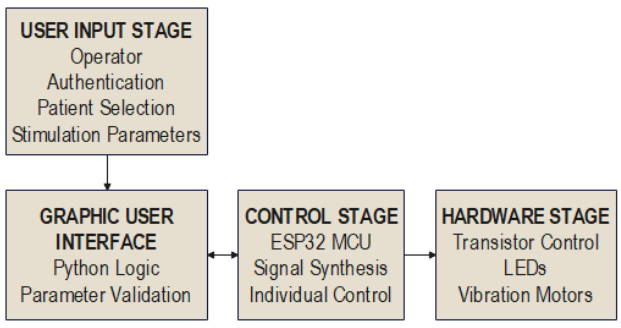


Figure 2. Schematic of the Proposed System Design.

IV. Preliminary Results

The preliminary phase of this research focused on the functional integration of the hardware components and the validation of the control software. At this stage, the system has successfully demonstrated the capability to generate controlled mechanical oscillations within the specified ranges.

A. Graphic User Interface (GUI) Implementation

The control layer was developed in Python using the Tkinter library, prioritized for its cross-platform portability, which is critical for real-time medical monitoring. To ensure clinical traceability and patient safety, the system is integrated with a local SQLite database, which maintains a secure record of authorized operators and specific patient therapy histories. The GUI follows a two-stage operational workflow:

Session Initialization: A dedicated welcome screen facilitates operator authentication. This ensures that every therapy session is linked to a specific health professional, allowing for accountability within the clinical environment.

Therapy Configuration: The main control panel enables the selection of a specific neonate and the independent configuration of four distinct stimulation channels. As seen in Figure 3, each channel provides control over waveform geometry, frequency (adjustable in 0.1 Hz increments), and a dedicated intensity slider.



Figure 3. Capture of the Python GUI.

The integration of a centralized count-down timer ensures that treatment durations remain consistent with clinical protocols. Communication with the hardware core is established via a high-speed serial interface, enabling the operator to adjust the stimulation parameters without introducing latency into the mechanical output.

B. Hardware Integration

The physical prototype was assembled on a compact perforated prototype board (7 x 9 cm) to evaluate the footprint required for incubator compatibility. The switching stage effectively managed the current requirements of the actuators, and the limit of 3.3 V was verified using a multimeter to ensure

compliance with the design constraints. The physical board is shown in Figure 4. This board size allows the implementation of another 4 channels to give more personalized treatment.

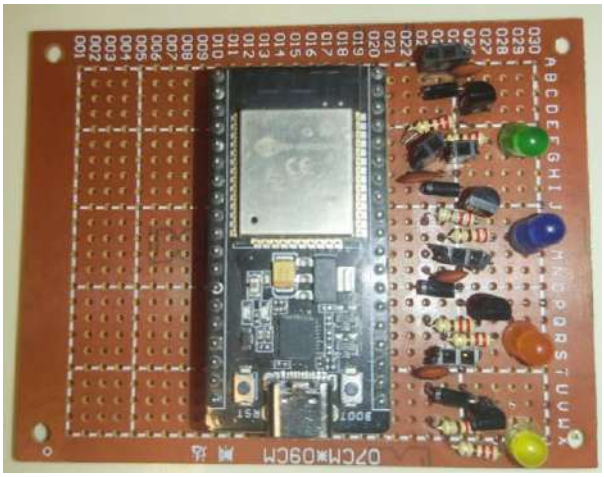


Figure 4. Physical system prototype

Preliminary bench tests (see Figure 5) confirmed that the system can maintain stable vibration frequencies up to 30 Hz. The mechanical response of the motors accurately followed the programmed waveforms, with the Sinusoidal and Triangular modes providing a smoother transition compared to the Square wave, giving the chance to try all these waveforms in a controlled way.

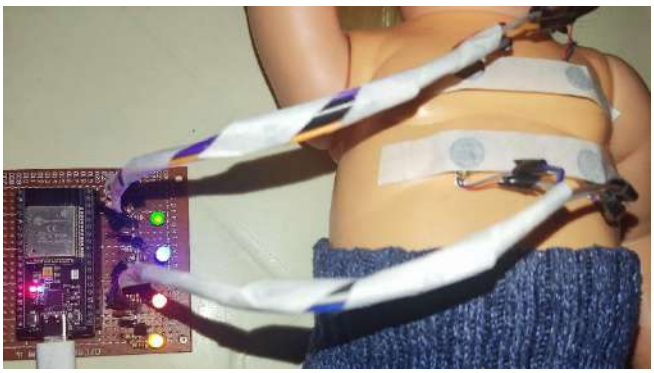


Figure 5. Bench test of the prototype on a mannequin

V. Conclusions

The development of this mechanical stimulation system prototype marks a significant step toward automating respiratory physiotherapy for preterm neonates. By integrating HFCC technology into a controlled electronic environment, this research demonstrates that it is possible to generate precise therapeutic oscillations tailored to the fragile physiology of the neonatal thoracic cage.

Compared to existing HFCWO systems that rely on bulky pneumatic compressors and expensive inflatable vests, this prototype offers a low-cost, modular alternative using localized mechanical actuators. By utilizing off-the-shelf components, the system significantly reduces implementation costs for a potential therapeutic device. The integrated Python/SQLite GUI further enhances usability by providing digital traceability, a feature often missing in traditional manual therapy or basic analog airway clearance devices.

This prototype functions as a first step in the development of a technological alternative to manual chest physiotherapy, with the potential to reduce the high demand for human resources in neonatal intensive care units. While further clinical validation and ergonomic refinements are required, the current system establishes a foundation for a new generation of neonatal airway clearance devices.

VI. Future Work

A. Portability and Electrical Safety

The next iteration will focus on developing a standalone embedded system, developing a small-sized PCB, eliminating the dependency on a PC. This unit will integrate a dedicated touchscreen interface for parameter adjustment and will be powered by a high-capacity Lithium-ion battery system.

B. Ergonomic Wearable Integration

To optimize mechanical coupling, the actuators will be integrated into a specialized ergonomic vest or wrap designed with hypoallergenic materials. This wearable approach ensures that the vibration motors maintain consistent contact with the neonate's thoracic cage in a static place while allowing for easy placement within limited incubator spaces.

C. Respiratory Rate Sensing

A significant addition will be the integration of a respiratory rate monitor based on bioimpedance pneumography. By incorporating non-invasive electrodes, the system will provide real-time feedback on the patient's respiratory status. This data will enable a closed-loop control architecture, where the intensity and frequency of the stimulation could be automatically modulated based on the neonate's physiological response and secretion thickness.

References

[1] Moss, T. J. (2005). The respiratory consequences of preterm birth. In *Proceedings of the Australian Physiological Society* (Vol. 36, pp. 23-28).

[2] Di Polito, A., Del Vecchio, A., Tana, M., Papacci, P., Vento, A. L., Campagnola, B., ... & Vento, G. (2021). Effects of early respiratory physiotherapy on spontaneous respiratory activity of preterm infants: study protocol for a randomized controlled trial. *Trials*, *22*(1), 492. https://doi.org/10.1186/s13063-021-05446-8

[3] Uzundurukan, A., Poncet, S., Boffito, D. C., & Micheau, P. (2024). Acoustic airway clearance devices: a systematic review of experimental and numerical studies. Biomedical Engineering Advances, 8, 100134. https://doi.org/10.1016/j.bea.2024.100134

[4] Gipsman, A.I., Bhandari, A. & Bhandari, V. Use of mucolytics and inhaled antibiotics in the NICU. J Perinatol 45, 5–12 (2025). https://doi.org/10.1038/s41372-024-02178-w

[5] Hullumani S, Qureshi I, Raghumahanti R, Desai S, Brahmane N, Warghat P. Effect of Physiotherapy Intervention on Improvement of Respiratory Function and Degree of Respiratory Distress Reduction in Neonates With Respiratory Distress Syndrome: Protocol for a

Systematic Review. JMIR Res Protoc. 2025 Sep 24;14:e71854. doi: 10.2196/71854. PMID: 40991768; PMCID: PMC12459735.

[6] de Morais Nunes A, Nascimento Sales Figueiredo Fernandes AT, Tévina de Castro Silva A, Guerra Azevedo I, Souza Monteiro K, Pereira SA. Effects of respiratory physiotherapeutic interventions on pulmonary mechanics, vital parameters and pain in newborns: A systematic review. Can J Respir Ther. 2025 Jun 24;61:145-156. doi: 10.29390/001c.140878.

[7] Chatburn, R. L. (2007). High-frequency assisted airway clearance. *Respiratory care*, *52*(9), 1224-1237.

AUTHORS



Mateo Sebastian Rojas Flores He is currently pursuing a bachelor's degree in biomedical engineering from Universidad Privada del Valle (UNIVALLE). He was Vice-chair at the IEEE Student Branch UNIVALLE (2025). He is learning CAD and 3D printing to enhance his projects. His research interests lie primarily in Medical Device Design, Lab Equipment and Biomaterials.



Eynar Calle Viles holds a Master's degree in Telecommunications, with a specialization in Communication Networks and Telematics, and a postgraduate degree in Internet of Things (IoT) from the National Institute of Telecommunications (INATEL), Brazil.

With more than 10 years of professional experience, he is a specialist in the design and implementation of networks, FTTH, IoT, and embedded systems. He has developed several technological innovations, including patent proposals such as the "Mobile Robot for Metal Detection" and the "Localization System for the Tupac Katari Satellite."

In addition, he has multiple publications and participations in international journals and conferences. His main research interests include antennas and propagation, information security, and IoT, actively contributing to the development of innovative technological solutions.

He currently serves as Research Coordinator at the Faculty of Informatics and Electronics at Universidad Privada del Valle (Univalle) and is President of the IEEE Computer Society Bolivia.



Tatiana Torrico Sanchez is a health professional with a strategic focus, leading the Neonatology Service at Hospital Univalle and the Clinical-Surgical Nursing Program at the Universidad Privada del Valle. My leadership is defined by proactivity, educational coaching, and the implementation of nursing reengineering processes aimed at strengthening multidisciplinary capacities in health. I drive innovative projects and practical solutions that generate a direct impact on the population and elevate institutional standards.

Avaliação comparativa dos níveis de estresse e ansiedade em estudantes de Engenharia Biomédica e Medicina da Universidade Privada do Valle, Cochabamba

Brayan Javier Mejía Saavedra, Marcelo Fabian Cruz Guizada, Melissa Andrea Herrera Tambo

Faculdade de Informática e Eletrônica – Univalle, Cochabamba - Bolivia
msb0031404@est.univalle.edu, cgm0031216@est.univalle.edu, htm0031168@est.univalle.edu
https://orcid.org/0009-0008-5378-4060, https://orcid.org/0009-0003-6573-7732, https://orcid.org/0009-0008-5085-9725

***Abstract*— The present study aimed to compare the levels and manifestations of academic stress and anxiety in Biomedical Engineering and Medicine students at the Universidad Privada del Valle (Cochabamba, Bolivia). Data from two samples of 50 students each were analyzed, evaluating four dimensions of stress: psychological, physiological, cognitive, and behavioral, using a Likert scale, complemented by an analysis of the living situation. The results show that, although both programs generate high levels of stress, these differ in nature and impact: in Biomedical Engineering, physiological manifestations (insomnia, tachycardia) and exhaustion associated with the logical-mathematical workload predominate, while in Medicine, stress acquires a more intrusive and emotionally complex character, with greater social isolation, dietary alterations, and ruminative thinking. Living alone was identified as an aggravating factor for emotional burnout in Medicine students, in contrast to Biomedical Engineering. These findings highlight the need to develop differentiated institutional psychological interventions, adjusted to the specific nature of each discipline, rather than generic student well-being strategies.**

***Index Terms*— Anxiety, Biomedical Engineering, Academic Stress, Mental Health, Medicine.**

***Resumo*— O presente estudo teve como objetivo comparar os níveis e manifestações do estresse acadêmico e da ansiedade em estudantes de Engenharia Biomédica e Medicina da Universidade Privada do Valle (Cochabamba, Bolívia). Foram analisados dados de duas amostras de 50 estudantes cada, avaliando quatro dimensões do estresse: psicológica, fisiológica, cognitiva e comportamental, por meio de uma escala Likert, complementada com a análise da situação de convivência. Os resultados evidenciam que, embora ambas as carreiras gerem níveis elevados de estresse, estes diferem em natureza e impacto: na Engenharia Biomédica predominam as manifestações fisiológicas (insônia, taquicardia) e o esgotamento associado à carga lógico-matemática, enquanto na Medicina o estresse adquire um caráter mais intrusivo e emocionalmente complexo, com maior isolamento social, alterações alimentares e pensamento ruminativo. Adicionalmente, identificou-se que viver sozinho atua como agravante do desgaste emocional nos estudantes de Medicina, em contraste com o observado na Engenharia Biomédica. Esses achados ressaltam a necessidade de desenvolver intervenções psicológicas institucionais diferenciadas, ajustadas à natureza específica de cada disciplina, em lugar de estratégias genéricas de bem-estar estudantil.**

***Palavras chave*— Ansiedade, Engenharia Biomédica, Estresse Acadêmico, Saúde Mental, Medicina.**

I. INTRODUÇÃO

Nos últimos anos, a saúde mental dos estudantes universitários tem adquirido crescente relevância no âmbito acadêmico. A transição para a universidade constitui uma mudança importante tanto em termos sociais quanto acadêmicos, que pode afetar a saúde mental considerando as exigências da aprendizagem, a aquisição de novas competências e a adaptação a um novo ambiente formativo [16]. Embora um nível moderado de estresse possa estimular o desejo de superação, um estresse excessivo tem um impacto negativo na saúde física e mental dos estudantes, afetando seu rendimento acadêmico. Os problemas de saúde mental com maior prevalência em universitários são a ansiedade, o estresse e a depressão, cujo principal efeito se reflete em um desempenho acadêmico baixo ou médio[17].

Este fenômeno manifesta-se com especial intensidade em cursos de alta exigência como Medicina e Engenharia Biomédica, embora com perfis distintos. Os estudantes de Medicina apresentam níveis significativos de estresse vinculados à responsabilidade de ter em suas mãos a vida das pessoas, somada à dinâmica exigente no interior das instituições formadoras. Por sua vez, o incremento na carga objetiva de trabalho acadêmico correlaciona-se com uma maior percepção de estresse, estados mais elevados de ansiedade e um maior número de problemas de saúde, sendo o tipo de titulação ou estudos um fator que determina diferenças qualitativas e quantitativas nas experiências subjetivas de estresse [18]. Esta diferença na natureza do estresse entre ambas as populações justifica a necessidade de uma análise comparativa.

O presente trabalho busca avaliar e comparar os níveis de estresse e ansiedade em estudantes de Engenharia Biomédica e Medicina da Universidade Privada del Valle (Cochabamba, Bolivia), por meio da análise de quatro dimensões: fatores psicológicos, fisiológicos, cognitivos e comportamentais. Espera-se que os resultados sirvam como base para o desenvolvimento de estratégias de intervenção adaptadas às particularidades de cada curso.

O estresse é uma resposta adaptativa do organismo às demandas do ambiente que superam os recursos percebidos pelo indivíduo [1]. Numa perspectiva cognitiva, Lazarus et al. [2] definem-no como uma relação particular entre o indivíduo e o seu ambiente, avaliada como ameaçadora ou que excede os seus recursos e coloca em risco o seu bem-estar. Lazarus [3] ampliou este conceito ao integrar as emoções como componentes centrais do processo de estresse, estabelecendo que a avaliação cognitiva determina a resposta emocional diante dos estressores. Intimamente ligada a esse fenômeno, a ansiedade constitui um estado emocional de apreensão e tensão perante situações percebidas como ameaçadoras, com manifestações cognitivas, fisiológicas e comportamentais [4].

Esses mesmos autores indicam que a ansiedade implica a ativação do sistema de ameaça na ausência de perigo objetivo, gerando pensamentos intrusivos e antecipatórios, ao passo que Cano-Vindel et al. [5] distinguem entre ansiedade-estado, como reação emocional transitória, e ansiedade-traço, como disposição pessoal estável.

Para a sua mensuração, utilizam-se instrumentos com propriedades psicométricas adequadas. O estresse percebido é avaliado com a Perceived Stress Scale (PSS) de Cohen et al. [8], cuja validade foi confirmada em contextos universitários latino-americanos [9]; a ansiedade é frequentemente quantificada com o Inventário de Ansiedade Estado-Traço (STAI) de Spielberger et al. [10], instrumentos utilizados em estudos longitudinais recentes com estudantes de ciências da saúde. Esses instrumentos revelam-se especialmente relevantes no âmbito acadêmico, onde o estresse adquire características particulares. Barraza [11] define o estresse acadêmico como aquele que surge em resposta às demandas próprias do processo de ensino-aprendizagem, e o seu Modelo Sistêmico Cognoscitivista propõe que os estudantes experimentam um desequilíbrio sistêmico ao perceberem as exigências escolares como estressores, desencadeando sintomas físicos, psicológicos e comportamentais.

II. MATERIAIS E MÉTODOS

A. Desenho do Estudo e Amostragem

Trata-se de um estudo quantitativo, transversal e comparativo, realizado na Universidade Privada del Valle (Univalle), campus Cochabamba. A amostra foi constituída por 100 estudantes, divididos em dois grupos de 50 alunos cada: Engenharia Biomédica e Medicina. Os participantes foram selecionados considerando a distribuição por semestre e a situação de moradia, distinguindo entre aqueles que vivem com a família ou sozinhos.

B. Instrumentos de Coleta de Dados

A coleta baseou-se em um questionário estruturado para informações sociodemográficas e acadêmicas. Para a mensuração das variáveis principais, utilizaram-se instrumentos validados:

- Estresse Percebido: Avaliado pela *Perceived Stress Scale* (PSS).
- Ansiedade: Quantificada pelo Inventário de Ansiedade Estado-Traço (STAI).
- Escala de Medida: Operacionalizada através de uma escala Likert de cinco pontos, variando de 0 (Nunca) a 4 (Sempre).

Os itens foram agrupados em quatro categorias fundamentais para análise:

1. Fatores Psicológicos.
2. Fatores Fisiológicos.
3. Fatores Cognitivos.
4. Fatores Comportamentais.

C. Procedimentos e Análise Estatística

Os dados brutos foram submetidos a um processo de padronização estatística para permitir a comparação entre as amostras. A análise incluiu:

- Cálculo de frequência percentual para cada item das escalas em ambos os cursos.
- Comparação das médias globais de estresse para avaliar o impacto da situação de moradia.
- Aplicação de testes de correlação estatística para identificar a força e a significância da associação entre o estresse e a ansiedade (como o coeficiente r e o valor p).

III. RESULTADOS E ANÁLISE ESTATÍSTICA

A. Análise dos Fatores

a.1) Fatores Psicológicos:

Os resultados indicam que os alunos de ambas as carreiras apresentam altos níveis de desgaste emocional em relação ao desempenho acadêmico, mas com manifestações e picos ligeiramente distintos devido à natureza de suas matrizes curriculares.

Observou-se que aproximadamente 70% dos estudantes de Engenharia Biomédica relataram sentir ansiedade de forma frequente ou sempre, contra 64% dos estudantes de Medicina. Essa diferença sugere que a constante exigência de resolução de problemas exatos, cálculos e entregas de projetos tecnológicos gera um pico pontual de ansiedade ligeiramente maior na Biomédica. Por outro lado, o desânimo e a sensação de sobrecarga constante afetaram 60% dos alunos de Medicina (frequente ou sempre) em comparação com 56% na Engenharia Biomédica. Essa elevada incidência de sobrecarga contínua na Medicina reflete o impacto do volume massivo de memorização, leituras extensas e o peso emocional atrelado à responsabilidade clínica nos primeiros contatos com pacientes. Ambos os grupos apresentam alta incidência de irritabilidade, comprometendo fortemente a motivação e o bem-estar psicológico geral.

TABELA I
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES PSICOLÓGICOS: ENGENHARIA BIOMÉDICA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Ansiedade quanto ao desempenho acadêmico	2	8	20	40	30
Desânimo / Sensação de sobrecarga	4	10	30	36	20
Irritabilidade ou mudanças de humo	6	14	28	36	16
Nervosismo ou inquietude	3	12	25	40	20

TABELA II
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES PSICOLÓGICOS: MEDICINA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Ansiedade quanto ao desempenho acadêmico	4	4	28	40	24
Desânimo / Sensação de sobrecarga	8	0	32	44	16
Irritabilidade ou mudanças de humor	4	12	44	36	4
Nervosismo ou inquietude	0	28	32	36	4

Estas tabelas apresentam a distribuição percentual das respostas dos alunos em relação aos principais itens avaliados no âmbito dos fatores psicológicos, permitindo uma visualização clara do perfil de desgaste emocional de cada curso.

a.2) Fatores Fisiológicos:

A manifestação física do estresse apresenta divergências significativas entre as duas carreiras, evidenciando como as diferentes rotinas acadêmicas afetam o corpo. A prevalência de insônia é notavelmente elevada na Engenharia Biomédica, com 50% dos alunos relatando dificuldades para dormir de forma frequente ou sempre, enquanto na Medicina esse índice é de 40%. A taquicardia e as palpitações também são consideravelmente mais comuns na Biomédica (35% contra apenas 16% na Medicina).

Por outro lado, os estudantes de Medicina sofrem um impacto físico muito mais severo em termos de tensão muscular e distúrbios gastrointestinais. Impressionantes 60% dos acadêmicos de Medicina relataram tensão muscular (frequente ou sempre), superando os 50% observados na Engenharia Biomédica. Os problemas digestivos (acidez, náuseas e síndrome do intestino irritável) afetam 44% dos estudantes de Medicina, contra 37% dos alunos de Biomédica. Esse aumento expressivo de sintomas gástricos e de tensão corporal nos estudantes de Medicina pode ser justificado pela rotina exaustiva de plantões, privação de sono não programada, e posturas rígidas durante o intenso volume de leitura e práticas anatômico-clínicas.

TABELA III
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES FISIOLÓGICOS: ENGENHARIA BIOMÉDICA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Insônia / Dificuldades para dormir	10	15	25	35	15
Tensão muscular	8	12	30	40	10
Problemas digestivos	20	18	25	30	7
Palpitações/ Taquicardia	15	20	30	25	10

TABELA IV
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES FISIOLÓGICOS: MEDICINA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Insônia / Dificuldades para dormir	4	20	36	24	16
Tensão muscular	4	12	24	36	24
Problemas digestivos	24	16	16	28	16
Palpitações / Taquicardia	24	24	36	12	4

Estas tabelas detalham as respostas relativas aos sintomas físicos, permitindo comparar como o corpo dos estudantes reage a matrizes curriculares de exatas/tecnologia em oposição às ciências biológicas/clínicas.

a.3) Fatores Cognitivos:

No âmbito cognitivo, os alunos de ambas as carreiras enfrentam desafios, mas de naturezas distintas. Os estudantes de Engenharia Biomédica demonstraram maior dificuldade para concentrar-se (40% de incidência frequente ou sempre, contra apenas 24% em Medicina). Essa diferença expressiva pode estar atrelada à exaustão mental gerada pelo raciocínio lógico e matemático contínuo exigido na área de exatas.

Por outro lado, a tendência a antecipar resultados negativos e a presença de pensamentos intrusivos são consideravelmente maiores na Medicina. Mais de 48% dos estudantes de Medicina relataram pensamentos intrusivos frequentes ou sempre (contra 40% em Biomédica) e 48% também antecipam resultados negativos com alta frequência (contra 45% em Biomédica). Isso indica que a pressão inerente ao cuidado com a saúde e à vida humana cria um ciclo de pensamentos ruminantes muito mais acentuado nos futuros médicos, dificultando o relaxamento de 44% dos estudantes mesmo fora dos momentos de avaliação.

TABELA V
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES COGNITIVOS: ENGENHARIA BIOMÉDICA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Dificuldade para concentrarse	12	18	30	28	12
Pensamentos intrusivos	10	20	30	25	15
Antecipação de resultados negativos	5	15	35	30	15
Dificuldade para relaxar	8	20	32	25	15

TABELA VI
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES COGNITIVOS: MEDICINA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES (%)	FREQUENTE (%)	SEMPRE (%)
Dificuldade de concentração	4	32	40	12	12
Pensamentos intrusivos	12	16	24	32	16
Antecipação de resultados negativos	12	12	28	36	12
Dificuldade para relaxar	12	28	16	28	16

a.4) Fatores Comportamentais:

Quanto aos comportamentos, a procrastinação se mostra como uma estratégia de enfrentamento comum e estatisticamente semelhante em ambas as áreas, afetando quase metade dos estudantes (45% na Biomédica e 44% na Medicina). Embora alivie temporariamente a tensão, esse comportamento agrava a pressão do prazo a longo prazo.

