



UFOP

Universidade Federal
de Ouro Preto

**Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Computação e Sistemas**

Sistema de Detecção Passiva de Presença Humana Baseado em Wi-Fi CSI com ESP32-S3 para Ambientes Internos

Lucas Brazzarola de Lima

TRABALHO DE CONCLUSÃO DE CURSO

ORIENTAÇÃO:

Samuel Souza Brito

COORIENTAÇÃO:

Igor Muzetti Pereira

Julho, 2026

João Monlevade—MG

Lucas Brazzarola de Lima

**Sistema de Detecção Passiva de Presença
Humana Baseado em Wi-Fi CSI com ESP32-S3
para Ambientes Internos**

Orientador: Samuel Souza Brito

Coorientador: Igor Muzetti Pereira

Monografia apresentada ao curso de Engenharia de Computação do Instituto de Ciências Exatas e Aplicadas, da Universidade Federal de Ouro Preto, como requisito parcial para aprovação na Disciplina “Trabalho de Conclusão de Curso II”.

Universidade Federal de Ouro Preto

João Monlevade

Julho de 2026

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

L732s Lima, Lucas Brazzarola de.

Sistema de detecção passiva de presença humana baseado em Wi-Fi CSI com ESP32-S3 para ambientes internos. [manuscrito] / Lucas Brazzarola de Lima. - 2026.

73 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Samuel Souza Brito.

Coorientador: Prof. Dr. Igor Muzetti Pereira.

Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia de Computação .

1. Aprendizado do computador. 2. Microcontroladores. 3. Redes de computadores. 4. Sensoriamento remoto. 5. Sistemas embarcados (Computadores). 6. Telecomunicações. 7. Teoria da Informação. I. Brito, Samuel Souza. II. Pereira, Igor Muzetti. III. Universidade Federal de Ouro Preto. IV. Título.

CDU 621.396:004.85

Bibliotecário(a) Responsável: Flavia Reis - CRB6/2431



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS EXATAS E APLICADAS
DEPARTAMENTO DE COMPUTAÇÃO E SISTEMAS



FOLHA DE APROVAÇÃO

Lucas Brazzarola de Lima

Sistema de Detecção Passiva de Presença Humana Baseado em Wi-Fi CSI com ESP32 para Aplicações em Ambientes Clínicos

Monografia apresentada ao Curso de Engenharia de Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Engenharia de Computação

Aprovada em 31 de julho de 2026

Membros da banca

Dr. Samuel Souza Brito - orientador - Universidade Federal de Ouro Preto
Dr. Igor Muzetti Pereira - coorientador - Universidade Federal de Ouro Preto
Dra. Janniele Aparecida Soares Araújo - avaliadora - Universidade Federal de Ouro Preto
Dr. Theo Silva Lins - avaliador - Universidade Federal de Ouro Preto

Samuel Souza Brito, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 24/08/2026



Documento assinado eletronicamente por **Samuel Souza Brito, PROFESSOR DE MAGISTERIO SUPERIOR**, em 24/08/2026, às 16:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1173679** e o código CRC **CDBE8EB9**.

Este trabalho é dedicado a você, Fábio. Que sua memória permaneça não apenas na saudade, mas também no propósito de buscar formas de cuidado capazes de evitar que outras famílias enfrentem uma perda semelhante.

Agradecimentos

Agradeço, primeiramente, à minha família, que sustentou esta caminhada de diferentes maneiras: à minha mãe, que enfrentou um de seus maiores medos ao me ver sair de casa para seguir meu próprio caminho e, com sua coragem, tornou possível o início desta jornada; ao meu pai, que sempre me deu forças para continuar, especialmente nos momentos em que o cansaço e as dificuldades fizeram a conclusão desta etapa parecer distante; e à minha namorada, que acreditou que eu conseguiria mesmo quando nem eu acreditava mais.

À República Cativoiro, pelo acolhimento, pelas amizades, por ter sido meu lar e pelo apoio nos momentos que inspiraram este trabalho, mas também naqueles em que, após a pandemia, retornar à universidade parecia já não fazer sentido. Vocês fizeram parte não apenas da minha formação acadêmica, mas também de quem me tornei durante esses anos.

Aos meus orientadores, que acompanharam minha luta para conciliar o desenvolvimento deste trabalho com as responsabilidades profissionais e que, ainda assim, sempre se mostraram compreensivos, disponíveis e dispostos a contribuir. A orientação e a confiança de vocês foram fundamentais para que este trabalho pudesse ser concluído.

Aos professores do ICEA que levaram a sério a responsabilidade de ensinar, que me mostraram por que tantos desistem de se formar em Engenharia da Computação, mas também me mostraram o caminho para que eu pudesse estar entre aqueles que chegam ao fim.