As diferenças mais drásticas revelaram-se no isolamento social e na alteração de hábitos vitais básicos. Em Medicina, impressionantes 56% dos alunos afirmam reduzir ativamente as interações sociais e se isolar quando sobrecarregados, em contraste com 45% em Engenharia Biomédica. Além disso, as alterações drásticas nos hábitos alimentares (como comer em excesso ou perder completamente o apetite devido ao estresse) atingem 60% dos estudantes de Medicina, uma margem brutalmente superior aos 40% observados na Biomédica. Devido à ausência de uma métrica direta sobre "uso de substâncias" na avaliação específica de Medicina, a análise para este grupo substituiu o indicador por "condutas repetitivas/compulsivas" para reduzir a ansiedade, manifestadas em 24% da amostra médica.

TABELA VII
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES COMPORTAMENTAIS: ENGENHARIA BIOMÉDICA

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES(%)	FREQUENTE (%)	SEMPRE (%)
Procrastinação	20	15	20	30	15
Isolamento social	18	12	25	30	15
Alterações nos hábitos alimentares	16	14	30	30	10
Uso de substâncias	40	20	20	15	5

TABELA VIII
DISTRIBUIÇÃO DAS RESPOSTAS PARA OS FATORES COMPORTAMENTAIS: MEDICINA.

ITEM	NUNCA (%)	RARA VEZ (%)	ÀS VEZES(%)	FREQUENTE (%)	SEMPRE (%)
Procrastinação	8	24	24	28	16
Isolamento social	4	20	20	48	8
Alterações nos hábitos alimentares	12	16	12	40	20
Uso de substâncias	20	36	20	16	8

a.5) Comparações por Situação de Moradia e Análise Estatística Global

A análise comparativa relacionada à situação de moradia revelou perfis demográficos e de estresse ligeiramente diferentes entre as carreiras. Enquanto na Engenharia Biomédica a esmagadora maioria (80%) vive com a família, na amostra de Medicina observou-se uma proporção maior de alunos morando sozinhos (36%).

Ao analisar as médias globais de estresse (escala de 0 a 4) por situação de moradia, um fenômeno interessante ocorreu na amostra de Medicina em comparação à Biomédica. No estudo prévio de Engenharia Biomédica, os alunos que viviam com a família apresentaram pontuações ligeiramente maiores de estresse em todas as categorias. Em contraste, os estudantes de Medicina que **vivem sozinhos** demonstraram pontuações superiores nos fatores psicológicos (2,58 contra 2,36), cognitivos (2,33 contra 2,02) e comportamentais (2,19 contra 2,09).

Essa inversão sugere que, para o estudante de Medicina, viver sozinho atua como um agravante do desgaste mental e do isolamento, possivelmente devido à falta de uma rede de apoio presencial para lidar com o intenso impacto emocional inerente ao início das práticas clínicas. Contudo, assim como na Biomédica, a análise estatística indica que o estresse é um fenômeno generalizado em toda a amostra, independentemente do ambiente residencial.

TABELA IX
DISTRIBUIÇÃO DOS ALUNOS POR SEMESTRE E SITUAÇÃO DE MORADIA

CURSO	VIVE COM A FAMÍLIA	VIVE SOZINHO	TOTAL
Engenharia Biomédica	40 (80%)	10 (20%)	50
Medicina	32 (64%)	18 (36%)	50

TABELA X
COMPARAÇÃO DOS MÉDIAS DOS FATORES POR SITUAÇÃO DE MORADIA

FATOR/VARIÁVEL	ENG. BIOMÉDICA VIVE COM A FAMÍLIA (MÉDIA ± DP)	ENG. BIOMÉDICA VIVE SOZINHO (MÉDIA ± DP)	MEDICINA VIVE COM A FAMÍLIA (MÉDIA ± DP)	MEDICINA VIVE SOZINHO (MÉDIA ± DP)
Psicológicos	2,55 ± 0,50	2,40 ± 0,60	2,36 ± 0,71	2,58 ± 0,77
Fisiológicos	2,60 ± 0,60	2,45 ± 0,70	2,11 ± 0,80	2,06 ± 0,95
Cognitivos	2,35 ± 0,50	2,20 ± 0,55	2,02 ± 0,97	2,33 ± 0,83
Comportamentais	2,70 ± 0,60	2,50 ± 0,65	2,09 ± 1,00	2,19 ± 0,88

Por fim, a análise global demonstra que a Engenharia Biomédica impõe um estresse com características mais "físicas" e de esgotamento comportamental/tensão constante (maiores médias globais nesses fatores), motivado pelo alto volume de entregas e rigor lógico-matemático. Por outro lado, o estresse na Medicina apresenta um caráter mais intrusivo, isolador e focado no alto impacto dos resultados antecipados, sendo consideravelmente agravado pela falta de apoio presencial nos alunos que residem sozinhos.

Essa inter-relação profunda entre as dimensões físicas, mentais e demográficas consolida a hipótese de que estratégias institucionais de enfrentamento do estresse não podem ser genéricas, devendo ser especificamente desenhadas para a natureza do rigor acadêmico de cada carreira.

Estudos recentes confirmam que a carga acadêmica, os métodos de avaliação e as preocupações vocacionais são os principais desencadeadores do estresse em contextos universitários [12].

Essa relação foi confirmada em estudantes de Engenharia Biomédica da Universidade Privada do Valle, onde foi reportada uma correlação forte entre ambas as variáveis ($r \approx 0{,}80$; $p < 0{,}001$), evidenciando que os estudantes com níveis elevados de estresse tendem invariavelmente a manifestar altos níveis de ansiedade [13]. Essas evidências ressaltam a importância de considerar a especificidade disciplinar no momento de desenvolver intervenções institucionais, visto que os perfis de estresse dos estudantes de Engenharia Biomédica diferem qualitativamente dos observados em outras carreiras universitárias, tanto em suas manifestações quanto em seus principais desencadeadores.

CONCLUSÃO

O presente estudo permitiu comparar os níveis e manifestações do estresse acadêmico entre estudantes de Engenharia Biomédica e Medicina da Universidade Privada do Valle, evidenciando que, embora ambas as carreiras gerem níveis elevados de estresse, estes diferem significativamente em sua natureza, expressão e impacto sobre os estudantes.

Os resultados demonstram que o estresse na Engenharia Biomédica manifesta-se predominantemente nas dimensões fisiológica e comportamental, com maior prevalência de insônia, taquicardia e esgotamento associado ao rigor lógico-matemático e à elevada carga de entregas. Em contraste, o estresse nos estudantes de Medicina adquire um caráter mais intrusivo e emocionalmente complexo, expresso em maior isolamento social, alterações nos hábitos alimentares, pensamentos ruminantes e antecipação de resultados negativos, fatores que se intensificam pela responsabilidade inerente ao cuidado da vida humana desde os primeiros semestres de formação clínica.

Um achado de particular relevância foi a influência da situação de moradia sobre os níveis de estresse. Enquanto na Biomédica viver com a família associou-se a pontuações mais elevadas de estresse, na Medicina ocorreu o fenômeno inverso: os estudantes que residem sozinhos apresentaram níveis significativamente mais elevados nos fatores psicológicos, cognitivos e comportamentais, o que sugere que a ausência de uma rede de apoio presencial atua como agravante do desgaste emocional próprio dessa carreira.

Esses achados reforçam a necessidade de que as instituições de educação superior desenvolvam estratégias de intervenção psicológica e programas de bem-estar estudantil diferenciados, adaptados à natureza específica do rigor acadêmico de cada disciplina. Uma abordagem genérica mostra-se insuficiente para atender à complexidade e heterogeneidade do estresse acadêmico em carreiras da área da saúde. Recomenda-se para estudos futuros ampliar o tamanho amostral, incorporar variáveis como o ano de curso, o gênero e o desempenho acadêmico, e validar os instrumentos utilizados na população boliviana, a fim de obter dados mais representativos e comparáveis em nível regional.

REFERÊNCIAS

[1] Alberto. Orlandini and J. Antonio. Garc-A Barreto, *El estrés*. FCE - Fondo de Cultura Económica, 2001.

[2] R. Stanley. Lazarus and Susan. Folkman, *Stress, appraisal, and coping*. Springer, 1984.

[3] R. S. . Lazarus, *Estrés y emoción : manejo e implicaciones en nuestra salud*. Editorial Desclée de Brouwer, 2011.

[4] M. G. Craske *et al.*, "Anxiety disorders," *Nat. Rev. Dis. Primers*, vol. 3, no. 1, p. 17024, May 2017, doi: 10.1038/nrdp.2017.24.

[5] A. Cano-Vindel and J. J. Miguel-Tobal, "Evaluación de la ansiedad desde un enfoque interactivo y multidimensional: El Inventario de Situaciones y Respuestas de Ansiedad (ISRA). [Evaluation of anxiety from an interactive and multidimensional focus: The Inventory of Situations and Responses of Anxiety (ISRA).]," *Psicología Contemporánea*, vol. 6, no. 1, pp. 14–21, 1999.

[8] S. Cohen, T. Kamarck, and R. Mermelstein, "A Global Measure of Perceived Stress," *J. Health Soc. Behav.*, vol. 24, no. 4, p. 385, Dec. 1983, doi: 10.2307/2136404.

[9] E. Remor, "Psychometric Properties of a European Spanish Version of the Perceived Stress Scale (PSS)," *Span. J. Psychol.*, vol. 9, no. 1, pp. 86–93, May 2006, doi: 10.1017/S1138741600006004.

[10] C. Donald. Spielberger and E. S. . Barratt, *Anxiety : current trends in theory and research*. Academic Press, 1972.

[11] Arturo Barraza Macías, "Un modelo conceptual para el estudio del estrés académico.," *Revista Electrónica De Psicología Iztacala*, 2010. Accessed: Mar. 09, 2026. [Online]. Available: https://revistas.unam.mx/index.php/repi/article/view/19028

[12] L. Avila-Carrasco *et al.*, "Anxiety, depression, and academic stress among medical students during the COVID-19 pandemic," *Front. Psychol.*, vol. 13, Jan. 2023, doi: 10.3389/fpsyg.2022.1066673.

[13] Beiter *et al.*, "The prevalence and correlates of depression, anxiety, and stress in a sample of college students," *J. Affect. Disord.*, vol. 173, pp. 90–96, Mar. 2015, doi: 10.1016/j.jad.2014.10.054.

[14] A. J. Onwuegbuzie and C. E. Daley, "Perfectionism and statistics anxiety," *Pers. Individ. Dif.*, vol. 26, no. 6, pp. 1089–1102, Jun. 1999, doi: 10.1016/S0191-8869(98)00214-1.

[15] Jensen, J. F. Mirabelli, A. J. Kunze, T. E. Romanchek, and K. J. Cross, "Undergraduate student perceptions of stress and mental health in engineering culture," *Int. J. STEM Educ.*, vol. 10, no. 1, p. 30, Apr. 2023, doi: 10.1186/s40594-023-00419-6.

[16] S. T. Trunce Morales *et al.*, «Niveles de depresión, ansiedad, estrés y su relación con el rendimiento académico en estudiantes universitarios», *Investig. En Educ. Médica*, vol. 9, n.º 36, pp. 8-16, dic. 2020, doi: 10.22201/fm.20075057e.2020.36.20229.

[17] M. P. G. López, D. A. J. López, I. C. V. Martínez, y M. E. L. Rivera, «Ansiedad, depresión y su incidencia en el rendimiento académico de estudiantes universitarios», *Entorno*, n.º 78, pp. 78-87, jul. 2024.

[18] R. G. Cabanach, A. Souto-Gestal, y V. Franco, «Escala de Estresores Académicos para la evaluación de los estresores académicos en estudiantes universitarios», *Rev. Iberoam. Psicol. Salud*, vol. 7, n.º 2, pp. 41-50, jul. 2016, doi: 10.1016/j.rips.2016.05.001.

AUTORES



Brayan Javier Mejía Saavedra é estudante do último ano de Engenharia Biomédica na Universidade Privada del Valle, sede Cochabamba, Bolívia, ingressando na universidade em 2022, onde também atua como pesquisador. Integra a Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB), exercendo o cargo de Secretário-Geral na Mesa Diretiva. Possui interesse em projetos que melhorem a qualidade de vida e a saúde das pessoas. Em 2024/2, participou de um intercâmbio pelo programa MARCA MERCOSUL na Universidade Estadual

Paulista (UNESP), no Brasil, integrando a iniciativa Crea Joven SP.



Marcelo Fabian Cruz Guizada é estudante de Engenharia Biomédica na Universidad Privada del Valle (Univalle), localizada em Cochabamba, Bolívia, onde ingressou em 2022. Também atua como pesquisador na instituição. Atualmente é vice-presidente da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB). Seus interesses estão voltados à pesquisa de novas tecnologias para melhorar a qualidade de vida das pessoas, contribuindo por meio da combinação de conhecimentos em engenharia, tecnologia e medicina.



Melissa Andrea Herrera Tambo é estudante do último ano de Engenharia Biomédica na Universidade Privada del Valle (UNIVALLE), localizada em Cochabamba, Bolívia. Integra a equipe de pesquisa UMA e é membro da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB), onde também ocupa cargo na Mesa Diretiva. Além disso, é membro do Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE). Seus interesses estão voltados ao desenvolvimento e à aplicação de soluções biomédicas inovadoras que contribuam para o bem-estar e a saúde das pessoas por meio da integração entre engenharia, tecnologia e ciências médicas.

Projeto e Integração de um Kit IoT de Sinais Vitais para Apoio à Triagem Remota na Plataforma DocIA

Brayan Javier Mejia Saavedra, Maria Laura Nuñez Jaillita, Mishel Bravo Almendras,
Sergio Antonio Vasquez Martinez & Luis Diego Ferreyra Contreras
Universidad Privada del Valle – UNIVALLE
msb0031404@est.univalle.edu, njm0031533@est.univalle.edu, bam0032019@est.univalle.edu,
vms0031149@est.univalle.edu, fcl0031576@est.univalle.edu

***Abstract*—The first contact between a patient and the healthcare system is a critical stage where objective physiological data can improve clinical decision-making. This paper presents a low-cost IoT vital signs kit integrated with DocIA, a digital health platform. The system combines a microcontroller-based device, a Flutter mobile application, and a Supabase cloud backend organized into edge, mobile, and cloud layers. Data acquisition includes heart rate, oxygen saturation, and body temperature, transmitted via WiFi and processed to avoid redundancy. Results demonstrate the technical feasibility of the system as a tool for remote triage and preliminary assessment, particularly in regions with limited access to healthcare infrastructure, such as Bolivia. While currently a prototype, the system provides a foundation for future enhancements, including clinical decision support and real-world validation.**

***Index Terms*— Internet of Things, remote patient monitoring, telemedicine, vital signs.**

***Resumo* —O primeiro contato entre um paciente e o sistema de saúde representa uma etapa crítica, na qual dados fisiológicos objetivos podem melhorar a tomada de decisões clínicas. Este artigo apresenta um kit de sinais vitais IoT de baixo custo integrado ao DocIA, uma plataforma digital de saúde. O sistema combina um dispositivo baseado em microcontrolador, um aplicativo móvel desenvolvido em Flutter e um backend em nuvem gerenciado pelo Supabase, organizado em camadas edge, mobile e cloud. A aquisição de dados inclui frequência cardíaca, saturação de oxigênio e temperatura corporal, transmitidos via WiFi e processados para evitar redundâncias. Os resultados demonstram a viabilidade técnica do sistema como ferramenta de triagem remota e avaliação preliminar, especialmente em regiões com acesso limitado à infraestrutura de saúde, como a Bolívia. Embora ainda em estágio de protótipo, o sistema estabelece uma base para futuras melhorias, incluindo suporte à decisão clínica e validação em ambientes reais.**

***Palavras-chave*— Internet das Coisas, monitoramento remoto de pacientes, sinais vitais, telemedicina.**

I. Introdução

O primeiro contato entre um paciente e o sistema de saúde constitui uma etapa crítica dentro do processo de atendimento médico. Durante esta fase inicial, comumente denominada triagem (triage), a equipe de saúde deve realizar uma avaliação rápida do estado do paciente com o objetivo de priorizar o atendimento conforme a gravidade do caso. No entanto, em muitos contextos, essa avaliação baseia-se principalmente em observações clínicas preliminares ou em informações fornecidas verbalmente pelo paciente, o que pode limitar a precisão da avaliação inicial, especialmente quando não se dispõe de medições objetivas de sinais vitais.

A ausência de ferramentas que permitam registrar e transmitir informações fisiológicas em etapas precoces do processo de atendimento pode gerar atrasos na tomada de decisões médicas e dificultar a priorização adequada dos pacientes. Este problema torna-se mais evidente em ambientes onde os serviços de saúde enfrentam altos níveis de demanda ou onde o acesso a equipamentos médicos especializados é limitado. Nesses cenários, contar com mecanismos que permitam obter dados fisiológicos de maneira rápida e remota pode contribuir para melhorar a eficiência dos processos de atendimento [1].

Em países como a Bolívia, onde a distribuição geográfica dos serviços de saúde apresenta desafios relacionados ao acesso a centros médicos especializados, as soluções baseadas em tecnologias digitais adquirem uma relevância particular. A incorporação de ferramentas que permitam registrar sinais vitais e transmiti-los a plataformas digitais pode apoiar estratégias de telemedicina, facilitando a avaliação preliminar de pacientes antes de uma consulta presencial e ampliando as possibilidades de monitoramento em regiões com acesso limitado à infraestrutura médica [2], [3].

Nos últimos anos, o desenvolvimento de tecnologias baseadas em Internet das Coisas (IoT) permitiu a criação de dispositivos capazes de capturar dados provenientes de sensores biomédicos e transmiti-los através de redes sem fio para plataformas digitais. Esses sistemas facilitam o monitoramento remoto de pacientes e permitem integrar informações fisiológicas em aplicativos móveis e sistemas de gestão clínica. A combinação de dispositivos IoT, aplicativos móveis e serviços na nuvem abre novas possibilidades para o desenvolvimento de soluções orientadas ao monitoramento de saúde e ao apoio na tomada de decisões médicas [4], [5].

Nesse contexto surge a DocIA, uma plataforma digital orientada a integrar ferramentas tecnológicas que facilitem a gestão de informações médicas e o monitoramento de pacientes. A incorporação de dispositivos IoT dentro desta plataforma permite ampliar suas capacidades por meio da captura direta de sinais vitais, os quais podem ser registrados e consultados através do aplicativo móvel do sistema [6], [7], [8].

O presente trabalho descreve o design e a integração de um kit IoT de sinais vitais conectado à plataforma DocIA. O sistema desenvolvido combina um dispositivo de aquisição baseado em microcontrolador, um aplicativo móvel desenvolvido em Flutter e uma infraestrutura backend que permite armazenar e gerenciar as medições obtidas. O objetivo desta integração é avaliar a viabilidade de utilizar dispositivos IoT de baixo custo como apoio para o monitoramento remoto e a obtenção de informações fisiológicas preliminares em cenários de triagem ou avaliação inicial de pacientes [9], [10].

O problema central abordado neste trabalho é a ausência de mecanismos acessíveis para capturar, transmitir e registrar sinais vitais objetivos em etapas iniciais de triagem remota. Embora existam propostas de monitoramento remoto baseadas em IoT, muitas delas permanecem restritas ao desenvolvimento do dispositivo ou à transmissão isolada dos dados. Assim, a principal contribuição deste estudo é apresentar a integração técnica de um kit IoT de sinais vitais com a plataforma DocIA, avaliando sua viabilidade preliminar quanto à aquisição de dados, comunicação sem fio, armazenamento em nuvem e redução de registros redundantes.

II. Materiais e Métodos

A. Arquitetura do Sistema DocIA

A plataforma DocIA foi projetada sob uma arquitetura distribuída que integra dispositivos IoT, um aplicativo móvel e serviços backend para a aquisição, processamento e armazenamento de sinais vitais. O objetivo desta arquitetura é permitir a captura de informações fisiológicas a partir de um dispositivo físico de monitoramento e sua integração dentro de um sistema digital acessível para profissionais de saúde e pacientes. A estrutura do sistema organiza-se em três níveis principais: o nível de dispositivos ou edge layer, representado pelo kit IoT encarregado de capturar os sinais vitais; o nível de interação ou mobile layer, representado pelo aplicativo móvel desenvolvido em Flutter; e o nível de serviços na nuvem ou cloud layer, encarregado de gerenciar o processamento, armazenamento e disponibilidade dos dados.

Dentro desta arquitetura, o kit IoT é responsável por realizar a aquisição de parâmetros fisiológicos por meio de sensores biomédicos conectados a um microcontrolador. Esses dados são transmitidos através da rede sem fio local para o aplicativo móvel, que atua como interface de interação com o usuário e como intermediário para a comunicação com os serviços backend. O aplicativo móvel envia as medições capturadas ao servidor por meio de solicitações a uma API de serviços web, que funciona como ponto central de comunicação entre os diferentes componentes do sistema.

O backend da DocIA encarrega-se de receber essas solicitações, processar os dados e armazená-los no banco de dados gerenciado via Supabase. Além do armazenamento de medições, o Supabase também fornece serviços de autenticação e gestão de usuários, permitindo controlar o acesso à informação médica registrada na plataforma. Esta arquitetura permite integrar dispositivos de monitoramento físico com ferramentas digitais de gestão médica, facilitando o acesso a informações fisiológicas em cenários de atendimento remoto e apoio ao processo de triagem.

B. Arquitetura do Kit IoT

O kit IoT desenvolvido para este sistema baseia-se em um microcontrolador programado através do ambiente de desenvolvimento Arduino, o qual atua como unidade central para a aquisição e transmissão de dados provenientes de sensores biomédicos. Este dispositivo encarrega-se de coordenar a leitura dos sensores, processar os valores obtidos e preparar a informação para sua posterior transmissão ao aplicativo móvel. A escolha de uma plataforma baseada em Arduino deve-se à sua facilidade de integração com sensores biomédicos, seu baixo custo e sua ampla disponibilidade em ambientes de prototipagem eletrônica.

O dispositivo incorpora diferentes sensores encarregados de medir parâmetros fisiológicos relevantes para a avaliação preliminar do estado do paciente, tais como frequência cardíaca, saturação de oxigênio e temperatura corporal. Estes sensores encontram-se conectados ao microcontrolador por meio de interfaces digitais ou analógicas, dependendo do tipo de sensor utilizado. O microcontrolador realiza leituras periódicas dos sensores e estrutura os dados obtidos em um formato adequado para sua transmissão.

Para a comunicação com o aplicativo móvel, o dispositivo conecta-se a uma rede sem fio via WiFi. Durante a fase atual de desenvolvimento do sistema, as credenciais de conexão à rede encontram-se configuradas diretamente no firmware do dispositivo, o que simplifica o processo de conexão durante a etapa de prototipagem. Esta abordagem permite estabelecer rapidamente um ambiente funcional para a captura e transmissão de dados, embora em versões futuras preveja-se incorporar mecanismos de configuração dinâmica de rede que permitam melhorar a usabilidade do sistema.

O protótipo físico do kit IoT DocIA é apresentado na Fig. 1. A estrutura física do kit foi desenvolvida com foco em portabilidade, baixo custo e integração com a plataforma móvel DocIA.



Fig. 1. Protótipo físico do kit IoT DocIA para aquisição de sinais vitais. O dispositivo integra sensores biomédicos para frequência cardíaca, saturação periférica de oxigênio e temperatura corporal, permitindo a transmissão dos dados ao aplicativo móvel da plataforma.

C. Descoberta de Dispositivos e Comunicação

A conexão entre o kit IoT e o aplicativo móvel é realizada por meio de um mecanismo de descoberta de dispositivos baseado na rede local. Quando o aplicativo móvel encontra-se conectado à mesma rede WiFi que o dispositivo, o sistema inicia um processo de busca por sensores ativos dentro da rede. Para identificar o dispositivo correto, o aplicativo utiliza o endereço MAC do kit IoT como identificador único, o que permite distinguir o sensor de outros dispositivos presentes na rede.

Uma vez detectado um dispositivo compatível, o aplicativo mostra sua disponibilidade dentro da interface e oferece ao usuário a opção de estabelecer uma vinculação com o sensor. Após completar este processo, o dispositivo começa a transmitir os dados de sinais vitais para o aplicativo móvel. Esta abordagem permite estabelecer uma comunicação direta entre o kit IoT e o aplicativo sem exigir configurações complexas por parte do usuário.

Durante a fase beta do sistema, o processo de conexão encontra-se simplificado mediante a inclusão manual das credenciais de rede dentro do firmware do dispositivo. Embora esta estratégia facilite o desenvolvimento e os testes do protótipo, versões futuras do sistema contemplam a implementação de mecanismos mais avançados de configuração de rede e emparelhamento de dispositivos.