Por fim, agradeço a todos que, de alguma forma, estiveram presentes ao decorrer desta (longa e extensa) caminhada.

“Any sufficiently advanced technology is indistinguishable from magic.”

— Arthur C. Clarke,
Profiles of the Future.

Resumo

Este trabalho investiga o uso de informações do estado do canal de uma comunicação Wi-Fi para detectar passivamente a presença humana e distinguir três condições de ocupação em um ambiente interno: ambiente vazio, presença estática e presença em movimento. O objetivo foi desenvolver e avaliar uma solução de baixo custo, sem câmeras e sem dispositivos vestíveis, adequada a situações nas quais o monitoramento sem contato e sem formação de imagens seja desejável. O sistema utiliza dois microcontroladores ESP32-S3, um responsável pela criação da rede e pela geração periódica de tráfego UDP e outro pela captura das amostras de *Channel State Information* (CSI), posteriormente encaminhadas ao computador por comunicação serial. Foram implementados mecanismos de aquisição, validação e armazenamento dos dados, além de um pipeline composto por pré-processamento, segmentação em janelas temporais, extração e seleção de características, treinamento de classificadores e inferência contínua com consolidação temporal das decisões. O conjunto de dados foi formado por 300 aquisições válidas, distribuídas igualmente entre as três classes, duas sessões e cinco zonas do ambiente, totalizando 149.141 pacotes CSI. Todas as aquisições envolvendo presença humana foram realizadas com um único participante. A avaliação empregou separação estratificada por arquivo em dez divisões, mantendo todas as janelas provenientes de uma mesma aquisição no mesmo subconjunto e ajustando as operações dependentes dos dados somente com os arquivos de treinamento. Entre os modelos comparados, foi selecionado o *Extra Trees* com 100 árvores e 160 características. O modelo apresentou acurácia média de $0,8433 \pm 0,0237$ e Macro F1 de $0,8402 \pm 0,0242$. A presença em movimento alcançou F1 de 0,9318, enquanto a presença estática obteve 0,7455, sendo frequentemente confundida com o ambiente vazio. Em uma sequência contínua, a consolidação temporal elevou a acurácia de 0,8914 para 0,9140. Os resultados demonstram a viabilidade da utilização de um único enlace Wi-Fi com dispositivos ESP32-S3 para monitoramento sem contato e sem formação de imagens no cenário experimental adotado. Entretanto, a dificuldade de diferenciar presença estática de ambiente vazio, a utilização de um único participante e ambiente, além da dependência da configuração de aquisição empregada, limitam a extensão dos resultados para outras condições sem nova avaliação ou calibração.

Palavras-chave: sensoriamento Wi-Fi. *Channel State Information*. ESP32-S3. detecção de presença. aprendizado de máquina.

Abstract

This work investigates the use of channel state information from Wi-Fi communication for passive human-presence detection and for distinguishing three occupancy conditions in an indoor environment: an empty environment, static presence, and movement. The objective was to develop and evaluate a low-cost solution without cameras or wearable devices, suitable for situations in which contactless and image-free monitoring is desirable. The system employs two ESP32-S3 microcontrollers: one is responsible for creating the network and periodically generating UDP traffic, while the other captures the Channel State Information (CSI) samples and forwards them to a computer through serial communication. Mechanisms for data acquisition, validation, and storage were implemented, as well as a pipeline comprising preprocessing, temporal window segmentation, feature extraction and selection, classifier training, and continuous inference with temporal decision consolidation. The dataset consisted of 300 valid acquisitions equally distributed among the three classes, two collection sessions, and five zones of the environment, totaling 149,141 CSI packets. All acquisitions involving human presence were performed with a single participant. The evaluation employed ten file-based stratified splits, keeping all windows originating from the same acquisition in a single subset and fitting all data-dependent operations using only the training files. Among the compared models, an Extra Trees classifier with 100 trees and 160 features was selected. The model achieved a mean accuracy of 0.8433 ± 0.0237 and a Macro F1-score of 0.8402 ± 0.0242 . Movement achieved an F1-score of 0.9318, whereas static presence obtained 0.7455 and was frequently confused with the empty environment. In a continuous sequence, temporal consolidation increased accuracy from 0.8914 to 0.9140. The results demonstrate the feasibility of using a single Wi-Fi link with ESP32-S3 devices for contactless and image-free monitoring under the adopted experimental conditions. However, the difficulty in distinguishing static presence from an empty environment, the use of a single participant and experimental environment, and the dependence on the adopted acquisition configuration limit the extension of the results to other conditions without further evaluation or calibration.

Keywords: Wi-Fi sensing. Channel State Information. ESP32-S3. presence detection. machine learning.