D. Aplicativo Móvel e Integração com API

O aplicativo móvel da DocIA foi desenvolvido utilizando o framework Flutter e constitui o principal ponto de interação entre o usuário, o dispositivo IoT e a infraestrutura do sistema. Através deste aplicativo, os usuários podem detectar sensores disponíveis na rede local, estabelecer a vinculação com o dispositivo e visualizar as medições de sinais vitais capturadas pelo kit IoT. Além de mostrar os valores atuais, o aplicativo permite consultar registros históricos armazenados na plataforma, o que facilita o acompanhamento da evolução dos parâmetros fisiológicos do paciente.

A comunicação entre o aplicativo móvel e os serviços backend é realizada exclusivamente por meio da API do sistema, a qual expõe diferentes endpoints para a gestão de dados e usuários. Através de solicitações HTTP, o aplicativo pode enviar novas medições de sinais vitais, consultar registros previamente armazenados e acessar a informação associada a cada paciente. Esta abordagem baseada em API permite desacoplar o aplicativo móvel da infraestrutura backend, facilitando a escalabilidade e a integração com outros componentes do sistema.

O aplicativo também gerencia os processos de autenticação de usuários utilizando os serviços fornecidos pelo Supabase. Por meio deste mecanismo, o sistema garante que o acesso à informação médica registrada na plataforma seja restrito a usuários autorizados. Dessa maneira, o aplicativo móvel cumpre um duplo papel dentro da arquitetura do sistema: atuar como interface de visualização de dados e servir como ponte de comunicação entre o dispositivo IoT e os serviços na nuvem.

A interface móvel utilizada para a visualização dos sinais vitais é apresentada na Fig. 2. Por meio dessa tela, o usuário pode acompanhar os valores capturados pelo kit IoT em tempo real, incluindo frequência cardíaca, SpO_2, temperatura corporal e temperatura ambiente. Essa integração permite que os dados fisiológicos adquiridos pelo dispositivo sejam apresentados de forma simples e acessível dentro da plataforma DocIA.



Fig. 2. Interface móvel da plataforma DocIA para visualização dos sinais vitais capturados pelo kit IoT, incluindo frequência cardíaca, SpO_2, temperatura corporal e temperatura ambiente.

E. Worker de Filtragem de Dados

Para otimizar o armazenamento de informação e evitar a geração de registros redundantes, o sistema incorpora um processo de filtragem de dados implementado por meio de um worker desenvolvido em Python. Este componente é executado como um serviço backend encarregado de analisar as medições recebidas antes de seu armazenamento definitivo no banco de dados.

O funcionamento do worker baseia-se na comparação entre cada nova medição recebida e o último registro armazenado para o mesmo dispositivo. Quando o sistema detecta que o valor recebido corresponde a uma medição diferente da previamente registrada, o dado é considerado válido e procede-se ao seu armazenamento no banco de dados. Caso a medição coincida exatamente com o registro anterior, o sistema omite seu armazenamento, evitando assim a acumulação de dados duplicados.

Este mecanismo permite reduzir significativamente a quantidade de registros gerados pelo sistema, especialmente em cenários onde os sensores reportam medições com alta frequência, mas com valores que permanecem constantes durante períodos prolongados. Como resultado, otimiza-se o uso do armazenamento no banco de dados e melhora-se a eficiência do sistema para o processamento posterior da informação.

F. Protocolo de Avaliação Técnica

A avaliação técnica do sistema foi realizada com 30 amostras de frequência cardíaca, saturação periférica de oxigênio e

temperatura corporal. As leituras obtidas pelo kit DocIA foram comparadas com equipamentos comerciais de referência: um oxímetro de pulso clínico Mindray PM-60 para frequência cardíaca e SpO_2, e um termômetro infravermelho Omron MC-720 para temperatura corporal. Para cada variável, calculou-se o erro relativo médio entre o valor medido pelo kit e o valor registrado pelo instrumento de referência.

Também foram avaliados aspectos funcionais da comunicação, incluindo tempo médio de pareamento entre o kit IoT e o aplicativo móvel, latência de transmissão e perda de pacotes em uma rede WiFi local. Por fim, analisou-se a eficiência do worker de filtragem por meio da comparação entre o número de medições transmitidas pelo dispositivo e o número de registros efetivamente armazenados no banco de dados Supabase.

III. Resultados

A avaliação do sistema DocIA concentrou-se em testar o desempenho do kit IoT, sua precisão na captura de dados e a eficiência da transmissão para a plataforma.

A. Precisão de Aquisição do Kit IoT

Para avaliar a precisão do hardware, foi realizado um ensaio de 30 amostras comparando as leituras do kit DocIA com equipamentos médicos de uso comum em centros de saúde de Cochabamba. Como instrumentos de referência, foram utilizados um oxímetro de pulso clínico Mindray PM-60 (para medir a frequência cardíaca e SpO_2) e um termômetro infravermelho Omron MC-720.

Os resultados, apresentados na Fig. 3,4 e 5, refletem um comportamento muito semelhante entre o protótipo e os equipamentos comerciais. O erro relativo médio obtido foi de 0,79% para a frequência cardíaca, 0,28% para a saturação de oxigênio e 0,16% para a temperatura corporal. Essas margens confirmam que o circuito e os sensores do kit conseguem capturar os sinais vitais com a estabilidade necessária para uma etapa de triagem inicial.

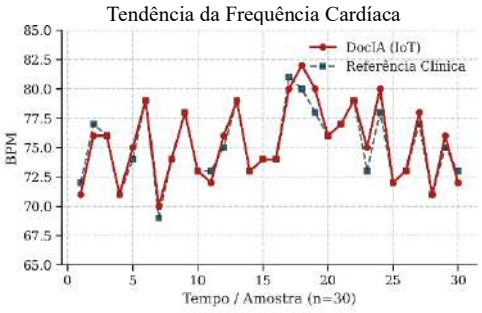


Fig. 3. Tendência da frequência cardíaca nas 30 amostras avaliadas. A linha sólida representa as medições obtidas pelo kit DocIA IoT, enquanto a linha tracejada representa os valores obtidos pelo equipamento clínico de referência.

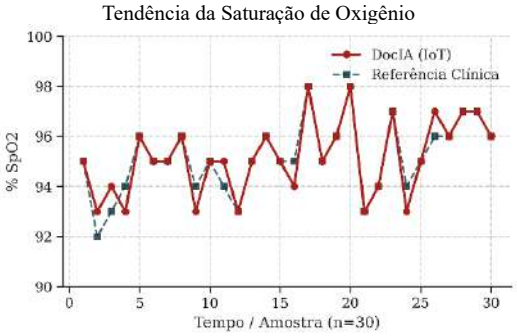


Fig. 4. Tendência da saturação periférica de oxigênio nas 30 amostras avaliadas. Observa-se comportamento semelhante entre as medições do kit DocIA IoT e os valores obtidos pelo oxímetro clínico de referência.

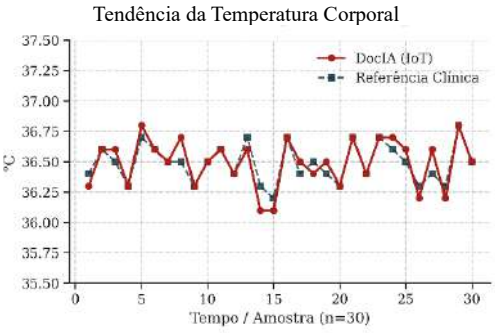


Fig. 5. Tendência da temperatura corporal nas 30 amostras avaliadas. A comparação entre o kit DocIA IoT e o termômetro clínico de referência evidencia estabilidade nas medições obtidas pelo protótipo.

B. Processamento na Borda (Edge) e Transmissão

Para otimizar o rendimento do kit e evitar gargalos na rede, o firmware do microcontrolador foi projetado para empacotar as três variáveis em um único payload estruturado. Esta consolidação reduziu a carga de processamento local. O módulo de comunicação WiFi do kit demonstrou alta eficiência, alcançando o pareamento por endereço MAC com o aplicativo móvel em uma média de 3,2 segundos. Durante o monitoramento contínuo, a latência de envio do hardware para a interface móvel manteve-se entre 150 e 250 milissegundos, com uma perda de pacotes inferior a 2% em um raio de 10 metros.

C. Eficiência do Kit na Arquitetura Cloud

A alta estabilidade e a baixa variância das leituras geradas pelo kit IoT foram fundamentais para otimizar as camadas superiores do sistema. Graças à ausência de ruído severo no hardware, o worker de filtragem pôde identificar com precisão os valores estáticos (especialmente a temperatura corporal) e descartar transmissões redundantes. Isso reduziu o volume total de inserções no banco de dados (Supabase) em 78%, demonstrando que o design robusto do kit reduz diretamente os custos operacionais de armazenamento na nuvem.

IV. Discussão

Os resultados obtidos no presente trabalho demonstram a viabilidade técnica de integrar um kit IoT de baixo custo a uma

plataforma digital de gestão médica para o monitoramento remoto de sinais vitais. A arquitetura distribuída implementada, composta por um dispositivo de aquisição baseado em microcontrolador, um aplicativo móvel desenvolvido em Flutter e uma infraestrutura backend gerenciada via Supabase, permitiu estabelecer um fluxo funcional de captura, transmissão e armazenamento de parâmetros fisiológicos. Esses resultados sugerem que a combinação de tecnologias IoT acessíveis com serviços na nuvem constitui uma estratégia tecnicamente viável para o desenvolvimento de soluções de telemedicina voltadas a contextos com recursos limitados, em concordância com os princípios de design centrado na acessibilidade tecnológica [4], [5].

Sob uma perspectiva arquitetônica, a separação do sistema em três camadas diferenciadas (edge, mobile e cloud) reflete uma abordagem amplamente adotada na literatura sobre sistemas IoT para saúde, onde a modularidade e o desacoplamento de componentes são considerados fatores-chave para a escalabilidade e manutenção da solução. Nesse sentido, o uso de uma API como ponto central de comunicação entre o aplicativo móvel e o backend favorece a extensibilidade do sistema, permitindo a futura incorporação de novos tipos de sensores ou interfaces sem modificar a lógica central da plataforma. Além disso, a implementação do worker de filtragem de dados em Python constitui uma contribuição relevante para a eficiência, ao reduzir a acumulação de registros redundantes, problema frequentemente reportado em sistemas IoT com alta frequência de amostragem [11].

Em comparação com outros estudos de monitoramento remoto via IoT, o sistema desenvolvido apresenta diferenciais vinculados à sua integração com uma plataforma clínica mais ampla. Enquanto diversas propostas na literatura limitam-se ao design de dispositivos isolados , o kit IoT deste trabalho faz parte de um ecossistema digital que inclui gestão de usuários, autenticação e registro histórico, o que amplia seu potencial de aplicação clínica real. No entanto, uma limitação atual reside nas credenciais de rede configuradas diretamente no firmware, o que restringe a portabilidade. Esta questão deverá ser abordada em versões posteriores através da incorporação de mecanismos de configuração dinâmica de rede.

Do ponto de vista de aplicabilidade local, o sistema adquire especial relevância na Bolívia, onde barreiras geográficas e a distribuição desigual dos serviços de saúde dificultam o acesso ao atendimento oportuno. A possibilidade de capturar sinais vitais de forma preliminar por um dispositivo de baixo custo constitui um mecanismo de apoio concreto para processos de triagem remota. Assim, o sistema alinha-se às estratégias nacionais e internacionais de incorporação de tecnologias digitais para ampliar a cobertura de atendimento em regiões com infraestrutura médica limitada [2], [3].

Finalmente, é importante reconhecer que o sistema encontra-se em etapa de prototipagem, e sua validação clínica formal ainda está pendente. A realização de testes com usuários reais em ambientes hospitalares é necessária para determinar a confiabilidade em condições de uso real. Trabalhos futuros deverão focar na realização de ensaios controlados para quantificar o desempenho do dispositivo, incorporar mecanismos adicionais de segurança para proteção de dados sensíveis e explorar a integração de algoritmos de suporte à decisão clínica [9], [10].

V. Conclusão

A integração do kit IoT com a plataforma DocIA representa um passo concreto em direção à democratização do acesso a ferramentas de monitoramento fisiológico em contextos onde a infraestrutura médica tradicional é insuficiente. Além do seu funcionamento técnico, o valor do sistema reside no seu potencial para transformar o processo de triagem: ao dispor de medições objetivas de sinais vitais antes ou durante uma consulta remota, o profissional de saúde acessa informações que, até então, dependiam exclusivamente da presença física do paciente.

No contexto boliviano, esta capacidade adquire uma dimensão especialmente significativa. O sistema abre a possibilidade de estender as coberturas de atendimento a comunidades geograficamente isoladas, onde um dispositivo de baixo custo poderia ser operado por pessoal de saúde comunitário e suas medições revisadas remotamente por um médico, reduzindo, assim, a necessidade de deslocamentos que, em muitos casos, são tardios ou inacessíveis.

Em relação ao desenvolvimento futuro, o sistema constitui uma base sobre a qual podem ser construídas extensões de considerável impacto clínico. A incorporação de algoritmos de suporte à decisão, capazes de alertar sobre valores fisiológicos fora das faixas normais, dotaria a plataforma de uma camada de inteligência diagnóstica que ampliaria sua utilidade além do registro de dados. Da mesma forma, a integração de novos sensores — como eletrocardiografia ou glicometria — permitiria expandir o perfil fisiológico capturado sem alterar a arquitetura central do sistema. Finalmente, a realização de estudos clínicos controlados constitui o passo necessário para transitar do protótipo funcional para uma ferramenta validada, com possibilidades reais de adoção institucional nos sistemas de saúde da região.

Referências

[1] B. M. Mahmmod, M. Abdulrazzaq Naser, A. H. Shanin Al-Sudani, M. Alsabah, H. J. Mohammed y H. Alaskar, «Patient Monitoring System Based on Internet of Things: A Review and Related Challenges With Open Research Issues,» IEEE Explore, 2024.

[2] G. Camacho-Leon, M. Faytong-Haro, K. Carrera, M. Molero, F. Melean, Y. Reyes, H. Mautong, I. De La Hoz y I. Cherrez-Ojeda, «A Narrative Review of Telemedicine in Latin America during the COVID-19 Pandemic,» Multidisciplinary Digital Publishing Institute, 22 07 2022.

[3] F. Saigí-Rubió , «Promoting telemedicine in Latin America in light of COVID-19,» National Library of Medicine, 10 03 2023.

[4] A. Khalad , M. Maslin , R. Fiza Abdul y Y. Yazriwati , «IoT-based remote monitoring system: A new era for patient engagement,» The Institution of Engineering and Technology, 29 08 2024.

[5] V. Dankan Gowda, J. Devendra, C. Senthil kumar, G. Madasamy Raja, D. Srinivas y H. Mohammad, «Impact of IoT on Remote Patient

Monitoring and Advancements in Telemedicine,» IEEE Explore, 04 10 2024.

[6] R. Alexandru, A. Ionut y C. Tudor, «Edge Computing in Healthcare: Innovations, Opportunities, and Challenges,» Multidisciplinary Digital Publishing Institute, 10 09 2024.

[7] A. Archip, N. Botezatu, E. Şerban, P.-C. Herghelegiu y A. Zală, «An IoT based system for remote patient monitoring,» IEEE Explore, 30 06 2016.

[8] H. Tung Yew, M. Fung Ng, S. Zhi Ping, S. Kheau Chung, A. Chekima y J. A. Dargham, «IoT Based Real-Time Remote Patient Monitoring System,» IEEE Explore, 16 04 2020.

[9] A. De La Torre, P. Diaz y R. Perdomo, «Analysis of the virtual healthcare model in Latin America: a systematic review of current challenges and barriers,» National Library of Medicine, 25 03 2024.

[10] K. Mahmoud, C. Jaramillo y S. Barteit, «Telemedicine in Low- and Middle-Income Countries During the COVID-19 Pandemic: A Scoping Review,» National Library of Medicine, 22 06 2022.

[11] M. Alshamrani, «IoT and artificial intelligence implementations for remote healthcare monitoring systems: A survey,» ScienceDirect, 09 2022.

AUTORES



Maria Laura Nuñez Jaillita é estudante de Engenharia de Sistemas na Universidad Privada del Valle (UNIVALLE). Atua como CTO da DocIA, projeto focado na integração de modelos de inteligência artificial em sistemas de software para aplicações em saúde digital. É membro da Seção Bolívia do IEEE, participando de conferências e atividades acadêmicas relacionadas à tecnologia e engenharia.



Brayan Javier Mejía Saavedra é pesquisador no centro UMA (Univalle) e acadêmico de Engenharia Biomédica e de Sistemas (UMSS). É cofundador da DocIA, uma startup de telemedicina que integra ecossistemas IoT e Inteligência Artificial para transformar a triagem remota em diagnósticos objetivos e escaláveis. Membro ativo do IEEE e Secretário-Geral da SCEIB, especializa-se no desenvolvimento de healthtechs de baixo custo e alta precisão. Em 2024, integrou a iniciativa de inovação Crea Joven SP (Brasil) durante intercâmbio na UNESP via programa MARCA MERCOSUL.



Mishel Bravo Almendras é estudante concluinte de Engenharia de Sistemas de Computação na Universidad Privada del Valle (Univalle), Cochabamba, Bolívia. É cofundadora da DocIA, uma startup de telemedicina que integra kits IoT para monitoramento de sinais vitais e inteligência artificial. Atua como desenvolvedora de software com foco em inteligência artificial, análise web, gestão de bancos de dados e segurança da informação. Sua abordagem autodidata a levou a explorar o hacking ético como complemento à sua formação técnica. Participa de pesquisas em diversos campos, com ênfase no desenvolvimento de soluções tecnológicas de impacto.



Sergio Antonio Vásquez Martínez é estudante do último ano de Engenharia Biomédica na Universidad Privada del Valle (Univalle), Cochabamba, Bolívia. Ingressou na universidade em 2022, onde também atua como pesquisador. Atualmente, é presidente da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB). É cofundador da startup DocIA, um projeto de telemedicina que integra kits IoT e inteligência artificial. Seus interesses concentram-se em projetos voltados à melhoria da qualidade de vida e bem-estar através da Engenharia Biomédica.



Luis Diego Ferreyra Contreras é estudante do último ano de Engenharia Eletrônica na Universidad Privada del Valle (UNIVALLE), Cochabamba, Bolívia. É membro da equipe de pesquisa UMA da Univalle e da Sociedade Científica de Eletrônica e Telecomunicações. Seus interesses de pesquisa incluem sistemas embarcados, Internet das Coisas (IoT), instrumentação eletrônica e monitoramento biomédico de sinais vitais. É cofundador da startup DocIA.

Estudo para Desenvolvimento de um Robô para a Categoria Mini Sumô

Alexandre B. Lugli, Andre R. Silva, Egídio R. Neto, Eugênio M. J. da Silva,
Luiz Fernando D. R. dos Santos, Maria Luiza M. Alves, Wanderson E. Saldanha
Instituto Nacional de Telecomunicações (Inatel)
baratella@inatel.br, andre.rodrigues@gea.inatel.br, egidioneto@inatel.br, eugenio@inatel.br,
luizf.santos@inatel.br, maria.alves@gel.inatel.br, wsaldanha@inatel.br

***Abstract*— This work presents the development of a Radio-Controlled Mini-Sumo robot created by the Robotbulls team. The project integrates mechanical design, embedded electronics, and control solutions while complying with the strict size, weight, and safety requirements of the category. The robot was built using high-torque motors, high-grip wheels, an ESP32-S3-WROOM microcontroller, and a custom PCB, with development supported by tools such as SolidWorks, EasyEDA, and VS Code. The results demonstrate that the Mini-Sumo RC category is an effective educational platform, enabling the practical application of multidisciplinary engineering concepts.**

***Index Terms*— Embedded Electronics,Mechanical Design,Mini-Sumo RC, Mobile Robotics.**

***Resumo*— Este trabalho apresenta o desenvolvimento de um robô Mini-Sumô Rádio Controlado criado pela equipe Robotbulls. O projeto integra soluções de mecânica, eletrônica embarcada e controle, respeitando as restrições de tamanho, peso e segurança da categoria. O robô foi construído com motores de alto torque, rodas de alta aderência, microcontrolador ESP32-S3-WROOM e uma PCB personalizada, utilizando ferramentas como SolidWorks, EasyEDA e VS Code. Os resultados demonstram que a categoria Mini-Sumô RC é uma plataforma eficaz de aprendizado, permitindo a aplicação prática de conhecimentos multidisciplinares de engenharia.**

***Palavras Chave*— Eletrônica Embarcada,Mini-Sumô RC, Projeto Mecânico, Robótica Móvel.**

I. Introdução

A equipe Robotbulls, reconhecida no ambiente acadêmico por sua dedicação ao desenvolvimento de tecnologias aplicadas à robótica de competição [1], tem ampliado sua atuação ao explorar categorias que desafiam tanto a engenharia quanto a capacidade estratégica dos competidores [2]. Ao longo de sua trajetória, a equipe consolidou uma cultura de inovação, cooperação e aprimoramento contínuo, utilizando cada projeto como oportunidade de aprofundar conhecimentos práticos e fortalecer as competências técnicas dos integrantes. Dentro desse contexto, a categoria Mini-Sumô Rádio Controlado surge como um campo fértil para experimentação e desenvolvimento [3], unindo complexidade técnica a um ambiente competitivo dinâmico.

As competições de Mini-Sumô RC se diferenciam por exigir não apenas um robô eficiente, mas também interação direta entre operador e máquina. Isso significa que o desempenho final depende tanto do projeto quanto da habilidade humana, tornando a modalidade um laboratório real para estudar responsividade, precisão e controle. A equipe Robotbulls, ao ingressar nessa categoria, se coloca diante de um desafio que envolve desde a concepção mecânica compacta até o domínio de sistemas eletrônicos avançados, passando por decisões estratégicas de pilotagem que influenciam diretamente o resultado dos combates.

A construção de um Mini-Sumô envolve lidar com limitações severas de peso [4], dimensões e espaço interno, o que obriga à seleção minuciosa de materiais e componentes que ofereçam equilíbrio entre leveza, resistência e desempenho. A estrutura do robô precisa suportar impactos intensos e repetidos, enquanto os motores devem fornecer força e aceleração suficientes para deslocar o adversário dentro de uma arena reduzida. Da mesma forma, o sistema de tração deve maximizar o atrito disponível, permitindo que a energia mecânica seja convertida em empurre efetivo. Paralelamente, a eletrônica embarcada precisa garantir comunicação rápida e controle preciso, evitando atrasos que possam comprometer a estratégia do operador.

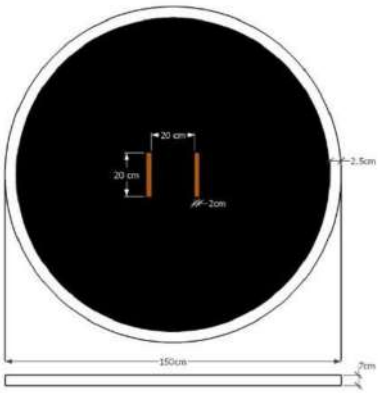


Fig. 1. Superfície circular (dohyō) destinada às disputas de mini-sumô.

Esse conjunto de exigências transforma o processo de desenvolvimento em um exercício multidisciplinar [2], no qual decisões relacionadas à mecânica, eletrônica, software e controle precisam dialogar entre si. Ao enfrentar esses desafios, os membros da Robotbulls entram em contato direto com problemas típicos da prática profissional ,como gerenciamento térmico, otimização de geometria estrutural, integração de módulos eletrônicos, confiabilidade do firmware e eficiência energética. Cada solução adotada reflete a capacidade da equipe de analisar, projetar e testar alternativas dentro de um ciclo contínuo de melhorias.

O presente trabalho tem como objetivo desenvolver um robô Mini-Sumô Rádio Controlado que represente a maturidade técnica da equipe Robotbulls, combinando robustez estrutu-

ral, eficiência mecânica e precisão de controle. O projeto integra materiais cuidadosamente selecionados, motores de alta performance, sistemas de tração especializados, microcontroladores modernos e ferramentas profissionais de modelagem e programação. O resultado esperado é um robô capaz de agir com rapidez, estabilidade e força, respondendo ao operador de forma previsível e eficiente mesmo sob condições intensas de combate.

Dessa forma, esta introdução estabelece não apenas a motivação da Robotbulls ao ingressar na categoria, mas também evidencia a relevância educacional e tecnológica do Mini-Sumô RC [2]. Ao transformar a competição em uma plataforma de aprendizado, o projeto demonstra como a robótica pode potencializar o desenvolvimento de engenheiros mais preparados, criativos e capazes de integrar teoria e prática em soluções de alto desempenho.

II. Fundamentação Teórica

A categoria de Robô Mini-Sumô Rádio Controlado integra um conjunto de competições de robótica que simulam o tradicional confronto do sumô japonês, no qual duas máquinas se enfrentam com o objetivo de empurrar o oponente para fora da arena circular, conhecida como dohyō. Diferentemente das modalidades autônomas, essa categoria caracteriza-se pela intervenção direta do operador por meio de rádio controle, o que faz com que o desempenho do robô dependa tanto das características de engenharia quanto da habilidade humana. Essa interação entre máquina e operador cria um cenário competitivo dinâmico, no qual decisões rápidas e precisão de movimentos tornam-se essenciais para o resultado dos combates.

As regras que regem a categoria são [4] estabelecidas para garantir equilíbrio e segurança entre os competidores. O robô deve respeitar dimensões máximas de 10 centímetros por 10 centímetros na base e peso total de até 500 gramas, incluindo todos os componentes. A arena utilizada normalmente apresenta 77 centímetros de diâmetro, sendo revestida em preto com borda branca, de modo que qualquer queda para fora do círculo resulta na derrota do robô durante o round. A comunicação entre operador e máquina deve ocorrer por meio de sistemas de rádio não interferentes, geralmente na faixa de 2.4 GHz [5], evitando conflitos entre diferentes equipes durante o evento. Antes de cada competição, uma inspeção técnica é realizada para garantir que o robô segue as normas de segurança, não apresenta dispositivos ofensivos ou danosos e opera dentro dos limites estabelecidos.

A categoria impõe diversas limitações que influenciam diretamente o processo de concepção do robô. As restrições de tamanho e peso obrigam as equipes a buscar soluções compactas e eficientes, o que envolve a escolha de motores de alto torque e baixo peso, baterias de elevada densidade energética e chassi elaborado com materiais leves, porém resistentes. A energia embarcada deve ser suficiente para manter o desempenho durante os rounds, sem recorrer a fontes externas, o que exige sistemas eletrônicos otimizados [6]. Além disso, é proibido o uso de armas ou mecanismos que possam causar danos ao adversário ou à arena, o que direciona o foco do desenvolvimento à tração, velocidade, estabilidade e confiabilidade dos componentes mecânicos e eletrônicos [7]. Embora sensores possam ser utilizados para auxiliar em pequenas funções, a atuação autônoma não pode prevalecer sobre o controle manual, preservando o caráter rádio controlado da modalidade.

Os desafios enfrentados pelas equipes ao projetar um Mini-Sumô Rádio Controlado abrangem diferentes áreas da engenharia. Um dos principais está relacionado à tração, pois a capacidade de aderência ao chão da arena define a eficiência do empurrão contra o adversário [8]. Materiais utilizados nos pneus, distribuição de peso e torque dos motores influenciam diretamente o comportamento do robô durante o combate. Outro desafio encontra-se no equilíbrio entre velocidade e estabilidade: enquanto maior velocidade favorece ataques rápidos, ela também aumenta o risco de perda de controle e queda para fora da arena; já uma configuração mais estável pode limitar respostas rápidas frente ao adversário. Além disso, a habilidade do operador desempenha papel determinante, pois o controle manual exige precisão, rápida interpretação do movimento do oponente e estratégias que evitem erros que possam resultar na derrota.

Problemas como aquecimento dos motores, interferências de rádio e organização interna dos componentes também se apresentam como obstáculos significativos, principalmente considerando o espaço extremamente reduzido para acomodar bateria, motores, controlador, receptor e demais circuitos. Esses desafios fazem do Mini-Sumô Rádio Controlado uma categoria que demanda refinamento técnico e domínio operacional. De modo geral, a categoria representa uma síntese de conhecimentos de mecânica, eletrônica, controle e estratégia. O respeito às regras e limitações, somado à superação dos desafios técnicos de projeto e operação, torna a modalidade uma importante ferramenta de aprendizado e desenvolvimento tecnológico. A complexidade envolvida na construção e no controle do robô evidencia a relevância da categoria para a formação de competidores capazes de integrar teoria e prática com eficiência, criatividade e precisão.

III. Materiais e Aplicação Prática

Os robôs da categoria Mini-Sumô RC englobam desafios complexos, desde o desenvolvimento da estrutura e design do robô até a escolha de materiais e componentes tanto para a parte mecânica, quanto para a eletrônica. Todos esses desafios são enfrentados com o objetivo de criar robôs rápidos e compactos que se confrontarão no ambiente competitivo.

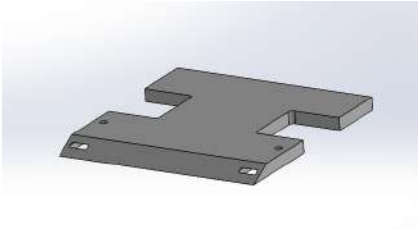


Fig. 2. Exemplo de uma base de um robô de mini-sumô.

O robô é desenvolvido para encontrar a sua máxima eficiência em confrontos, utilizando ligas de aço que trazem resistência, estabilidade e durabilidade para a estrutura, juntamente de polímeros que formam o suporte para os componentes

que compõem o robô, assim, trazendo um balanceamento entre leveza e resistência.

A. *Motores*

Os motores utilizados no robô possuem duas finalidades diferentes, sendo elas a locomoção e o posicionamento das lâminas frontais laterais do robô. Para a locomoção, foi escolhido o Motor Maxon [6] DCX16 devido à sua alta performance, sendo conhecido por sua alta densidade de potência e eficiência, essenciais para maximizar a força de empurre mantendo o peso baixo. A escolha correta do motor é vital para garantir que o robô atinja o equilíbrio entre agilidade e força de tração. Já para o posicionamento das lâminas laterais, foi escolhido o Micro Servo [9] Motor 9g SG90 por ser compacto e leve, e por permitir realizar um controle angular preciso das lâminas frontais laterais do robô.



Fig. 3. Motor Maxon DCX16.



Fig. 4. Micro Servo Motor 9g SG90.

B. *Sistema de controle*

O coração eletrônico e o cérebro do robô são representados por uma Placa de Circuito Impresso (PCB) customizada, que integra o microcontrolador ESP32-S3-WROOM. Este conjunto atua como unidade central de processamento, sendo capaz de adquirir dados de sensores, executar algoritmos de tomada de decisão, e gerar sinais de comando para os motores em tempo real. O ESP32-S3-WROOM foi escolhido pela sua robustez, flexibilidade e capacidade de processamento superior, características cruciais para o processamento rápido necessário durante o combate.

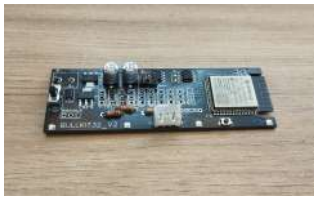


Fig. 5. Placa de circuito impresso integrado com a ESP32-S3-WROOM.



Fig. 6. Exemplo do microcontrolador ESP32-S3-WROOM.

C. *Rodas*

As rodas utilizadas são um formadas por um conjunto de dois componentes, que são as próprias rodas feitas de poliuretano de alta aderência, garantindo que o torque gerado pelos motores Maxon seja efetivamente convertido em força de tração no Dohyō. E o segundo componente é o cubo de roda, que é a interface de contato entre o motor e a arena, o atualmente utilizado é modelo StickyMAX 32mm comercialmente vendido pela empresa Robocore. O desempenho da tração e da força de atrito é maximizado por este componente.



Fig. 7. Exemplo de Roda com o Cubo StickyMax 32mm.

D. *Lâmina*

Para o Mini-Sumô, a lâmina é um dos pontos mais importantes durante um combate, pois existem diversos pontos a serem tratados nesse aspecto, como a sua resistência, a sua espessura, ou até mesmo a sua afiação, além de ser o primeiro ponto de contato com o robô adversário. Tendo como base os pontos passados, foi escolhida as Lâminas Negras Kanzawa, que são feitas de aço temperado de fabricação japonesa, o que proporciona uma alta resistência, ademais, possuem um fio de corte extremamente fino e afiado, gerando uma vantagem para passar sob a lâmina do robô adversário.



Fig. 8. Lâmina Negra Kanzawa.

E. *Alimentação e Inicialização*

A autonomia e controle da partida são garantido por alguns componentes específicos, que são as baterias e o MicroStart. Para obter um desempenho constante o robô é alimentado por duas baterias 2S LiPo de 300mAh e 75c da marca Tattu [10], sendo o ideal para suportar as demandas de tensão e corrente dos motores. E para que se possa inicializar o robô no momento do combate, o MicroStart funciona como um botão de inicialização remoto, o que permite que o robô seja ativado de forma segura e autônoma, dando início à sua lógica de combate sem intervenção manual.

F. *Ferramentas de Desenvolvimento*

O desenvolvimento completo do robô Mini-Sumô exige uma suíte de ferramentas de engenharia que cobrem as esferas mecânica, eletrônica e de software, garantindo a integração eficiente do projeto. Para a concepção e detalhamento do design



Fig. 9. Bateria 2S 300mAh da marca Tattu.

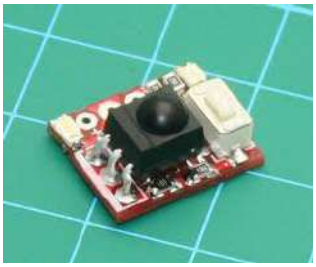


Fig. 10. MicroStart vendido comercialmente pela JSumo.

estrutural e mecânico do robô, o software SolidWorks é empregado [11] , permitindo a prototipagem virtual do chassi e componentes para otimização de massa e resistência a colisões. Já no domínio eletrônico, que inclui o design da placa de circuito impresso (PCB) que hospeda o microcontrolador ESP32-S3-WROOM, o EasyEDA é utilizado [12] para esquematizar e rotear a eletrônica, garantindo que os módulos de potência e sensores sejam integrados de forma compacta e funcional. Por fim, o desenvolvimento do firmware e da lógica de controle é realizado no ambiente de desenvolvimento integrado (IDE) VS-Code [13], que fornece recursos avançados de codificação e debugging necessários para otimizar o desempenho do robô em tempo real.

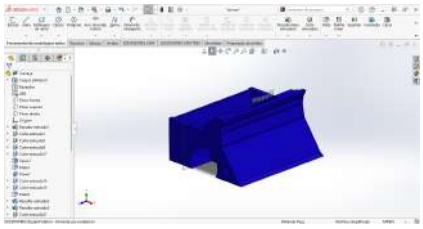


Fig. 11. Exemplo de uma carcaça desenvolvida no software SolidWorks.



Fig. 12. Código desenvolvido no software VS-Code para realizar o controle do robô.

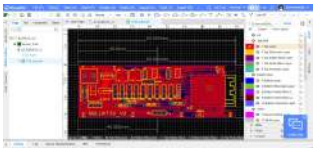


Fig. 13. Esquemático da placa utilizada no robô, desenvolvido na plataforma EasyEDA.

IV. Testes e Resultados

A etapa de testes possui severa importância para o desenvolvimento dos robôs competitivos, para que problemas possam ser notados durante essa fase, e assim corrigidos para que o robô chegue refinado na etapa da competição real.

Durante os testes do robô de três lâminas foi possível notar diversas vantagens sobre outros robôs durante os combates, como uma área de abrangência maior no Dohyō (arena), possibilitando uma maior facilidade em atingir o adversário, além de alcançar o adversário de ângulos diferentes devido às lâminas aumentarem o alcance e serem anguladas. A capacidade de rampar é crucial, e o design de lâminas múltiplas pode potencializar essa tática. Entretanto, também foram notados alguns defeitos que exigem aprimoramento, relacionados à mecânica e eletrônica do robô.

Na mecânica, observou-se a perda de balanço quando as lâminas ainda não foram abaixadas, o que sugere um problema no centro de gravidade do robô que precisa ser corrigido. Além disso, houve a falta de força do Micro Servo Motor para levantar as lâminas e seus suportes, indicando que o torque do motor é insuficiente para a carga.

Na eletrônica, o aquecimento em excesso da placa, chegando a reiniciar o robô devido à alta temperatura, aponta para uma demanda de corrente superior à capacidade dos componentes ou a algum erro no design do circuito. Problemas de superaquecimento e reinicialização podem comprometer a lógica de controle.

Apesar dos defeitos citados, o robô de três lâminas provou ser altamente vigoroso em combate, conseguindo remover os adversários do Dohyō com facilidade, junto ao fato de ser uma ideia pouco implementada no cenário de mini-sumô brasileiro.

V. Conclusão

O desenvolvimento do robô Mini-Sumô Rádio Controlado pela equipe RobotBulls materializa a maturidade técnica e a cultura de inovação da equipe, transformando os desafios da categoria em uma plataforma rica de aprendizado e aprimoramento em engenharia. Este projeto demonstrou como a aplicação de conceitos de mecânica, eletrônica, controle e estratégia converge para a criação de soluções de alto desempenho em um ambiente de competição.



Fig. 14. PaysanBull

O Mini-Sumô RC exige uma síntese de conhecimentos, e o protótipo desenvolvido integrou com sucesso componentes de processamento avançado, sistemas de tração especializada e um design ofensivo. O uso de ferramentas profissionais de modelagem mecânica, projeto eletrônico e desenvolvimento de firmware garantiu a integração eficiente entre as esferas do projeto.

Referências

[1] *Instituto Nacional de Telecomunicações (Inatel)*. https://inatel.br. Acesso em: 09 dez. 2025.

[2] R. Santiago Jr. *Robótica Educacional: Conceitos e Aplicações no Ensino de Engenharia*. Editora Érica, 2018.

[3] Saeed B. Niku. *Robótica: Sistemas e Aplicações*. LTC, 2019.

[4] *Regras Oficiais da Categoria Mini Sumô – Robocore*. https://www.robocore.net/regras. Acesso em: 09 dez. 2025.

[5] *Flysky FS-i6X 2.4GHz – Manual Técnico*. https://www.flysky-cn.com/i6x. Acesso em: 09 dez. 2025.

[6] Maxon Motor. *Motores DC e Brushless: Guia Técnico*. 2020. URL: https://www.maxongroup.com.

[7] João Silva e Marcos Costa. *Sistemas Embarcados e Eletrônica Aplicada*. Érica, 2018.

[8] David Halliday, Robert Resnick e Jearl Walker. *Fundamentos de Física - Mecânica*. 10ª ed. LTC, 2016.

[9] *Funcionamento de Servomotores – Guia Técnico*. https://www.servomotors.net/how-it-works. Acesso em: 09 dez. 2025.

[10] *Guia Técnico de Baterias LiPo*. https://oscarli.com/lipo-guide. Acesso em: 09 dez. 2025.

[11] *SolidWorks – Documentação e Guias Técnicos*. https://www.solidworks.com. Acesso em: 09 dez. 2025.

[12] EasyEDA Tech. *EasyEDA – Documentation and User Manual*. Disponível em: ¡https://docs.easyeda.com¿. 2023.

[13] Microsoft Corporation. *Visual Studio Code – Documentation*. Disponível em: ¡https://code.visualstudio.com¿. 2023.

AUTORES



Maria Luiza Muquem Alves é graduanda do quarto período em Engenharia Elétrica pelo INATEL – Instituto Nacional de Telecomunicações, membro atuante da Robotbulls — equipe de robótica do INATEL — e bolsista de Iniciação Científica.



Andre Rodrigues Silva é graduando do quarto período em Engenharia de Controle e Automação pelo INATEL – Instituto Nacional de Telecomunicações, membro atuante da Robotbulls — equipe de robótica do INATEL — e bolsista de Iniciação Científica.

Desenvolvimento de Padrão para Robô de Combate Categoria Cupim 454g

Caio de C. Yarouhas, David B. J. do Prado
Alexandre B. Lugli, Eugênio M. J. da Silva, Egídio R. Neto,
Luiz Fernando D. R. dos Santos e Wanderson E. Saldanha
caio.castro@ges.inatel.br, david.prado@gea.inatel.br, baratella@inatel.br, eugenio@inatel.br,
egidio.neto@inatel.br, luizf.santos@inatel.br, wsaldanha@inatel.br

Instituto Nacional de Telecomunicações – Inatel

***Abstract* — Robotics competitions have stimulated the development of small-scale combat robots, encouraging the application of concepts from mechanics, electronics and embedded systems. This work presents the development and assembly process of a combat robot in the Cupim category, focusing on the construction techniques, materials and electronic components required for its operation. The robot structure is primarily built using wood-based materials, following the traditional constraints of the category, and integrates propulsion, control and power systems designed to operate within strict weight and size limits.**

***Index Terms* — Combat robotics, Cupim robot, mechanical design, embedded electronics.**

***Resumo* — As competições de robótica têm estimulado o desenvolvimento de robôs de combate de pequeno porte, promovendo a aplicação de conceitos de mecânica, eletrônica e sistemas embarcados. Este trabalho apresenta o processo de desenvolvimento e montagem de um robô da categoria Cupim, com foco nas técnicas de construção, materiais utilizados e componentes eletrônicos necessários para seu funcionamento. A estrutura do robô é construída principalmente em madeira, conforme as características da categoria, integrando sistemas de locomoção, controle e alimentação projetados para operar dentro de limites restritos de peso e dimensões.**

***Palavras Chave* — Robótica de combate, robô Cupim, projeto mecânico, eletrônica embarcada.**

I. INTRODUÇÃO

Robôs de combate são sistemas eletromecânicos projetados para competir em arenas controladas como exibido na figura 1, nos quais diferentes projetos se enfrentam utilizando estratégias de ataque, defesa e mobilidade. Essas plataformas exigem a integração de diversos subsistemas, incluindo estrutura mecânica, sistemas de locomoção, controle eletrônico e gerenciamento de energia.

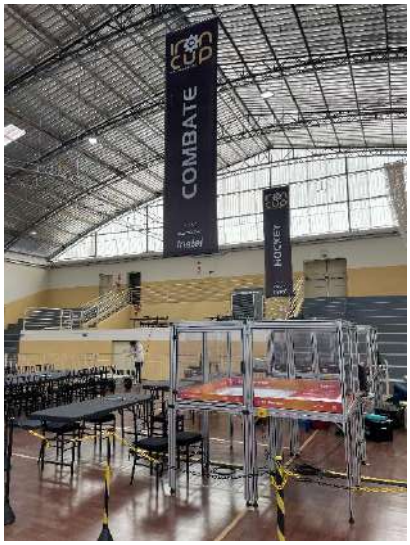


Figura 1

A categoria Cupim destaca-se por utilizar predominantemente materiais de madeira na estrutura do robô, incentivando projetos leves, simples e acessíveis. Essa característica impõe desafios específicos de engenharia, pois o robô deve ser resistente o suficiente para suportar impactos durante os combates, ao mesmo tempo em que mantém boa mobilidade e estabilidade.

O desenvolvimento de um robô dessa categoria envolve decisões relacionadas à escolha dos materiais, distribuição de massa, seleção de motores, baterias e sistemas de controle. Cada componente influencia diretamente o desempenho do robô dentro da arena.

Este artigo apresenta os principais elementos envolvidos na construção de um robô Cupim, descrevendo os materiais utilizados, os componentes eletrônicos empregados e a integração dos sistemas necessários para o seu funcionamento.

II. FUNDAMENTAÇÃO TEÓRICA

A. Regras e limitações

Na categoria Cupim, os robôs devem respeitar limites de peso e dimensões, definidos pelas regras da competição. Esses limites exigem que o projeto seja cuidadosamente planejado para garantir que todos os componentes necessários possam ser integrados sem ultrapassar as restrições estabelecidas.

Os combates ocorrem em arenas fechadas e possuem duração máxima dois minutos. Durante esse período, os robôs devem demonstrar capacidade de mobilidade, resistência estrutural e eficiência em suas estratégias de ataque ou defesa.

A estrutura do robô é tradicionalmente construída utilizando madeira como material principal, o que influencia diretamente o projeto mecânico. O uso desse material exige soluções de engenharia que garantam rigidez estrutural e distribuição adequada de cargas durante impactos.

B. Estrutura mecânica

A estrutura do robô Cupim é construída principalmente em madeira, sendo o MDF *(Medium Density Fiberboard)* o material utilizado. Este apresenta boa rigidez estrutural, baixo custo e facilidade de usinagem, permitindo cortes precisos e montagem simples.

O chassi do robô é geralmente composto por:

- base estrutural
- laterais de proteção
- suportes internos para componentes eletrônicos.

A montagem dessas partes pode ser realizada utilizando parafusos, cola estrutural ou ambos, garantindo maior resistência mecânica durante os combates.

Apesar de ser leve, a madeira possui limitações em relação à resistência a impactos. Por esse motivo, o projeto do chassi deve considerar reforços estruturais e distribuição adequada de massa, evitando pontos de concentração de esforço.

III. COMPONENTES DO ROBÔ

A. Sistema de locomoção

A movimentação do robô é realizada por motores elétricos DC, responsáveis por acionar as rodas. Esses motores são conectados diretamente ou por meio de sistemas de transmissão.

As rodas são normalmente fabricadas em plástico ou borracha, sendo frequentemente revestidas com material de maior aderência para melhorar a tração na arena.

O posicionamento das rodas e a distribuição de peso influenciam diretamente a estabilidade e a capacidade de manobra do robô.

B. Sistema de alimentação

O sistema de energia do robô é composto por baterias recarregáveis, geralmente do tipo LiPo (polímero de lítio). Essas baterias são escolhidas por apresentarem alta densidade energética e baixo peso.

A tensão fornecida pela bateria pode variar durante o uso, sendo necessário o uso de reguladores de tensão para alimentar componentes eletrônicos que operam em níveis específicos.

C. Sistema de controle eletrônico

O controle do robô é realizado por um microcontrolador, responsável por interpretar comandos e controlar os motores.

Placas baseadas em ESP32 ou Arduino podem ser utilizadas para esse fim. Esses dispositivos permitem o processamento dos sinais recebidos do sistema de controle e o acionamento dos atuadores do robô.

O microcontrolador também pode ser responsável por funções adicionais, como monitoramento de energia ou controle de mecanismos de ataque.

D. Conexões elétricas

A comunicação entre os componentes é feita por meio de fios elétricos e conectores, que distribuem energia e sinais de controle dentro do robô.

A escolha da bitola dos fios deve considerar a corrente elétrica necessária para alimentar motores e circuitos eletrônicos.

E. Sistema de segurança

Para garantir a operação segura do robô, é instalada uma chave geral liga/desliga, responsável por interromper ou permitir o fluxo de energia para o sistema.

Além disso, um LED indicador pode ser utilizado para informar visualmente quando o robô está energizado.

IV. MATERIAIS UTILIZADOS

A construção de um robô Cupim envolve diferentes componentes mecânicos e eletrônicos que trabalham de forma integrada.

1. Estrutura em MDF

O MDF é utilizado como material principal na construção do chassi do robô. Suas principais vantagens são:

- facilidade de corte
- baixo custo
- boa rigidez estrutural
- facilidade de usinagem

O MDF compõe principalmente:

- base do robô
- laterais do chassi
- suportes internos

- Para aumentar a resistência, é comum utilizar reforços estruturais conforme o espaçamento descrito nas regras.

2. Motores DC

Os motores elétricos DC são responsáveis pela locomoção do robô, acionando diretamente as rodas. Eles devem possuir:

- torque suficiente para movimentar o robô
- velocidade adequada para manobras rápidas
- boa eficiência energética

3. Rodas

As rodas garantem a mobilidade do robô dentro da arena. Normalmente são utilizadas rodas com:

- revestimento em borracha
- boa aderência
- resistência ao desgaste

Uma boa tração é essencial para empurrar o adversário e manter estabilidade.

4. Baterias

As baterias LiPo (polímero de lítio) são amplamente utilizadas devido à sua alta densidade energética. Elas fornecem energia para:

- motores
- microcontrolador
- circuitos auxiliares

Sua escolha deve considerar:

- tensão nominal
- capacidade (mAh)
- taxa de descarga

5. Regulador de tensão

O regulador é utilizado para reduzir e estabilizar a tensão da bateria, normalmente variando entre 6,4 V e 8,4 V, para 5 V, valor necessário para alimentar o microcontrolador.
Isso garante funcionamento estável dos circuitos eletrônicos.

6. Microcontrolador ESP32

O ESP32 é o controlador central do robô. Ele é responsável por:

- processar comandos
- controlar motores
- realizar comunicação sem fio

O ESP32 opera com:

- tensão de alimentação de 5 V
- sinais lógicos de 3,3 V

Também possui conectividade Bluetooth e Wi-Fi, permitindo controle remoto.

7. Fios e conectores

Os fios elétricos realizam a interligação entre os componentes. Fios mais grossos são utilizados em circuitos que conduzem correntes maiores, principalmente na alimentação dos motores.

8. Chave liga/desliga

A chave é utilizada para controlar o fluxo de energia no sistema.

- Posição O → circuito aberto (desligado)
- Posição I → circuito fechado (ligado)

9. LED indicador

O LED funciona como indicador visual de funcionamento do robô.
Quando aceso, indica que o sistema está energizado.

V. CONCLUSÃO

O desenvolvimento de um robô da categoria Cupim envolve a integração de diversos sistemas mecânicos e eletrônicos dentro de restrições rigorosas de peso e tamanho, a figura 2. A utilização de madeira como material estrutural exige soluções de engenharia que garantam resistência suficiente para suportar impactos durante os combates.

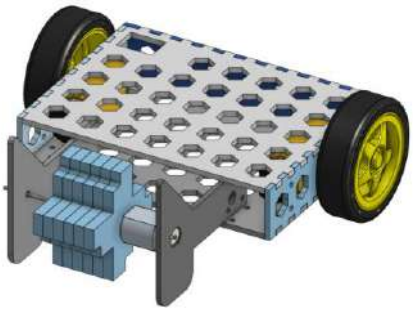


Figura 2

A seleção adequada de motores, baterias, microcontroladores e componentes eletrônicos é fundamental para garantir o funcionamento eficiente do robô. A montagem cuidadosa do chassi e a organização interna dos sistemas também contribuem para melhorar o desempenho e a confiabilidade durante as competições.

Dessa forma, o projeto de robôs Cupim representa um desafio técnico que envolve conhecimentos de mecânica, eletrônica e sistemas embarcados, resultando em plataformas compactas e funcionais para aplicações em robótica de combate.

REFERÊNCIAS

[1] RoboCore, *Regras – Combate de Robôs*. Disponível em: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Combate.pdf

[2] RoboCore, *Critério de Julgamento – Combate de Robôs*. Disponível em: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Crit%C3%A9rio_Julgamento+-+Combate.pdf

[3] RoboCore, *Regras – Combate Cupim*. Disponível em: https://robocore-eventos.s3.sa-east-1.amazonaws.com/public/Regras+-+Combate+Cupim.pdf

[4] RoboCore, *Portal Combate de Robôs*. Disponível em: https://www.combatederobos.com.br/

Architecture of an Autonomous Vehicle with ROS for Mapping and Navigation in Static Environments

Carlos Leonardo Lazzari, Murilo Ribeiro Bonato, Kleyton Hoffmann
Universidade do Oeste de Santa Catarina - UNOESC
carlos.leonardo290403@gmail.com, murilobonato03@gmail.com, kleyton.hoffmann@unoesc.edu.br

***Abstract*—This work presents the architecture of a 1/10 autonomous vehicle for waypoint navigation in a static indoor environment. The system was implemented on the NVIDIA® Jetson Orin Nano™ platform and utilizes the Robot Operating System (ROS). The prototype adopts an Ackermann geometry and integrates a 2D LiDAR, an IMU, and encoders for perception, localization, mapping, and real-time navigation, employing the SLAM Toolbox and Nav2 libraries. The navigation architecture uses behavior trees, the AMCL algorithm, the Hybrid A* planner, and the MPPI controller for trajectory generation and the computation of velocity and steering commands. The results demonstrate the correct functioning of the system and sensor integration, enabling simultaneous localization and mapping of the robot in static environments, as well as the capability to generate and follow trajectories to predefined waypoints while respecting the vehicle's kinematic constraints.**

***Index Terms*—Autonomous Vehicle Navigation, SLAM, Motion and Path Planning, Robot Operating System (ROS).**

I. Introduction

As modern engineering practices and computational technologies evolve, advanced methods for perception, navigation, and autonomous control have been increasingly integrated into vehicles and robots. Despite technological, regulatory, and economic challenges, such systems have become widespread in sectors such as logistics, healthcare, agriculture, and manufacturing, enhancing efficiency and automation [1]. In general, an autonomous robot is defined as a system capable of perceiving its environment, planning its actions, and acting independently to achieve predefined objectives [2].

A key enabler of this development is the Robot Operating System (ROS), an open-source middleware framework introduced in 2007. ROS provides a structured communication architecture along with libraries and tools for the development of robotic systems. Its functionalities include support for perception, mapping, planning, navigation, simulation, debugging, and data visualization, with native support for C++ and Python [3].

In Brazil, the RoboCore Trekking PRO competition promotes the development of autonomous mobile robots capable of navigating known environments with diverse obstacles in the shortest possible time, serving as an experimental and educational platform that fosters research in the field [4].

In this context, this work presents the development of an integrated hardware and software architecture based on ROS 2 for an autonomous mobile robot designed for competitive environments. The proposed system incorporates mapping, localization, trajectory planning, and navigation techniques in static scenarios, using sensor data from a 2D LiDAR (Light Detection and Ranging), wheel encoders, and an Inertial Measurement Unit (IMU).

II. Odometry System

The odometry system uses real-time sensor data to continuously estimate the robot's position, orientation, and trajectory relative to a fixed reference frame, enabling the reconstruction of its motion over time [5].

In this work, for the proposed system with Ackermann steering, the linear velocity v_x is approximated from measurements obtained from the rear wheel encoders, assuming negligible slip conditions. The vehicle heading ϕ is derived from the IMU orientation. At each time interval Δt, the vehicle position (x', y') is updated using Equations 1 and 2, where the linear velocity is integrated along the estimated heading.

$$x'_t = x_t + v_x \cos(\phi)\, \Delta t \tag{1}$$

$$y'_t = y_t + v_x \sin(\phi)\, \Delta t \tag{2}$$

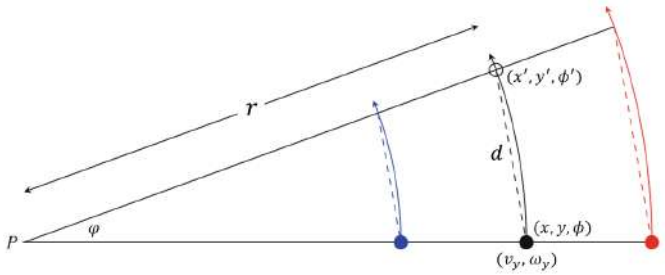


Fig. 1. Simplified geometric odometry model for a vehicle with Ackermann steering performing a left turn in the XY plane. Adapted from [5].

III. Simultaneous Localization and Mapping (SLAM)

SLAM is the process of building a map of an unknown environment, representing relevant features such as reference landmarks and obstacles, while simultaneously estimating the robot's location and state [6].

The SLAM architecture is typically divided into a *front-end*, responsible for processing sensor data and estimating the robot's pose, and a *back-end*, which performs optimization based on factor graphs, as illustrated in Figure 2, to refine the estimates, minimize errors, and generate consistent maps. The interdependencies between variables and loop closure detection reduce odometry drift and enable the recognition of previously visited locations [7, 8].

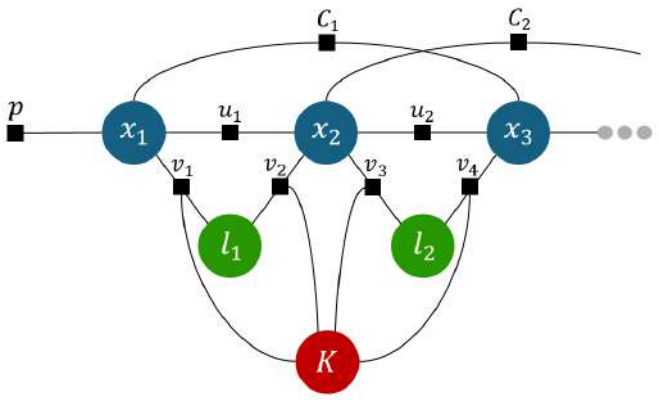


Fig. 2. SLAM represented as a factor graph, where: **(I) Blue circles** – robot poses (x_1, x_2, ...); **(II) Green circles** – reference landmarks (l_1, l_2, ...); **(III) Red circle** – calibration parameters (K); **(IV) Black dots** – probabilistic factors: odometry (u), LiDAR observations (z), loop closures (c), and prior factors (p) [9].

In LiDAR-based SLAM systems, the relative motion between consecutive scans is commonly estimated using scan matching techniques, which align point clouds by minimizing a distance metric between observations. Classical approaches, such as the Iterative Closest Point (ICP) algorithm [10], estimate the rigid transformation between a source cloud P and a target cloud Q through a least-squares formulation based on point correspondences.

The rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$ are obtained by minimizing the mean squared distance between matched points, as expressed in Equation 3. The source point cloud is transformed according to $p'_i = \mathbf{R} \cdot p_i + \mathbf{T}$, and this process is iterated until convergence.

$$E(\mathbf{R}, \mathbf{T}) = \frac{1}{n} \sum_{i=1}^{n} \|p'_i - q_i\|^2 \tag{3}$$

In this work, mapping and localization were performed using the SLAM Toolbox package, which implements 2D pose-graph-based SLAM with odometry and LiDAR data [11]. The system uses correlative scan matching in the front-end and pose graph optimization in the back-end, enabling real-time occupancy grid mapping with multi-session support. The system operates online and asynchronously, with pose estimates refined through nonlinear least-squares optimization solved using the Ceres Solver library [12, 13].

IV. Autonomous Navigation System

The autonomous navigation system of the proposed platform was implemented using the Navigation2 (Nav2) framework, which provides modules for localization, trajectory planning, and control in 2D environments. The system follows the standard architecture of the framework, as presented in Figure 3 [14, 15].

A. Adaptive Monte Carlo Localization (AMCL)

The Adaptive Monte Carlo Localization (AMCL) algorithm is a probabilistic technique used in this work to estimate the robot's pose within a previously constructed map generated by SLAM algorithms [17].

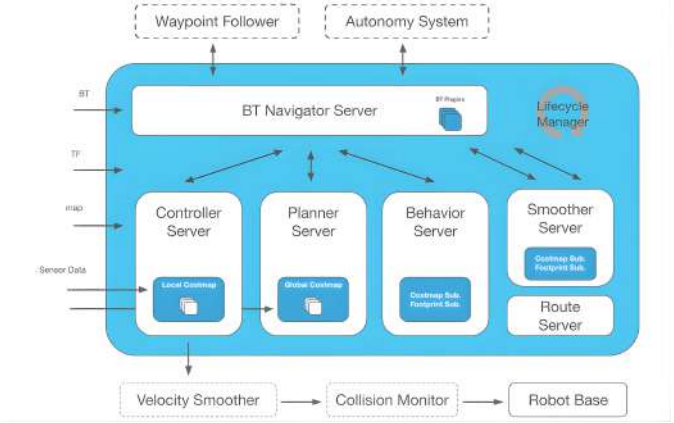


Fig. 3. Architecture of the Nav2 library [16].

The process is based on a distribution of hypotheses that represent possible robot positions and consists of three main steps: the prediction of particles according to the robot's motion, the correction of their weights based on sensor readings, and the resampling of low-probability hypotheses. These steps allow the algorithm to determine the robot's current position and to recover from localization errors caused by inaccurate sensor estimates [18].

B. Layered Costmap

Initially, a layered costmap is generated on top of the original map using the algorithm proposed by [19]. This costmap is composed of an ordered list of layers, each assigning a traversal cost to different regions of the environment. These costs reflect the difficulty of navigating through each area, considering the presence and proximity of obstacles, terrain characteristics, and operational constraints. Such information guides trajectory planning by favoring lower-cost paths and improving the overall quality of the route, as illustrated in Figure 4.

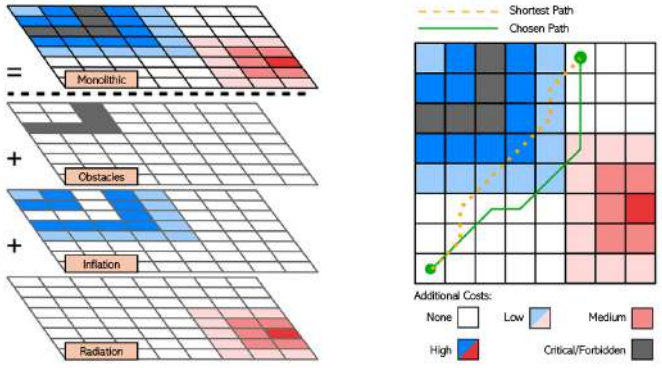


Fig. 4. Example of a layered costmap generated from the combination of multiple layers. Adapted from [20].

C. Global and Local Trajectory Planners

The global planner used in this work is based on the Hybrid A* algorithm, generating a kinematically feasible path over the global costmap [21]. After the global planner defines a viable route, the local planner employed is the Model Predictive Path Integral (MPPI) controller for nonlinear systems. It computes

control actions (linear and angular velocities) through predictive sampling, iteratively selecting and optimizing trajectories at each control cycle using a cost function that accounts for collisions, curvature, speed limits, steering angle, acceleration, and other criteria [22, 23]. At each iteration, the trajectory with the lowest cost (green line) is selected, and the first control action of this trajectory is applied to the system, as illustrated in Figure 5.

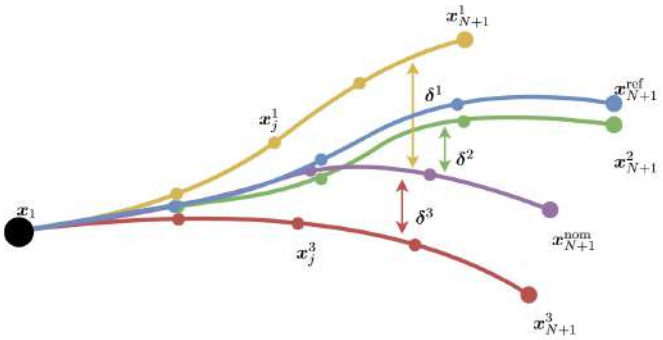


Fig. 5. Generation and evaluation of sampled trajectories by the MPPI controller [24].

V. Project Scope

The objective of this work is to develop a 1/10-scale autonomous vehicle, with dimensions of up to 500×500×500 mm, for competitions in the Trekking PRO category. The project implementation and technical details are available in [25].

The prototype must be capable of mapping static environments under operator control and, subsequently, localizing itself, generating and following trajectories autonomously to predefined points, completing the course within a maximum time of 10 minutes, in accordance with the competition rules.

A. Developed Prototype

The prototype of the autonomous vehicle developed at a 1/10 scale, including its structure, sensor placement, and electrical connections, is shown in Figure 6.



Fig. 6. Side view of the developed prototype.

VI. Hardware Architecture

Figure 7 presents the block diagram of the hardware architecture of the developed prototype, based on the NVIDIA® Jetson Orin Nano development board, which serves as the central processing, control, and interface unit.

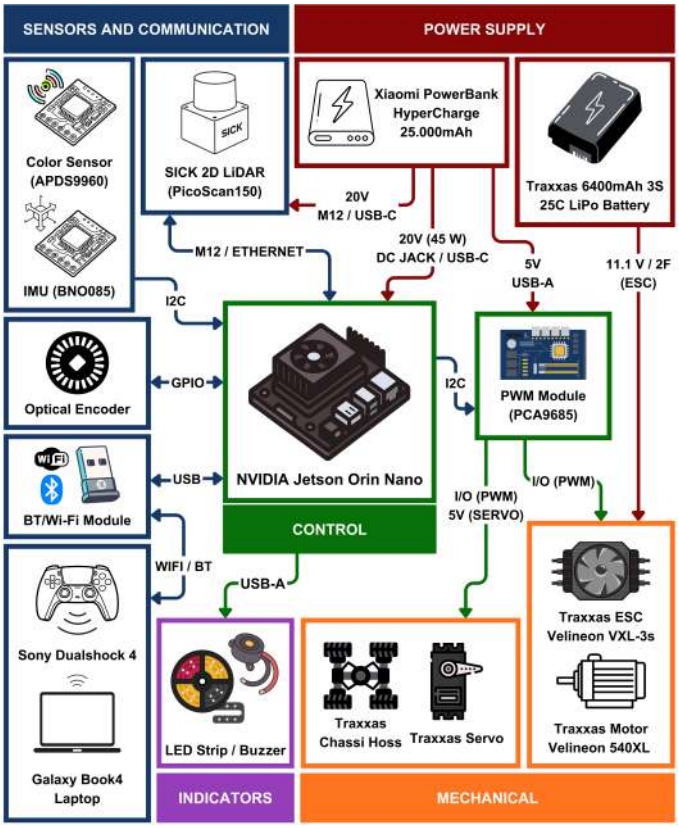


Fig. 7. Block Diagram – Hardware architecture of the developed prototype.

VII. System Configuration

The system was deployed on a Linux environment running the ROS 2 Humble Hawksbill distribution, with the package structure implemented in Python. Communication between nodes was carried out using the Eclipse Cyclone DDS middleware.

VIII. System Operating Architecture

The vehicle's control architecture was designed to be modular and flexible, enabling robust operation and smooth transitions between manual and autonomous modes under competition constraints. Figure 8 illustrates the system initialization, including sensor configuration, acquisition, and processing. After initialization, the user selects the operating mode via ROS, and only the corresponding modules are activated, ensuring continuous data exchange between components through DDS.

In Manual Mode, illustrated by the flowchart in Figure 8, the vehicle is directly controlled by the operator, while the SLAM algorithm performs complete mapping of the competition environment, including objectives, obstacles, and constraints. At the end of this operation, the generated map, which includes the image, metadata, pose graph, and point orientations, is saved, and the waypoints selected in RViz are stored for use during the autonomous mode.

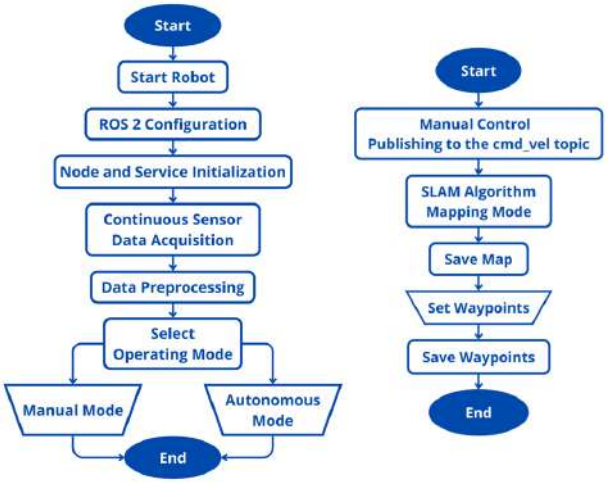


Fig. 8. Illustrative flowcharts: on the left, the vehicle initialization process and Operating Mode selection; on the right, manual vehicle operation.

In Autonomous Mode, illustrated in Figure 9, the system uses the previously generated map and LiDAR data, operating SLAM in localization mode to estimate the robot's pose. The Nav2 behavior tree *Navigate To Pose With Replanning and Recovery* is used for trajectory planning and control, continuously updating the path and executing motion commands until all user-defined waypoints are reached. The system continuously monitors the modules, triggers recovery behaviors in case of failures, and applies safety mechanisms, including timeout limits and stall detection. Upon reaching waypoints, predefined actions are executed, such as activating visual and auditory signals.

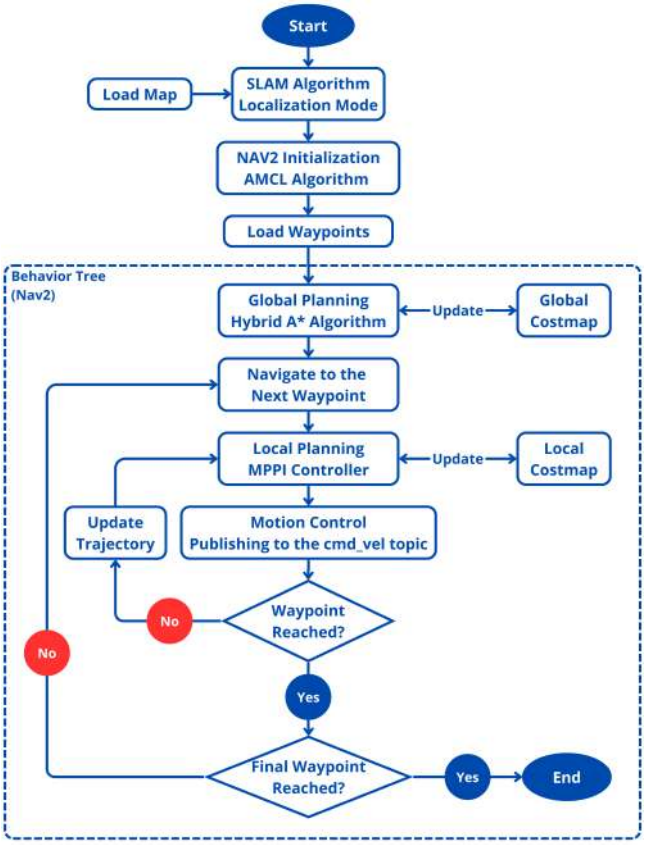


Fig. 9. Illustrative flowchart of the vehicle's autonomous operation.

IX. Results and Discussion

The main results of the autonomous vehicle tests are presented below. The experiments were conducted in the hallways on the first floor of Building 7 at UNOESC, Joaçaba, SC, Brazil, in a scaled environment simulating typical competition navigation conditions. Figure 10 shows the environment cost map overlaid on the map generated with the SLAM Toolbox during Manual Mode.

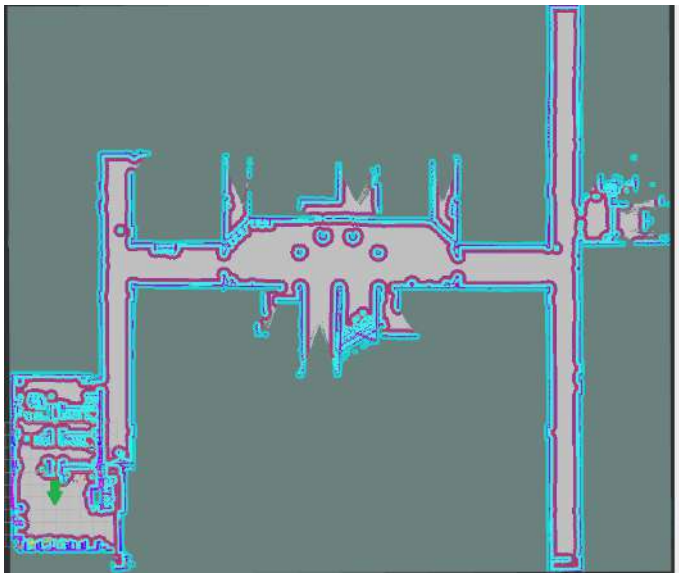


Fig. 10. Environment cost map.

The map origin point $([x, y, \theta] = [-38.3, -3.63, 0])$ is indicated by a green arrow representing the robot's initial pose. The environment is represented as an occupancy grid map, with a resolution of 0.05 m/cell, whose spatial extent is approximately 45×54 m, corresponding to the building layout. The map contains two main layers: the static layer, which describes fixed structures of the environment, and the inflation layer, which expands obstacle regions within a radius of $0.55\,\mathrm{m}$ to ensure safe navigation.

Figure 11 shows the initial route planned by the navigation system, highlighted in red, between the reference points in the previously constructed environment map.

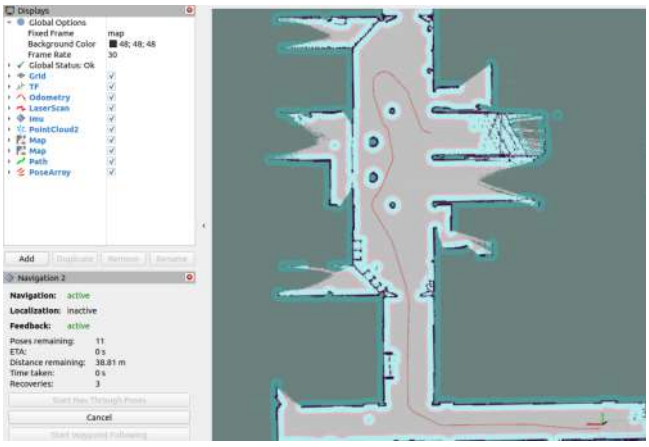


Fig. 11. Initial route planned by Nav2 between the reference points.

Figure 12 shows the route followed by the robot in Autonomous Mode, from the position $(-2.65, 4.96, 0.0)$ to $(-16.04, 25.29, 0.0)$, recorded from accumulated odometry data and pose estimates. The actual trajectory follows the planned route and passes through the reference points, with minor deviations caused by MPPI controller corrections, localization uncertainty, recovery actions, and a slight initial offset from the ideal path.

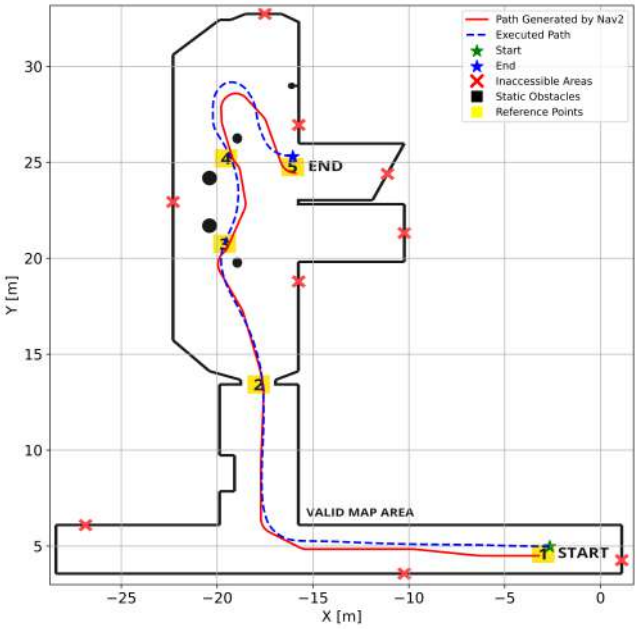


Fig. 12. Comparison between the planned route and the route executed by the robot.

Table I presents the performance results comparing the planned route with the trajectory executed by the robot.

TABLE I
TEST PERFORMANCE — COMPARISON BETWEEN THE PLANNED ROUTE AND THE EXECUTED TRAJECTORY

METRIC	VALUE	UNIT
RMSE Position	3.785	m
MAE Position	3.137	m
Distance traveled (actual)	44.924	m
Distance traveled (planned)	38.815	m
Trajectory efficiency	86.4	%
Average linear speed	0.365	m/s
Maximum linear speed (configured)	2.000	m/s
Total time	166.64	s
Mean heading change	1.214	°
Max heading change	10.344	°
Mean angular velocity	13.077	°/s

The executed trajectory resulted in a trajectory efficiency of 86.4%, indicating that approximately 13.6% of the motion corresponds to deviations from the planned path. These deviations can be attributed to continuous replanning, obstacle avoidance maneuvers, and trajectory smoothing performed by the local controller. The total navigation time of 166.64 s was influenced not only by the traveled distance but also by initialization procedures and recovery behaviors. Additionally, the relatively low average linear speed indicates a conservative control strategy focused on stability and safety in indoor environments.

The positional errors, with an RMSE of 3.785 m and an MAE of 3.137 m, reflect spatial deviations between the planned and executed trajectories, primarily caused by odometry drift, localization uncertainty, and control-induced adjustments. Although differences in trajectory length may indirectly contribute to these deviations, the errors remained bounded and did not compromise navigation performance.

The heading metrics indicate smooth trajectory evolution, with a mean change of $1.214°$ and a maximum variation of $10.344°$, associated with localized corrections during sharper maneuvers. The average angular velocity of $13.077°/s$ suggests moderate rotational motion consistent with controlled navigation in constrained environments.

X. CONCLUSION

This work presented the development of a 1/10-scale autonomous vehicle designed for Trekking PRO competitions, integrating hardware and software through the ROS middleware. The system demonstrated robust performance in constructing 2D maps of large-scale indoor environments composed of multiple interconnected areas using SLAM algorithms, enabling reliable localization throughout the exploration process. Furthermore, the control system was capable of generating feasible and smooth trajectories and autonomously following them between predefined points in a static environment.

The models, calculations, and parameters employed in the vehicle's localization, planning, and control algorithms were validated considering the platform's geometry and kinematic constraints. Additionally, the integration of sensors and software packages was successfully achieved, adhering to the standards defined by REP specifications and ROS guidelines.

REFERENCES

[1] Marius Misaros, Ovidiu Stan, Ionut-Catalin Donca, and Liviu-Cristian Miclea. "Autonomous Robots for Services — State of the Art, Challenges, and Research Areas". In: *Sensors* 23 (May 2023), p. 4962. DOI: 10.3390/s23104962.

[2] Jenay M. Beer, Arthur D. Fisk, and Wendy A. Rogers. "Toward a framework for levels of robot autonomy in human-robot interaction". In: *Journal of Human-Robot Interaction* 3.2 (July 2014), pp. 74–99. DOI: 10.5898/JHRI.3.2.Beer.

[3] Steven Macenski, Tully Foote, Brian Gerkey, Chris Lalancette, and William Woodall. "Robot Operating System 2: Design, architecture, and uses in the wild". In: *Science Robotics* 7.66 (May 2022). ISSN: 2470-9476. DOI: 10.1126/scirobotics.abm6074.

[4] RoboCore. *Trekking - Regras*. https://events.robocore.net/. 2023.

[5] Mordechai Ben-Ari and Francesco Mondada. "Robotic Motion and Odometry". In: *Elements of Robotics*. Cham: Springer, 2018, pp. 63–93. ISBN: 978-3-319-62532-4. DOI: 10.1007/978-3-319-62533-1_5.

[6] Gamini Dissanayake, Shoudong Huang, Zhan Wang, and Ravindra Ranasinghe. “A review of recent developments in Simultaneous Localization and Mapping”. In: *2011 6th International Conference on Industrial and Information Systems, ICIIS 2011 - Conference Proceedings* (Aug. 2011). DOI: 10.1109/ICIINFS.2011.6038117.

[7] Kurt Konolige, Giorgio Grisetti, Rainer Kümmerle, Wolfram Burgard, Benson Limketkai, and Regis Vincent. “Efficient Sparse Pose Adjustment for 2D mapping”. In: *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2010, pp. 22–29. DOI: 10.1109/IROS.2010.5649043.

[8] F. Lu and E. Milios. “Globally Consistent Range Scan Alignment for Environment Mapping”. In: *Autonomous Robots* 4.4 (1997), pp. 333–349. DOI: 10.1023/A:1008854305733.

[9] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, and John J. Leonard. “Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age”. In: *IEEE Transactions on Robotics* 32.6 (2016), pp. 1309–1332. DOI: 10.1109/TRO.2016.2624754.

[10] Andrea Censi. “An ICP variant using a point-to-line metric”. In: *2008 IEEE International Conference on Robotics and Automation*. 2008, pp. 19–25. DOI: 10.1109/ROBOT.2008.4543181.

[11] Steve Macenski and Igor Jambrecic. *slam_toolbox*. https://github.com/SteveMacenski/slam_toolbox. SLAM Toolbox for lifelong mapping and localization in potentially massive maps with ROS. 2021.

[12] Steve Macenski and Ivona Jambrecic. “SLAM Toolbox: SLAM for the dynamic world”. In: *Journal of Open Source Software* 6.61 (2021), p. 2783. DOI: 10.21105/joss.02783. URL: https://doi.org/10.21105/joss.02783.

[13] S. Agarwal et al. *Ceres Solver*. http://ceres-solver.org. Google Research, Accessed: 2025-09-28. June 2016.

[14] Steve Macenski, Francisco Martín, Ruffin White, and Jonatan Ginés Clavero. “The Marathon 2: A Navigation System”. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2020. URL: https://github.com/ros-planning/navigation2.

[15] Steve Macenski, Tom Moore, David V. Lu, Alexey Merzlyakov, and Michael Ferguson. “From the Desks of ROS Maintainers: A Survey of Modern and Capable Mobile Robotics Algorithms in the Robot Operating System 2”. In: *Robotics and Autonomous Systems* 168 (Oct. 2023), p. 104493. ISSN: 0921-8890. DOI: 10.1016/j.robot.2023.104493.

[16] Navigation2 Team. *Navigation2 Documentation*. https://docs.nav2.org/. Acesso em: 5 out. 2025. 2025.

[17] Dieter Fox and Wolfram Burgard. “Markov Localization for Reliable Robot Navigation and People Detection”. In: *Lecture Notes in Computer Science* (Aug. 2000). DOI: 10.1007/10705474_1.

[18] Francisco Martín Rico, José Miguel Guerrero Hernández, Rodrigo Pérez Rodríguez, Juan Diego Peña Narváez, and Alberto García Gómez-Jacinto. *Open Source Robot Localization for Non-Planar Environments*. 2024. arXiv: 2309.12744 [cs.RO].

[19] David V. Lu, Dave Hershberger, and William D. Smart. “Layered Costmaps for Context-Sensitive Navigation”. In: *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2014, pp. 709–715. DOI: 10.1109/IROS.2014.6942636.

[20] Andrew West, Thomas Wright, Ioannis Tsitsimpelis, Keir Groves, Malcolm J. Joyce, and Barry Lennox. “Real-Time Avoidance of Ionising Radiation Using Layered Costmaps for Mobile Robots”. In: *Frontiers in Robotics and AI* 9 (2022). DOI: 10.3389/frobt.2022.862067.

[21] Lim Chee Wang, Lim Ser Yong, and M.H. Ang. “Hybrid of Global Path Planning and Local Navigation Implemented on a Mobile Robot in Indoor Environment”. In: *Proceedings of the IEEE International Symposium on Intelligent Control*. 2002, pp. 821–826. DOI: 10.1109/ISIC.2002.1157868.

[22] Grady Williams, Paul Drews, Brian Goldfain, James M. Rehg, and Evangelos A. Theodorou. “Aggressive Driving with Model Predictive Path Integral Control”. In: *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*. 2016, pp. 1433–1440. DOI: 10.1109/ICRA.2016.7487277.

[23] Max Schwenzer, Muzaffer Ay, Thomas Bergs, and Dirk Abel. “Review on Model Predictive Control: An Engineering Perspective”. In: *The International Journal of Advanced Manufacturing Technology* 117.5 (Nov. 2021), pp. 1327–1349. ISSN: 1433-3015. DOI: 10.1007/s00170-021-07682-3.

[24] Michal Minařík, Robert Pěnička, Vojtěch Vonásek, and Martin Saska. “Model Predictive Path Integral Control for Agile Unmanned Aerial Vehicles”. In: *Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2024, pp. 13144–13151. DOI: 10.1109/IROS58592.2024.10802756.

[25] Carlos Leonardo Lazzari. *Trekking 2025: Projeto real de um veículo autônomo com ROS 2 Humble*. https://github.com/CarlosLeonardo29/trekking_amr. UNOESC, 2025.

Robô Didático como Complemento Terapêutico para o Desenvolvimento Social de Crianças com Transtorno do Espectro Autista (TEA)

Bianca Camila Garnica Pericon, Ervin Emanuel Orellana Pardo
Universidad Privada del Valle - UNIVALLE
bianca.garnicapericon@alumni.univalle.edu, ervin.orellanapardo@alumni.univalle.edu
ORCID: 0009-0006-1010-4040, 0009-0008-3927-0733

***Abstract*—**This paper presents the development of a didactic robot designed as a complementary therapeutic tool for children aged 4 to 8 with Level 1 Autism Spectrum Disorder (ASD) in Cochabamba, Bolivia. The system utilizes an architecture based on ESP32 and Raspberry Pi to integrate vision, motion, and audio sensors with actuators that provide gestural and acoustic feedback. Six playful dynamics were implemented using computer vision (MediaPipe), QR codes, and generative artificial intelligence (OpenAI) via MQTT wireless communication. Field tests, conducted with four participants from the "Paulo VI" Educational Unit, employed mixed-methods data collection techniques. The results indicated improvements in participation, sustained attention, and spontaneous interaction, with the user experience being positively valued by both the families and the institution. The project demonstrates the feasibility of integrating emerging technologies to strengthen the development of social and communicative skills within the local context.

***Index Terms*—**Autism Spectrum Disorder, therapeutic robotics, didactic robot, artificial intelligence, computer vision, inclusive education, human-robot interaction.

***Resumo*—**Este artigo apresenta o desenvolvimento de um robô didático concebido como ferramenta terapêutica complementar para crianças de 4 a 8 anos com Transtorno do Espectro Autista (TEA) Grau 1 em Cochabamba, Bolívia. O sistema utiliza uma arquitetura baseada em ESP32 e Raspberry Pi para integrar sensores de visão, movimento e áudio com atuadores que proporcionam feedback gesticular e sonoro. Foram implementadas seis dinâmicas lúdicas que empregam visão computacional (MediaPipe), códigos QR e inteligência artificial generativa (OpenAI) mediante comunicação sem fio MQTT. Os testes de campo, realizados com quatro participantes da Unidade Educativa "Paulo VI", utilizaram técnicas mistas de coleta de dados. Os resultados indicaram melhorias na participação, atenção sustentada e interação espontânea, com a experiência do usuário sendo avaliada positivamente pelas famílias e pela instituição. O projeto demonstra a viabilidade de integrar tecnologias emergentes para fortalecer o desenvolvimento de habilidades sociais e comunicativas no contexto local.

***Palavras chave*—**Transtorno do Espectro Autista, robótica terapêutica, robô didático, inteligência artificial, visão computacional, educação inclusiva, interação humano-robô.

I. Introdução

O Transtorno do Espectro Autista (TEA) define-se como uma condição do neurodesenvolvimento caracterizada por déficits persistentes na comunicação e interação social, bem como por padrões repetitivos de comportamento [1]. Desde sua conceitualização inicial por Kanner, identificou-se que a reciprocidade social e a empatia são áreas de comprometimento crítico que aumentam o risco de isolamento [2]. Em nível global, a Organização Mundial da Saúde (OMS) estima uma prevalência de 1 para cada 100 crianças [3], enquanto nos Estados Unidos os números alcançam 1 em cada 36 nascimentos [4]. Na Bolívia, a situação apresenta uma lacuna significativa: estima-se que existam mais de 140.000 pessoas com TEA, mas menos de 10% conta com um diagnóstico completo, e apenas 2.046 estão registradas oficialmente perante o Ministério da Saúde e Esportes [5, 6].

Em Cochabamba, a falta de acesso a diagnósticos e o desconhecimento social limitam o fortalecimento de habilidades desde as primeiras etapas do desenvolvimento, as quais são fundamentais para a adaptação a ambientes educativos [7]. A ausência de participação em atividades grupais restringe o crescimento emocional e cognitivo, aumentando a probabilidade de comportamentos problemáticos, ansiedade e baixa autoestima [8, 9] . Diante deste cenário, surge o Zumi, um robô didático projetado para crianças de 4 a 8 anos com TEA Grau 1, concebido para complementar as terapias convencionais por meio de um ambiente seguro e motivador.

O projeto fundamenta-se na Terapia Assistida por Robôs para o Autismo (RAAT), a qual tem demonstrado benefícios no desenvolvimento de habilidades emocionais e comunicativas em ambientes controlados [10]. Especificamente, o uso de robôs zoomórficos e antropomórficos estimula a imitação e o contato visual, reduzindo a ansiedade ao proporcionar uma interação previsível e menos sobrecarregada do que a humana [11, 12]. O Zumi adota a figura de um coelho para priorizar a aceitação sensorial e emocional do paciente, servindo como uma ponte tecnológica para a reciprocidade social [12, 13].

II. Delineamento Metodológico

A. Abordagem e Tipo de Pesquisa

A investigação foi desenvolvida sob uma abordagem mista, o que permitiu uma avaliação abrangente dos resultados obtidos [11]. O componente quantitativo mediu níveis de participação, atenção sustentada e o cumprimento de instruções, enquanto a abordagem qualitativa fundamentou-se em entrevistas não estruturadas, observações diretas e na análise da percepção dos pais sobre o progresso comportamental. O estudo foi de caráter explicativo e descritivo, orientado a identificar as relações de causa e efeito derivadas do uso do robô, estabelecendo previamente uma linha de base para caracterizar as habilidades iniciais da população estudada [7, 11].

B. Métodos e Técnicas de Pesquisa

O processo desenvolveu-se mediante uma combinação de métodos teóricos e empíricos distribuídos em três fases. Na fase de desenvolvimento, aplicaram-se os métodos dedutivo e analítico para adaptar as teorias gerais da robótica educativa ao design terapêutico do Zumi. Durante os testes de campo, empregou-se a observação sistemática e participante, na qual a pesquisadora guiou as sessões para capturar reações em tempo real. Finalmente, na fase de análise, utilizou-se o método indutivo para generalizar conclusões a partir dos comportamentos documentados. As fontes de informação integraram bases primárias, como o DSM-5 [1] e a consultoria clínica da Lic. Gladys Mena Solíz, complementadas por literatura secundária especializada em Terapia Assistida por Robôs (RAAT) [6, 10].

C. População e Amostra

A população-alvo consistiu em crianças de 4 a 8 anos com diagnóstico de TEA Nível 1, etapa considerada crítica para a consolidação de habilidades sociais e comunicativas [1, 8]. A amostra foi composta por quatro estudantes da Unidade Educativa "Paulo VI" em Cochabamba, Bolívia: dois meninos e duas meninas com perfis diversos, incluindo dificuldades de atenção, hipersensibilidade sensorial e ansiedade diante de estímulos novos. Cabe destacar a inclusão excepcional de uma participante com TEA Nível 2, realizada para fins investigativos, visando avaliar a adaptabilidade do sistema diante de demandas de suporte mais intensas e documentar respostas comportamentais sob diversos estímulos [10].

D. Procedimento e Intervenção Institucional

O procedimento de intervenção estruturou-se a partir de uma fase de articulação institucional com a direção do estabelecimento e o núcleo psicopedagógico. Após a obtenção dos termos de consentimento livre e esclarecido e do estabelecimento de um vínculo de confiança com as famílias, aplicaram-se questionários pré-intervenção para documentar as capacidades iniciais. A intervenção consistiu em 12 sessões individuais programadas, nas quais foram utilizadas técnicas de pesquisa de campo para registrar a interação com o robô, finalizando com a aplicação de questionários pós-intervenção aos pais para contrastar o progresso e a eficácia das seis dinâmicas terapêuticas implementadas.

III. Projeto de Engenharia

A. Arquitetura Geral e Diagrama de Blocos

O funcionamento do sistema fundamentou-se em uma arquitetura de rede distribuída que permitiu a interação coordenada entre quatro tipos de usuários: administradores, equipe de suporte técnico, profissionais da saúde e pacientes (crianças). Como ilustrado no Diagrama de Blocos (ver Figura 1), a plataforma centralizou a gestão em um servidor web hospedado sob o protocolo HTTPS, garantindo uma comunicação bidirecional mediante o protocolo MQTT, ideal por sua leveza em dispositivos de recursos limitados [14].

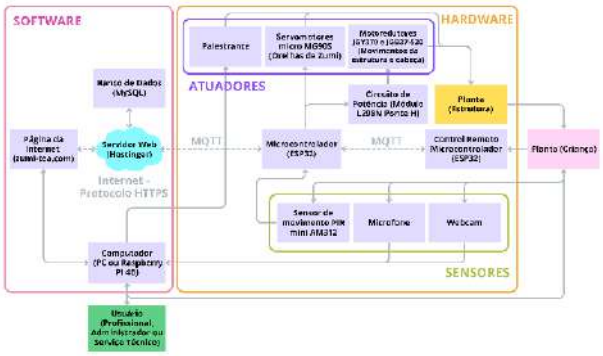


Fig. 1. Diagrama de blocos do projeto.

O fluxo de dados permitiu que o Administrador gerenciasse o banco de dados MySQL e as estatísticas de uso, enquanto o Profissional programou atividades e realizou o acompanhamento clínico. Por sua vez, o Suporte Técnico validou a operatividade dos periféricos a partir da Raspberry Pi 4B. As crianças interagiram diretamente com a planta física por meio de sensores (câmera, microfone e PIR AM312) e atuadores, fechando o laço de feedback sensorial com a interface web.

B. Projeto Mecânico e Estrutural

O sistema mecânico foi projetado sob critérios de ergonomia e aceitação sensorial [13], utilizando uma figura zoomórfica (coelho) modelada no Fusion 360, como observado na Figura 2, e fabricada mediante impressão 3D com filamento PLA+.

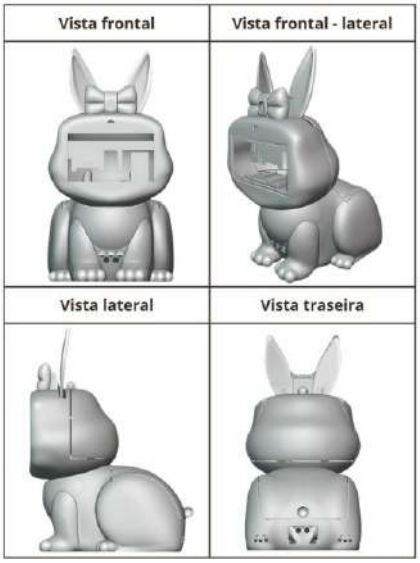


Fig. 2. Projeto da estrutura robótica em 3D.

A organização interna foi disposta em três níveis funcionais para otimizar o centro de gravidade e facilitar a manutenção, conforme observado na Figura 3.

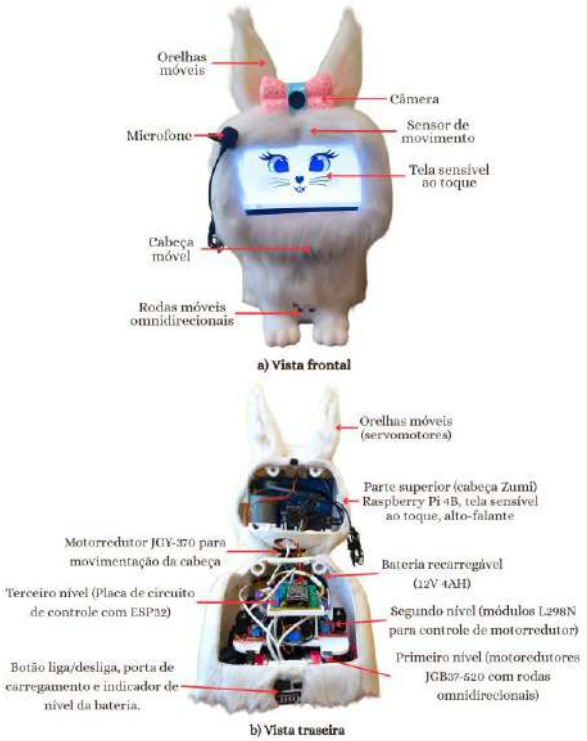


Fig. 3. Partes e disposição interna do robô finalizado.

C. Modelo Cinemático

Com o propósito de alcançar mobilidade omnidirecional, foi implementado o modelo matemático de cinemática direta para um robô com quatro rodas Mecanum omnidirecionais, o qual define a relação entre as velocidades angulares das quatro rodas Mecanum (60x30 mm) e a velocidade resultante do chassi no espaço cartesiano. Tal modelo foi testado por meio de simulações no MATLAB (ver Figura 4), utilizando a matriz de transformação cinemática para calcular as componentes de velocidade global do robô: frontal, lateral e angular [15].

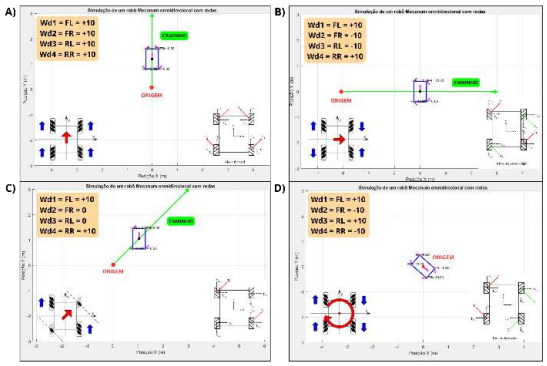


Fig. 4. Simulação de movimentos do robô omnidirecional no MATLAB. A) Movimento Frontal B) Movimento Lateral Direita C) Movimento Diagonal Direita D) Rotação Horária.

Os resultados obtidos na simulação confirmaram a precisão do sistema em diversas trajetórias, incluindo deslocamentos frontais, laterais, diagonais a 45° e rotações sobre o eixo central em ambos os sentidos. Além disso, esses testes permitiram estabelecer um período de amostragem de 0,1 s, parâmetro fundamental para garantir movimentos fluidos e estáveis durante a execução das dinâmicas terapêuticas, evitando ruídos mecânicos distratores que pudessem afetar a sensibilidade sensorial da criança [6].

D. Firmware do Sistema de Controle

O núcleo de gestão residiu em um ESP32, que atuou como a unidade de controle central do robô, cuja capacidade dual-core permitiu o gerenciamento de sensores e atuadores simultaneamente [14]. O firmware, desenvolvido no ambiente Arduino IDE, foi estruturado de forma modular para garantir a execução concorrente de tarefas críticas. O programa gerenciou a leitura de sensores, a modulação por largura de pulso (PWM) para os motores e o controle de posição dos servomotores.

A implementação incluiu bibliotecas especializadas para a conectividade Wi-Fi e o protocolo MQTT (PubSubClient), permitindo uma latência mínima na recepção de comandos. Uma característica técnica relevante foi o sistema de segurança de temporização, o qual desativava os atuadores se não fosse detectada uma nova instrução em um intervalo definido, evitando deslocamentos indefinidos. Da mesma forma, utilizou-se a biblioteca Preferences para o armazenamento não volátil na memória flash, permitindo que parâmetros críticos, como ângulos das orelhas e velocidades máximas, persistissem após reinicializações do sistema.

E. Sistema de Controle Remoto

Desenvolveu-se um periférico de controle independente baseado em um ESP32 de 30 pinos (ver Figura 5). Diferentemente dos sistemas de supervisão, este dispositivo foi concebido para potencializar o senso de agência e a autonomia da criança, permitindo que ela interaja diretamente com a estrutura física por meio de um joystick analógico e botões de ação rápida.

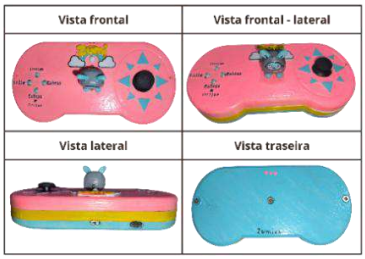


Fig. 5. Diferentes vistas do controle remoto.

O sistema processou leituras analógicas (0-4095) mapeadas para uma faixa de 0-255 para o controle de velocidade. Essa configuração facilitou que a criança controlasse não apenas o deslocamento omnidirecional, mas também a gesticulação das orelhas e a rotação da cabeça. O feedback visual foi gerenciado por um LED RGB, que notificava o estado da conexão com o broker MQTT, garantindo que o usuário tivesse ciência do estado operacional do sistema durante a atividade.

F. Interface Gráfica de Usuário (GUI)

A interface de usuário foi implementada como uma plataforma multiplataforma desenvolvida em HTML, PHP e JavaScript, hospedada em um servidor Hostinger sob protocolos HTTPS com o domínio “zumi-tea.com”. O sistema integrou um ecossistema de APIs e serviços, incluindo MediaPipe para visão computacional, utilizando o modelo de poses que detecta 33 pontos-chave do corpo e o modelo de mãos que identifica 21 pontos por mão [15], Amazon Polly para síntese de voz natural [16] y GPT-4o-mini p ara a geração de diálogos adaptativos e amigáveis [17].

A persistência das informações foi organizada em um banco de dados de seis tabelas principais utilizando MySQL, sob um esquema entidade-relacionamento que vinculou os profissionais aos seus respectivos pacientes e atividades programadas.

A arquitetura de comunicação sem fio fundamentou-se na troca de mensagens através do tópico "zumi/topic", utilizando a biblioteca Paho MQTT para assegurar conexões WebSocket estáveis a partir do navegador. Esta infraestrutura tecnológica permitiu que o robô funcionasse como um sistema integrado capaz de reportar telemetria e receber instruções complexas com uma confiabilidade de 96% em condições de rede moderadas.

O design da GUI foi estruturado para quatro perfis de usuário com permissões hierárquicas:

1) **Administradores:** Gerenciaram o registro de usuários (Profissionais, Suporte Técnico e Crianças), a programação de atividades e as contas armazenadas no banco de dados.
2) **Profissionais:** Gerenciaram o registro de pacientes, a programação de atividades e a análise de estatísticas de progresso.
3) **Suporte Técnico:** Acessou ferramentas de diagnóstico para validar a integridade de sensores e atuadores diretamente da Raspberry Pi 4B.
4) **Crianças:** Utilizaram uma interface simplificada com botões circulares e métodos de acesso amigáveis baseados em iconografia (animais ou cores).

G. Desenvolvimento de Dinâmicas Lúdico-Terapêuticas

Sob a consultoria da Psicóloga Gladys Mena Solíz, foram executadas seis dinâmicas baseadas em algoritmos de visão computacional e processamento de linguagem natural:

1) **Sentindo com o Zumi:** Implementou a biblioteca html5-qrcode para que a criança escaneasse pictogramas de emoções vinculados à personagem "Sofía", facilitando o reconhecimento emocional.
2) **Contando com o Zumi:** Empregou o MediaPipe Hands para detectar dedos levantados de 1 a 10, estimulando a coordenação mão-olho.
3) **Zumi Diz:** Utilizou o MediaPipe Pose para validar a imitação de 10 posturas corporais, trabalhando a motricidade grossa e a atenção.
4) **Conversando com o Zumi:** Integrou a API da OpenAI (GPT-4o-mini) e Amazon Polly. Ofereceu modalidades de conversação livre e resolução de problemas, adaptando as respostas de forma acolhedora via engenharia de prompts.
5) **Estátua! (Congelado):** Utilizou o sensor PIR AM312 da cabeça para monitorar o movimento durante períodos de silêncio musical, estimulando o autocontrole e a atenção sustentada.
6) **Movimente-se com o Zumi:** A criança imitou os deslocamentos do robô. O sistema analisou mudanças na altura do torso e posição horizontal via MediaPipe para validar a orientação espacial.

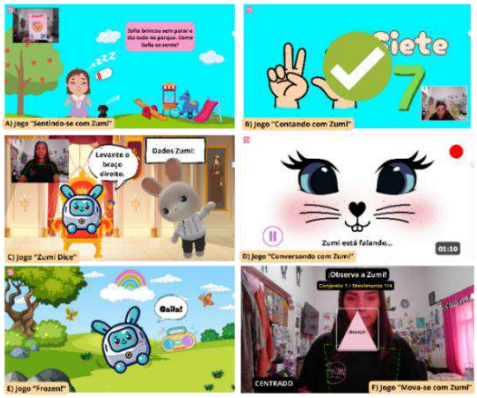


Fig. 6. Capturas de tela das atividades lúdico-terapêuticas desenvolvidas.

H. Custos de Desenvolvimento e Manutenção Operacional

O custo total de desenvolvimento do projeto foi de 4.578,46 Bs, o que representa um investimento de 656,88 USD, calculados sob a taxa de câmbio de 6,97 Bs vigente durante a aquisição de suprimentos em 2024 na Bolívia. Este montante incluiu componentes eletrônicos, filamentos 3D, manufatura de PCBs, modelagem 3D, hospedagem web e domínio institucional. Quanto à sustentabilidade operacional, o design modular e o uso de materiais de fabricação aditiva permitem que a manutenção seja eficiente e de baixo custo, facilitando a reposição de peças estruturais perante eventuais desgastes. Da mesma forma, a arquitetura baseada em hardware aberto reduz a manutenção técnica principalmente a atualizações periódicas de software e firmware na unidade de processamento central (Raspberry Pi), assegurando a continuidade do serviço terapêutico com um investimento mínimo a longo prazo.

IV. Resultados e Discussões

A. Desempenho Técnico e Funcionalidade do Protótipo

Previamente às sessões de interação, realizaram-se testes de estresse no hardware e software para garantir a estabilidade do sistema. Os componentes sensoriais (sensor PIR AM312, câmera web e microfone) mostraram uma taxa de resposta ideal sob condições de iluminação e ruído controladas na Unidade Educativa "Paulo VI". A integração do protocolo MQTT permitiu uma latência imperceptível na execução de comandos da interface web para o ESP32, garantindo que o feedback gesticular (movimento de orelhas e cabeça) fosse coerente com o progresso dos jogos, fator-chave na Terapia Assistida por Robôs [10, 18].

Verificou-se que o uso de rodas Mecanum, testado previamente em simulações no MATLAB, permitiu uma navegação fluida no ambiente educativo, o que foi crítico para manter o interesse visual das crianças sem gerar ruídos mecânicos distratores. Essa estabilidade técnica foi a base que permitiu ao grupo alcançar uma pontuação inicial média de 61,7 na primeira sessão (ver Figura 7), demonstrando que a confiabilidade do hardware influencia diretamente na disposição inicial da criança em relação à terapia.

B. Análise do Desenvolvimento de Habilidades Sociais e Comunicativas

O impacto da intervenção foi avaliado mediante a normalização das pontuações em uma escala de 100 pontos e fichas de acompanhamento com escalas de avaliação personalizadas. Como detalhado na Tabela I, o grupo de estudo mostrou uma tendência ascendente constante em todas as dimensões.

TABELA I
PROGRESSO MÉDIO DE HABILIDADES SOCIAIS E COGNITIVAS

Dimensão Avaliada	Pontuação Inicial Média	Pontuação Final Média	Melhora Percentual
Iniciativa Comunicativa e Interativa	48.3	95.0	+96.6 %
Participação e Engajamento	58.3	93.7	+60.7 %
Atenção Sustentada	58.3	91.7	+57.1%
Seguimento de Instruções	62.5	95.8	+53.3%
Reação Emocional	65.6	90.6	+38.1%

É notável que a Iniciativa Comunicativa e Interativa apresentou o incremento mais significativo de todo o estudo, com uma melhora percentual de 96,6%, passando de uma pontuação de 48,3 para 95,0. Esse avanço é atribuído à mediação do modelo GPT-4o-mini [17], que permitiu aos participantes transitar de uma comunicação reservada para uma iniciativa espontânea em direção ao robô. Esses resultados coincidem com estudos prévios que sugerem que robôs com comportamento previsível atuam como catalisadores para reduzir a ansiedade social [6, 10].

A evolução cronológica do grupo, representada na Figura 7, evidencia um progresso sustentado: após um crescimento acelerado nas primeiras três sessões (de 61,7 para 84,4), o grupo alcançou um platô de consolidação a partir da sessão 9, mantendo pontuações superiores a 93,8.

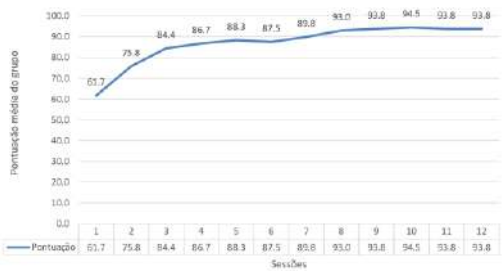


Fig. 7. Pontuação média do grupo ao longo das sessões de intervenção.

C. Avaliação da Motricidade e do Autocontrole

As dinâmicas lúdicas baseadas em visão computacional demonstraram a eficácia dos algoritmos do MediaPipe. Em jogos como "Contando com o Zumi" e "Zumi Diz", o sistema alcançou níveis de precisão superiores a 90% na detecção de poses e gestos manuais [15]. Essa precisão técnica foi fundamental para que o Seguimento de Instruções alcançasse uma média final de 95,8 (Tabela I), uma vez que o reforço positivo imediato do robô fortaleceu o aprendizado por associação.

Quanto ao autocontrole, o jogo Estátua! (Congelado), apoiado pelo sensor PIR AM312, permitiu documentar uma melhora progressiva na inibição motora e na atenção sustentada [8]. Embora inicialmente tenham sido observadas dificuldades devido à curiosidade gerada, a Figura 7 reflete como a adaptação à dinâmica permitiu que as pontuações do grupo se estabilizassem acima dos 93 pontos nas etapas finais, alcançando uma sincronização quase total com os estímulos musicais.

D. Experiência do Usuário e Percepção Familiar

A experiência do usuário foi avaliada como "altamente satisfatória" pela instituição e pelas famílias. Nos questionários de satisfação direcionados às crianças, obteve-se uma classificação unânime de satisfação máxima (100/100), ressaltando que as atividades foram percebidas como brincadeiras e não como intervenções clínicas sob pressão, o que se reflete na dimensão de Participação e Engajamento, que registrou uma melhora de 60,7% (Tabela I).

A percepção dos pais, coletada nas pesquisas pós-intervenção, coincidiu no fato de que o robô cativou a atenção das crianças de maneira mais eficaz do que outros dispositivos tecnológicos convencionais. No entanto, a análise qualitativa revelou que o sucesso da terapia depende também da comunicação transdisciplinar; em casos onde a percepção familiar divergiu da observação escolar, identificou-se a importância de integrar relatórios automatizados por meio da plataforma web para alinhar as expectativas de progresso.

V. CONCLUSÃO

Os resultados deste estudo piloto sugerem que o robô "Zumi" atua como um catalisador para reduzir a ansiedade social, oferecendo um ambiente seguro e previsível que facilita a exploração de novas formas de comunicação. A previsibilidade do comportamento robótico permite que a criança minimize a sobrecarga sensorial, funcionando como uma ponte tecnológica para interações humanas mais complexas no futuro. A inclusão de um participante com TEA Nível 2 indicou que o sistema é resiliente e adaptável, embora demande maior personalização e o uso estratégico de suportes visuais para abordar fenômenos específicos como a ecolalia. A sinergia entre a inteligência artificial generativa e a visão computacional não apenas superou as limitações de ferramentas estáticas, mas também favoreceu a atenção sustentada e o senso de agência, permitindo que a criança assuma um papel ativo em seu próprio processo terapêutico.

Esta arquitetura técnica e de gestão web apresenta-se como um modelo promissor para democratizar o acesso a tecnologias assistivas em centros de educação especial na região. Como trabalhos futuros, recomenda-se a realização de estudos longitudinais que avaliem a generalização e a retenção de habilidades a longo prazo em contextos cotidianos, além do fortalecimento da tríade terapêutica — criança, profissional e família — por meio de sistemas de relatórios automatizados que facilitem o acompanhamento transdisciplinar. No âmbito técnico, propõe-se otimizar a cinemática omnidirecional e refinar os modelos de linguagem para a neurodiversidade, consolidando soluções escaláveis e éticas para a inclusão educativa integral.

Referências

[1] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. American Psychiatric Association, 2013. doi: 10.1176/appi.books.9780890425596.

[2] L. Kanner, "THE CONCEPTION OF WHOLES AND PARTS IN EARLY INFANTILE AUTISM," *American Journal of Psychiatry*, vol. 108, no. 1, pp. 23–26, Jul. 1951, doi: 10.1176/ajp.108.1.23.

[3] Organización Mundial de la Salud (OMS), "Autismo." Accessed: Jan. 23, 2025. [Online]. Available: https://www.who.int/es/news-room/fact-sheets/detail/autism-spectrum-disorders

[4] G. N. Soke, M. J. Maenner, D. Christensen, M. Kurzius-Spencer, and L. A. Schieve, "Prevalence of Co-occurring Medical and Behavioral Conditions/Symptoms Among 4- and 8-Year-Old Children with Autism Spectrum Disorder in Selected Areas of the United States in 2010," *J. Autism Dev. Disord.*, vol. 48, no. 8, pp. 2663–2676, Aug. 2018, doi: 10.1007/s10803-018-3521-1.

[5] BoliviaVerifica, "Ángeles con Esperanza, la iniciativa que aporta a mejorar la calidad de vida de personas autistas." Accessed: Jan. 23, 2025. [Online]. Available: https://boliviaverifica.bo/angeles-con-esperanza-la-iniciativa-que-aporta-a-mejorar-la-calidad-de-vida-de-personas-autistas/

[6] Ministerio de Salud y Deportes Bolivia, "Día Mundial de Concienciación sobre el Autismo," Facebook. Accessed: Jan. 24, 2025. [Online]. Available: https://www.facebook.com/SaludDeportesBo/posts/descubre-el-mundo-a-trav%C3%A9s-de-sus-ojos-%C3%BAnicos-en-bolivia-est%C3%A1n-registradas-2046-/832514925585457/

[7] X. B. Lin, C. G. Lim, and T.-S. Lee, "Social Deficits or Interactional Differences? Interrogating Perspectives on Social Functioning in Autism," *Front. Psychiatry*, vol. 13, Apr. 2022, doi: 10.3389/fpsyt.2022.823736.

[8] The University of Kansas, "Social difficulties in autism spectrum disorder." Accessed: Jan. 23, 2025. [Online]. Available: https://educationonline.ku.edu/community/social-difficulties-in-autism-spectrum-disorder

[9] R. Ellingsen, C. Bolton, and E. Laugeson, "Evidence-Based Social Skills Groups for Individuals with Autism Spectrum Disorder Across the Lifespan," 2017, pp. 343–358. doi: 10.1007/978-3-319-62995-7_20.

[10] J.-J. Cabibihan, H. Javed, M. Ang, and S. M. Aljunied, "Why Robots? A Survey on the Roles and Benefits of Social Robots in the Therapy of Children with Autism," *Int. J. Soc. Robot.*, vol. 5, no. 4, pp. 593–618, Nov. 2013, doi: 10.1007/s12369-013-0202-2.

[11] C. Yáñez *et al.*, "Uso terapéutico de robótica en niños con Trastorno del Espectro Autista," *Andes Pediatrica*, vol. 92, no. 5, p. 747, Nov. 2021, doi: 10.32641/andespediatr.v92i5.2500.

[12] B. Szymona *et al.*, "Robot-Assisted Autism Therapy (RAAT). Criteria and Types of Experiments Using Anthropomorphic and Zoomorphic Robots. Review of the Research," *Sensors*, vol. 21, no. 11, p. 3720, May 2021, doi: 10.3390/s21113720.

[13] J. Schell, *The Art of Game Design*. A K Peters/CRC Press, 2014. doi: 10.1201/b17723.

[14] EMQX Platform Docs, "Connect via Eclipse Paho Java." Accessed: Jan. 24, 2025. [Online]. Available: https://docs.emqx.com/en/cloud/latest/connect_to_deployments/java_sdk.html

[15] H. Taheri, B. Qiao, and N. Ghaeminezhad, "Kinematic Model of a Four Mecanum Wheeled Mobile Robot," *Int. J. Comput. Appl.*, vol. 113, no. 3, pp. 6–9, Mar. 2015, doi: 10.5120/19804-1586.

[16] Google AI for Developers, "Guía de soluciones de MediaPipe." Accessed: Jan. 23, 2025. [Online]. Available: https://ai.google.dev/edge/mediapipe/solutions/guide?hl=es-419

[17] Amazon Web Services, "Amazon Polly Developer Guide," 2025, Accessed: Sep. 21, 2025. [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://docs.aws.amazon.com/pdfs/polly/latest/dg/polly-dg.pdf

[18] OpenAI *et al.*, "GPT-4 Technical Report," Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.08774

[19] B. Scassellati, Henny Admoni, and M. Matarić, "Robots for Use in Autism Research," *Annu. Rev. Biomed. Eng.*, vol. 14, no. 1, pp. 275–294, Aug. 2012, doi: 10.1146/annurev-bioeng-071811-150036.

Autores



Bianca Camila Garnica Pericon é Engenheira Biomédica graduada pela Universidad Privada del Valle - UNIVALLE de Cochabamba, Bolívia (2025). Desde 2023, é membro ativo da Sociedade de Estudantes de Engenharia Biomédica (SCEIB) da UNIVALLE e faz parte da UMA, a unidade de pesquisa tecnológica da universidade. Seu projeto de pesquisa — vinculado ao presente artigo — foi reconhecido com o Primeiro Lugar como Melhor Projeto Internacional na FETIN 2024 (INATEL, Brasil), Primeiro Lugar em Inovação Transformadora nos prêmios KAMAY 2025 da Federação de Entidades Empresariais Privadas de Cochabamba, e o Prêmio de Excelência em Inovação Tecnológica (UNIVALLE, 2024). É cofundadora da Wexi, plataforma dedicada à economia circular de componentes eletrônicos na Bolívia, e sua atuação profissional visa à integração de soluções biomédicas avançadas e à promoção da inclusão por meio da engenharia na saúde.



Ervin Emanuel Orellana Pardo é Engenheiro Biomédico, com formação em instrumentação médica, imagens médicas, processamento de sinais biomédicos, biomecânica, inteligência artificial aplicada à saúde e engenharia clínica. Desde 2023, é membro ativo da Sociedade Científica de Estudantes de Engenharia Biomédica (SCEIB), participando de atividades acadêmicas e de pesquisa. Possui experiência no desenvolvimento de páginas web e na aplicação de tecnologias voltadas para a área da saúde. Além disso, é cofundador da Wexi, uma plataforma focada na economia circular de componentes eletrônicos na Bolívia, orientada para a reutilização tecnológica e a redução de resíduos eletrônicos.

Sistema Inteligente de Monitoramento de Apiários na Bolívia Baseado em IA/ML com IoT e LPWAN

Edzel Alejandro Escobar Cueto, Patricia N Flores-Ponce, Eynar Calle Viles, Edgar R Ramos Silvestre Universidad Privada del Valle
ece2018016@est.univalle.edu, fpp1006073@est.univalle.edu, ecallevi@univalle.edu, eramoss@univalle.edu
https://orcid.org/0009-0002-4220-6472 , https://orcid.org/0009-0003-0638-0390, https://orcid.org/0000-0003-1477-2813, https://orcid.org/0000-0001-7583-5852

***Abstract*— This paper presents the design and implementation of a collaborative and distributed artificial intelligence system based on the Internet of Things (IoT) and machine learning for intelligent environmental monitoring. The proposed architecture integrates distributed sensors, low-power wide-area networks (LPWAN), and computer vision models to collect, process, and analyze environmental data in real time, enabling continuous information flow to an intelligent analytics platform.**

The system incorporates cooperative AI mechanisms in which IoT nodes operate as autonomous agents capable of sensing environmental conditions, sharing data, and supporting distributed decision-making through a swarm intelligence approach. Deep learning algorithms based on YOLOv5 are employed for visual pattern recognition and behavioral analysis, while supervised learning techniques are used for predicting critical events.

The architecture was validated through a beekeeping case study, achieving 94% accuracy in bee detection under variable environmental conditions. Results demonstrate the effectiveness of distributed AI for remote monitoring systems and highlight its potential for broader applications in smart environmental management.

***Index Terms*— *Artificial Intelligence, machine learning, distributed intelligence, deep learning*.**

***Resumo*— Este artigo apresenta o projeto e a implementação de um sistema de inteligência artificial colaborativa e distribuída baseado na Internet das Coisas (IoT) e em aprendizado de máquina para monitoramento inteligente de ambientes naturais. A arquitetura proposta integra sensores distribuídos, redes de longa distância e baixo consumo de energia (LPWAN) e modelos de visão computacional para coletar, processar e analisar dados ambientais em tempo real, permitindo um fluxo contínuo de informações para uma plataforma inteligente de análise.**

O sistema incorpora mecanismos de IA cooperativa, nos quais os nós IoT atuam como agentes autônomos capazes de perceber o ambiente, compartilhar informações e apoiar processos de tomada de decisão distribuída por meio de um modelo de inteligência de enxame. Algoritmos de *deep learning* baseados em YOLOv5 são utilizados para reconhecimento de padrões visuais e análise comportamental, enquanto técnicas de aprendizado supervisionado auxiliam na predição de eventos críticos.

A arquitetura foi validada em um estudo de caso aplicado à apicultura, alcançando 94% de acurácia na detecção de abelhas em condições ambientais variáveis.

***Palavras-Chave*— *Inteligência Artificial, aprendizado de máquina, inteligência distribuída, deep learning*.**

I. Introdução

A apicultura é uma atividade essencial para a segurança alimentar e a estabilidade ecológica devido ao papel fundamental das abelhas na polinização [2]. No entanto, as práticas tradicionais de monitoramento baseiam-se em inspeções manuais que demandam muito tempo, são imprecisas e reativas às mudanças ambientais [4][5].

A inteligência artificial (IA) evoluiu de sistemas centralizados para arquiteturas colaborativas e distribuídas, nas quais múltiplos agentes compartilham informações e cooperam para resolver problemas complexos [1][9]. Nesse contexto, a integração da IA com redes de sensores IoT e comunicação LPWAN permite o desenvolvimento de sistemas capazes de aprender, adaptar-se e tomar decisões em tempo real, mesmo em ambientes com recursos limitados ou conectividade reduzida [10][11].

II. Projeto do Sistema

O sistema tem como objetivo otimizar a gestão apícola por meio da integração de dispositivos IoT, redes de comunicação de longo alcance e algoritmos de visão computacional. Sua finalidade é coletar, analisar e visualizar dados ambientais e biológicos das colmeias em tempo real, garantindo a saúde e a produtividade das colônias de abelhas.

A. Arquitetura Geral do Sistema

O sistema está estruturado em cinco módulos funcionais:

1. **Módulo de aquisição de dados:** composto por sensores de temperatura, umidade, peso e áudio, além de uma câmera de vídeo responsável por registrar o fluxo de abelhas na entrada da colmeia. Esses dados são gerenciados por um Raspberry Pi 5, que atua como o nó principal de processamento.

2. **Módulo de transmissão:** utiliza o protocolo LoRaWAN sob a tecnologia LPWAN, implementado com placas Heltec HTCC-AB01,

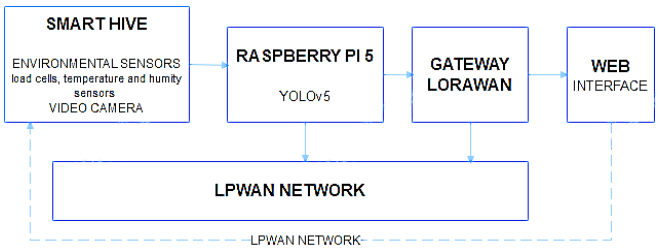
permitindo comunicação de longo alcance com baixo consumo de energia [7].

3. **Módulo de gerenciamento de rede:** um *gateway* LoRaWAN centraliza as informações recebidas dos nós e as transmite a um servidor remoto utilizando o protocolo MQTT [11].

4. **Módulo de armazenamento e processamento:** os dados são recebidos por um *Virtual Private Server* (VPS) configurado com um *broker* EMQX e o painel de controle Vesta Control Panel, garantindo a segurança, integridade e disponibilidade das informações.

5. **Módulo de interface web:** desenvolvido em PHP, JavaScript e Node.js, apresenta os dados por meio de dashboards interativos, gráficos dinâmicos e alertas automatizados acessíveis a partir de qualquer dispositivo conectado à Internet [6].

Figura 1. Diagrama de funcionamento do sistema, mostrando o comportamento operacional e a forma como os dados são coletados para a execução da tarefa definida pela unidade de processamento



B. *Integração de inteligência Artificial*

O sistema implementa um modelo de visão computacional baseado no YOLOv5, treinado com imagens capturadas de colmeias em condições reais por meio da plataforma Roboflow. Técnicas de *data augmentation* foram aplicadas para melhorar a capacidade de generalização do modelo, alcançando um mAP@0.5 de 0,94, precisão de 0,95 e recall de 0,90, assegurando desempenho confiável na detecção de abelhas e na análise de padrões comportamentais.

O modelo foi treinado diretamente no Raspberry Pi 5, explorando suas capacidades de processamento e a otimização proporcionada pelo framework PyTorch. Essa abordagem possibilita a execução local (*edge computing*), sem depender integralmente da conectividade com a Internet, garantindo continuidade operacional mesmo em áreas rurais com cobertura limitada.

C. *Fluxo Operacional do Sistema*

O sistema opera por meio de cinco etapas principais:

1. **Captura de dados:** sensores e a câmera coletam variáveis ambientais e visuais.
2. **Pré-processamento:** os dados são filtrados e formatados localmente.
3. **Análise por IA:** o modelo YOLOv5 realiza detecções no fluxo de vídeo, contabilizando abelhas e analisando sua atividade.
4. **Transmissão de dados:** métricas resumidas são enviadas ao servidor central por meio de MQTT/LPWAN.
5. **Visualização e alertas:** a interface web permite que os apicultores visualizem dados históricos e recebam notificações sobre eventos críticos, como variações de temperatura, perda de abelhas ou anomalias na colmeia.

D. *Implementação do Sistema*

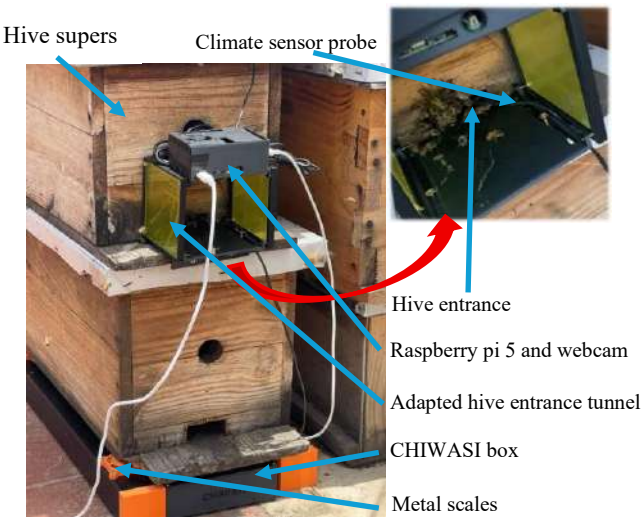
Durante os testes de campo realizados no apiário "Flores", em Cochabamba, conforme apresentado na Figura 2, o sistema demonstrou desempenho estável, com uma taxa de transmissão de dados bem-sucedida superior a 95% e uma acurácia de detecção visual de abelhas de 94%.

Os testes de calibração do sensor de peso apresentaram uma regressão linear com erro médio de **3%**, validando sua confiabilidade para a medição da produção de mel.

O sistema também registrou frequências sonoras entre 258 Hz e 441 Hz, associadas a estados de tranquilidade ou alerta nas abelhas, permitindo correlacionar o comportamento da colônia com as condições ambientais e a atividade interna da colmeia.

Figura 2. Implementação do sistema na colmeia para monitoramento e posicionamento da câmera na entrada da colmeia para a contagem de abelhas.

III. RESULTADOS



A. *Desenvolvimento do Protótipo*

O protótipo foi desenvolvido seguindo uma abordagem metodológica iterativa que incluiu:

- **Análise e projeto:** definição de requisitos em colaboração com apicultores e especialistas, juntamente com o desenvolvimento da arquitetura do sistema [6].

- **Implementação de Hardware:** Montagem dos nós IoT equipados com sensores e módulos de comunicação LPWAN, integrando microcontroladores responsáveis pelo processamento local dos dados.

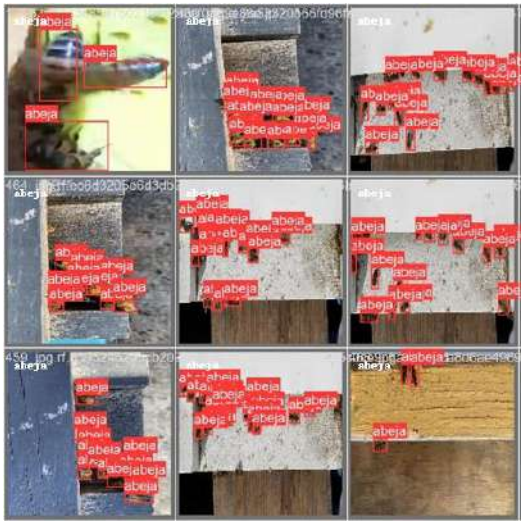
- **Desenvolvimento do Software:** Programação de algoritmos de machine learning e estratégias de inteligência de enxame, juntamente com a implementação de uma interface web para visualização em tempo real das informações coletadas.

B. Testes e Avaliação do Sistema

Os testes realizados em laboratório e em campo demonstraram que:

- rede LPWAN possibilitou transmissão contínua de dados sem perdas, mesmo em ambientes com conectividade limitada [7]. O sistema alcançou com sucesso um alcance de comunicação de 4,6 km entre o *Gateway* instalado e o nó de monitoramento (colmeia), mesmo em área urbana do município de Cercado, garantindo a estabilidade da rede.

Figura 3. Imagens rotuladas e identificadas pelo algoritmo de reconhecimento de abelhas durante sua passagem pela entrada da colmeia, permitindo a contagem e o rastreamento do número de abelhas que entram e saem.



- Os algoritmos de inteligência artificial atingiram elevados índices de acurácia na detecção de anomalias, permitindo respostas antecipadas a condições adversas [8][9]. Conforme apresentado na Figura 3, as imagens rotuladas e identificadas representam as saídas do algoritmo de IA e sua aplicação nos apiários.

- A integração do sistema de contagem e identificação por meio de visão computacional permitiu estimar com precisão a população ativa de abelhas na colmeia e favoreceu uma tomada de decisão mais robusta frente às mudanças ambientais dinâmicas [6]. O registro diário do tráfego de abelhas apresentou 3.600 entradas e 3.633 saídas, evidenciando uma diferença de 33 abelhas que não retornaram à colmeia. Essa variação é consistente com cenários típicos de perda natural, mortalidade ou dispersão, demonstrando a utilidade do sistema para monitoramento contínuo do comportamento populacional e detecção de possíveis anomalias em tempo real.

C. Análise da Matriz de Confusão

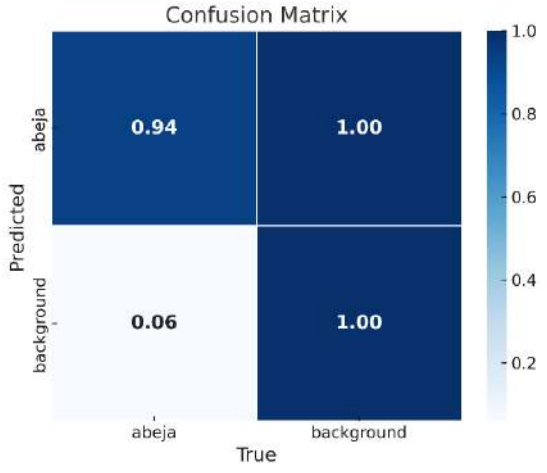
A matriz de confusão obtida durante o processo de validação do modelo YOLOv5 fornece informações relevantes sobre seu desempenho na detecção da classe “bee”. Embora o modelo tenha sido treinado exclusivamente para identificar essa única classe, a presença da categoria “background” na matriz decorre da natureza inerente dos algoritmos de detecção de objetos, que também consideram regiões onde nenhum objeto é detectado como fundo.

Os resultados indicam que o modelo apresenta 94% de taxa de detecção correta de abelhas e 6% de casos em que abelhas foram incorretamente classificadas como background. Por outro lado, todas as instâncias correspondentes ao background foram corretamente identificadas, refletindo 100% de precisão para essa categoria.

Esse comportamento sugere que, embora o modelo apresente desempenho robusto, ainda existem situações em que a identificação das abelhas pode ser comprometida, possivelmente devido a condições ambientais ou à variabilidade das imagens utilizadas no treinamento. Portanto, recomenda-se revisar as amostras do *dataset* e otimizar o processo de rotulagem para reduzir essas classificações incorretas.

A Figura 4 apresenta a matriz de confusão normalizada do modelo.

Figura 4. Matriz de confusão normalizada do modelo, utilizada para identificar as variáveis detectadas durante o processo de reconhecimento de imagens por meio da câmera.

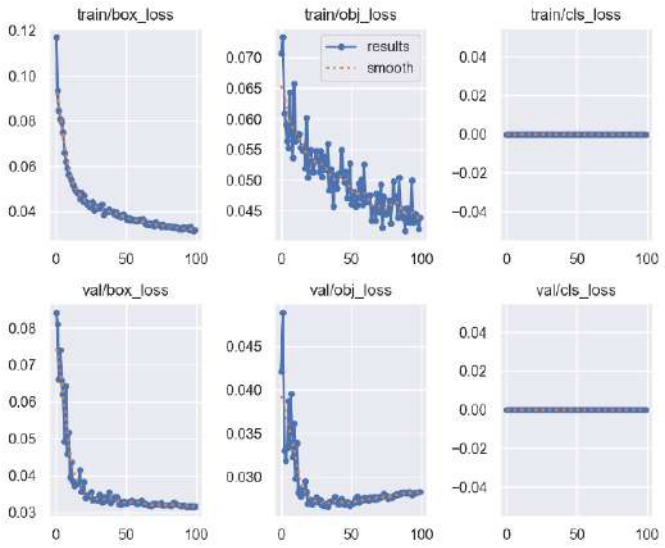


D. Análise das Funcoes de Perda

As funções de perda indicam os erros de predição do modelo durante as fases de treinamento e validação, conforme

apresentado na Figura 5. As seguintes perdas foram analisadas:

Figura 5. Gráficos das funções de perda, apresentando o comportamento das diferentes funções de erro do modelo.



- **Bounding Box Loss (train/box_loss and val/box_loss):** Essa perda mede o erro na predição das coordenadas das bounding boxes. Tanto a perda de treinamento quanto a de validação apresentaram tendência consistente de redução, estabilizando-se em torno de 0,03, o que indica aprendizado eficiente na localização precisa das abelhas.

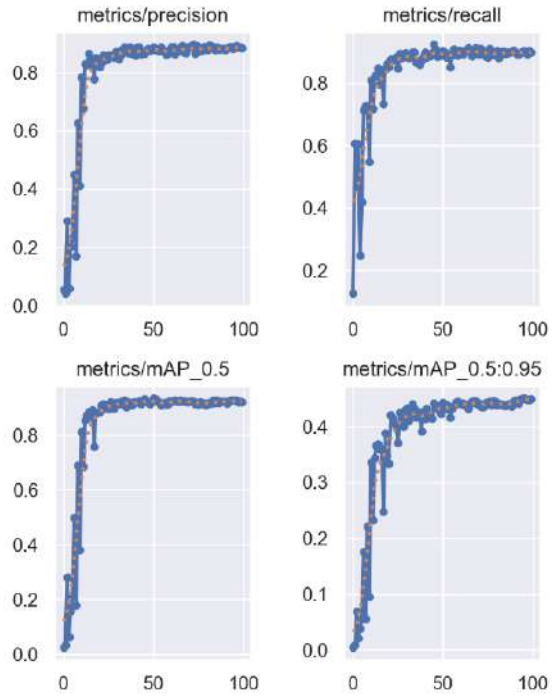
- **Objectness Loss (train/obj_loss and val/obj_loss):** Avalia a capacidade do modelo de detectar a presença de objetos dentro das *bounding boxes*. Essa perda diminuiu progressivamente nos dados de treinamento e validação, indicando que o modelo se tornou eficaz em distinguir entre presença e ausência de abelhas.

- **Classification Loss (train/cls_loss and val/cls_loss):** Como o conjunto de dados continha apenas uma única classe (**"bee"**), a perda de classificação permaneceu **igual a zero durante todo o treinamento**, confirmando que o modelo aprendeu efetivamente a única classe sem ambiguidades de classificação.

E. Métricas de Desempenho

As seguintes métricas foram utilizadas para avaliar a capacidade de detecção do modelo:

Figura. 6. Métricas de desempenho do modelo de detecção de objetos baseadas em *precisão*, *recall* e *mAP*.



- **Precision (metrics/precision):** A *precisão* mede a proporção de detecções verdadeiramente positivas em relação ao total de predições positivas. Essa métrica aumentou rapidamente e estabilizou-se próxima de 0,9, indicando alta taxa de predições corretas.

- **Recall (metrics/recall):** O *recall* mede a proporção de positivos reais que foram corretamente identificados. Esse indicador também atingiu aproximadamente 0,9, demonstrando a eficiência do modelo em detectar a maioria das abelhas presentes no conjunto de dados.

- **Mean Average Precision (metrics/mAP_0.5):** Essa métrica, calculada com limiar IoU = 0,5, estabilizou-se em torno de 0,9, refletindo o alto desempenho do modelo na detecção de objetos sob um critério moderado de sobreposição.

- **Mean Average Precision (metrics/mAP_0.5:0.95):** Essa métrica mais rigorosa, que considera limiares IoU entre 0,5 e 0,95, estabilizou-se em aproximadamente 0,45. Embora inferior ao mAP_0.5, esse valor é considerado aceitável e indica desempenho razoável mesmo sob critérios mais estritos de sobreposição.

A figura 6 apresenta os gráficos das métricas de desempenho do modelo de detecção de objetos. Neste sistema, o modelo detecta as abelhas que entram e saem da colmeia.

IV. CONCLUSÕES

A implementação do sistema integrado de IoT com Inteligência Artificial e tecnologia LPWAN demonstrou desempenho altamente eficiente durante os testes realizados

no apiário Flores, em Cochabamba, onde o sistema operou de forma contínua e estável em condições reais de campo.

O modelo de visão computacional baseado em YOLOv5 alcançou 94% de acurácia na detecção de abelhas, com mAP@0.5 próximo de 0,94, validando a robustez do modelo mesmo sob variações de iluminação, movimento e distância.

No que se refere à transmissão de dados, os módulos LoRaWAN com Heltec HTCC-AB01 mantiveram uma taxa de sucesso superior a 95% a uma distância de 4,6 km entre o ponto de instalação do *Gateway* (utilizando The Things Network – TTN para gerenciamento) e o nó de monitoramento (colmeia), mesmo em área urbana com presença de edificações. Essa infraestrutura permitiu a transmissão confiável de dados ambientais, como temperatura, umidade, peso e áudio, para o servidor VPS configurado com EMQX.

O sensor de peso calibrado apresentou erro médio de 3%, situando-se dentro de faixas aceitáveis para monitoramento da produção de mel. De forma semelhante, o sistema de áudio registrou frequências entre 258 Hz (estado normal de atividade) e 441 Hz (estado de alerta, estresse ou mudanças abruptas) no interior da colmeia, possibilitando correlacionar o comportamento biológico com as condições ambientais.

A plataforma web hospedada no VPS possibilitou a visualização dinâmica de dados históricos, a geração de alertas automáticos e a supervisão remota em tempo real, contribuindo significativamente para a melhoria da tomada de decisão pelos apicultores.

De forma geral, o sistema respondeu de maneira rápida, estável e eficiente, demonstrando que a integração de IA com nós IoT e redes LPWAN não apenas reduz a necessidade de inspeções manuais e o estresse nas abelhas, mas também fornece um módulo preditivo capaz de antecipar anomalias.

Esse avanço tecnológico é particularmente relevante no contexto da apicultura boliviana, um setor em crescimento que ainda depende amplamente de métodos tradicionais de monitoramento, o que limita a detecção precoce de doenças, a avaliação da produtividade e a gestão preventiva de riscos ambientais.

Como sistemas de monitoramento em tempo real baseados em tecnologias inteligentes ainda não existem no país, este projeto representa um marco pioneiro e uma base sólida para a transição rumo a uma apicultura digitalizada, sustentável e orientada por dados. Seus resultados constituem uma contribuição inovadora para a Bolívia, abrindo caminho para futuras pesquisas, a implantação de redes de apiários inteligentes e a integração de ferramentas avançadas, como análise acústica e monitoramento multi-colmeias em larga escala.

Este trabalho demonstra que a convergência entre Inteligência Artificial, IoT e LPWAN não apenas otimiza a gestão apícola, mas também estabelece um precedente para o desenvolvimento de soluções tecnológicas inovadoras que impulsionem o futuro da apicultura na Bolívia.

AGRADECIMENTOS

Os autores agradecem à Universidad Privada del Valle pelo apoio institucional, bem como aos tutores e pesquisadores que contribuíram para o desenvolvimento do projeto. Este trabalho foi inspirado pelo esforço conjunto de estudantes, docentes e apicultores locais, que tornaram possível a validação experimental do sistema.

REFERÊNCIAS

[1] Eskov, A., “The Origin and Organization of the Bee Colony Apis mellifera”, 2020.
[2] Manual Apícola Tapacarí, *Manual Apícola Tapacarí: Técnicas y Prácticas en Apicultura Moderna*, Tapacarí, Bolivia, 2021.
[3] Pérez, “Anatomía y Fisiología de la Abeja: Un Estudio Integral”.
[4] Matriz-FC-Apicultores, *Matriz de los Apicultores de Bolivia*, 2015.
[5] Artículo de Opinión Bolivia, “Crecimiento de la Apicultura en Cochabamba”, Voces de la Llajta, 2023.
[6] Calle, E. et al., “Implementación de un Sistema Integral IoT para el Monitoreo Avanzado de Apiarios Incorporando Inteligencia Artificial y Tecnología LPWAN”, Tesis de Grado, Universidad Privada del Valle, 2024.
[7] Raza, S., Kulkarni, P., y Sooriyabandara, M., “Low Power Wide Area Networks: An Overview”, *IEEE Communications Surveys & Tutorials*, 2017.
[8] Kennedy, J., y Eberhart, R., “Particle Swarm Optimization”, en *Proceedings of the IEEE International Conference on Neural Networks*, 1995.
[9] Blum, C., “Swarm Intelligence: Introduction and Applications”, en *Swarm Intelligence*, 2005.
[10] Li, S., Xu, L. D., y Zhao, S., “The Internet of Things: A Survey”, *Information Systems Frontiers*, 2015.
[11] Adelantado, F., et al., “Understanding the Limits of LoRaWAN”, *IEEE Communications Magazine*, 2017.



Eng. Edzel Alejandro Escobar Cueto Engenheiro de Telecomunicações graduado pela Universidad Privada del Valle (UNIVALLE), onde também concluiu o Diplomado em Desenvolvimento Web Full Stack, fortalecendo suas competências em tecnologias de software e desenvolvimento de soluções digitais.
Possui experiência na área de comercialização de tecnologia FTTH, insumos, ferramentas e equipamentos especializados na empresa Bolcompany Ltda., atuando na gestão comercial e técnica de soluções voltadas ao setor de telecomunicações. Além disso, possui conhecimentos em Internet das Coisas (IoT), inteligência artificial e sistemas de comunicação de longo alcance e baixo consumo energético, aplicados ao desenvolvimento de soluções tecnológicas inovadoras.
No campo da pesquisa, foi vencedor do Prêmio de Pesquisa CTI 2025, na categoria Biodiversidade e Mudança Climática – Jovens Pesquisadores, com o projeto *“Implementação de um sistema integral IoT para o monitoramento avançado de apiários incorporando inteligência artificial e tecnologia LPWAN”*.
Atualmente atua como representante comercial da Bolcompany Ltda., integrando sua formação técnica com o desenvolvimento de soluções tecnológicas aplicadas.



Eng. Patricia Nataly Flores Ponce, Engenheira Biomédica graduada pela Universidad del Valle, Cochabamba, Bolívia. Sua formação acadêmica está orientada ao desenvolvimento de tecnologias aplicadas à área biomédica, com ênfase em instrumentação eletrônica, sistemas inteligentes e soluções baseadas na Internet das Coisas (IoT).

Realizou estudos de formação complementar em docência em educação superior com enfoque por competências, voltados à inovação educacional e à formação universitária em áreas tecnológicas. Seu trabalho acadêmico concentra-se na pesquisa e no desenvolvimento de dispositivos tecnológicos para monitoramento e treinamento na área da saúde.

É autora de trabalhos de pesquisa relacionados a sistemas inteligentes aplicados à medicina, destacando o desenvolvimento de um sistema inteligente de treinamento em reanimação cardiopulmonar (RCP) com monitoramento remoto, baseado em sensores, comunicação IoT e análise de dados em tempo real. Suas pesquisas foram apresentadas em eventos acadêmicos internacionais e são direcionadas ao uso de tecnologias emergentes para aprimorar os processos de formação e assistência em saúde.

Atualmente desenvolve projetos de pesquisa em tecnologias biomédicas, inteligência artificial aplicada a sistemas de monitoramento e dispositivos médicos inteligentes, além de participar de iniciativas acadêmicas e científicas voltadas à inovação tecnológica no campo da engenharia biomédica.



Ms. C. Eynar Calle Viles é mestre em Telecomunicações, com especialização em Redes de Comunicações e Telemática, e possui pós-graduação em Internet das Coisas (IoT) pelo Instituto Nacional de Telecomunicações (INATEL), Brasil.

Com mais de 10 anos de experiência, é especialista em projeto e implementação de redes, FTTH, IoT e sistemas embarcados. Desenvolveu diversas inovações tecnológicas, incluindo propostas de patentes, como o “Robô Móvel para Detecção de Metais” e o “Sistema de Localização para o Satélite Tupac Katari”.

Além disso, possui várias publicações e participações em revistas e congressos internacionais. Suas principais áreas de interesse incluem antenas e propagação, segurança da informação e IoT, contribuindo ativamente para o desenvolvimento de soluções tecnológicas inovadoras.

Atualmente, atua como coordenador de pesquisa na Faculdade de Informática e Eletrônica da Universidad Privada del Valle (Univalle) e é presidente da IEEE Computer Society Bolivia.



M. Sc. Edgar Roberto Ramos Silvestre. Engenheiro Eletrônico graduado pela Universidad Mayor de San Simón (UMSS), Cochabamba – Bolívia. Mestre em Tecnologia de Computação pelo CIDETEC – IPN (Centro de Inovación y Desarrollo Tecnológico en Cómputo – Instituto Politécnico Nacional), Cidade do México, México. Possui diplomados em Educação Virtual e Docência Universitária pela Universidad del Valle. Atualmente cursa o Doutorado em Educação Superior na Universidad del Valle (UNIVALLE).

Realizou cursos em Robótica Aplicada no CNAD (Centro Nacional de Actualización Docente), Cidade do México, México; Supervisão e Automação Totalmente Integrada no SENA (Servicio Nacional de Aprendizaje), Cali – Colômbia; e IoT Applied to Smart Home no INATEL (Instituto Nacional de Telecomunicações), Brasil. Também possui certificações em Inteligência Artificial (HCIA AI) e Internet das Coisas (HCIA IoT).

Atua como docente em universidades públicas e privadas, participando de projetos de pesquisa e publicando artigos em revistas nacionais e internacionais. Atualmente é diretor do Departamento de Eletrônica e Telecomunicações da Faculdade de Informática e Eletrônica da Universidad del Valle.