Lista de ilustrações

Figura 1 – Representação didática da propagação do sinal, da diferença entre RSSI e CSI e das variações associadas às condições de ocupação.	20
Figura 2 – Visão geral dos fluxos de treinamento, exportação e operação em tempo real	27
Figura 3 – Representação do ambiente e do posicionamento dos dispositivos durante as aquisições	29
Figura 4 – Formação de janelas temporais deslizantes com tamanho W e passo S .	35
Figura 5 – Fluxo de geração e armazenamento das amostras processadas	39
Figura 6 – Fluxo da validação separada por arquivo	41
Figura 7 – Arquitetura geral do sistema	48
Figura 8 – Organização da aquisição Channel State Information (CSI) por meio de fila e tarefa de saída	53
Figura 9 – Interface utilizada para visualização, diagnóstico e armazenamento das aquisições CSI	57
Figura 10 – Interface gráfica utilizada na inferência em tempo real	60
Figura 11 – Macro F1 obtido para as combinações de tamanho de janela e passo . .	63
Figura 12 – Matriz de confusão normalizada do classificador selecionado	65
Figura 13 – Comparação temporal entre o estado real, a classificação bruta e o estado consolidado	67

Lista de tabelas

Tabela 1	–	Características extraídas de cada subportadora	36
Tabela 2	–	Principais parâmetros da rede e do tráfego gerado pelo ponto de acesso	50
Tabela 3	–	Organização do quadro de amostra do protocolo CSI2	54
Tabela 4	–	Resumo do conjunto de dados utilizado nos experimentos	62
Tabela 5	–	Comparação entre os principais classificadores avaliados	63
Tabela 6	–	Desempenho do classificador para cada classe	64
Tabela 7	–	Desempenho do classificador em função da zona de coleta	66
Tabela 8	–	Efeito da consolidação temporal sobre o cenário contínuo	66

AUC-PR	Area Under the Precision-Recall Curve
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BiLSTM	Bidirectional Long Short-Term Memory
BSSID	Basic Service Set Identifier
CRC	Cyclic Redundancy Check
CSI	Channel State Information
DTR	Data Terminal Ready
ESP-IDF	Espressif IoT Development Framework
FreeRTOS	Free Real-Time Operating System
HT20	High Throughput com largura de canal de 20 MHz
HT-LTF	High Throughput Long Training Field
IP	Internet Protocol
L-LTF	Legacy Long Training Field
LSTM	Long Short-Term Memory
MAC	Media Access Control
MCS	Modulation and Coding Scheme
OFDM	Orthogonal Frequency-Division Multiplexing
RSSI	Received Signal Strength Indicator
RTS	Request to Send
STBC	Space-Time Block Coding
UART	Universal Asynchronous Receiver-Transmitter
UDP	User Datagram Protocol
WPA2-PSK	Wi-Fi Protected Access 2 with Pre-Shared Key
XGBoost	Extreme Gradient Boosting

Sumário

1	INTRODUÇÃO	14
1.1	Objetivos	15
1.2	Justificativa	15
1.3	Organização do trabalho	17
2	REVISÃO BIBLIOGRÁFICA	18
2.1	Sensoriamento humano em ambientes internos	18
2.2	Sensoriamento baseado em CSI	18
2.3	Aquisição e processamento dos dados	21
2.4	Trabalhos relacionados	23
2.5	Síntese da revisão	25
3	METODOLOGIA	26
3.1	Procedimento experimental e formação do conjunto de dados	27
3.1.1	Ambiente de aquisição	28
3.1.2	Definição das classes	29
3.1.3	Sessões de coleta	30
3.1.4	CrITÉrios de aceitação das aquisições	30
3.2	Pré-processamento dos dados	31
3.2.1	Conversão e validação dos dados	31
3.2.2	Remoção de valores atípicos	32
3.2.3	Suavização temporal	33
3.2.4	Normalização das subportadoras	34
3.2.5	Redução de redundância	34
3.2.6	Segmentação em janelas temporais	34
3.3	Extração e seleção de características	35
3.3.1	Extração das características	36
3.3.2	Construção do vetor de características	37
3.3.3	Seleção pelo Fisher Score	38
3.4	Treinamento e validação dos classificadores	38
3.4.1	Geração do conjunto de características	39
3.4.2	Classificadores avaliados	40
3.4.3	Validação separada por arquivo	40
3.4.4	Registro e organização dos experimentos	44
4	DESENVOLVIMENTO	46

4.1	Visão geral da solução desenvolvida	46
4.1.1	Requisitos e decisões de projeto	46
4.1.2	Arquitetura de hardware e software	47
4.1.3	Fluxo geral dos dados	48
4.2	Desenvolvimento do ponto de acesso	49
4.3	Desenvolvimento do receptor CSI	51
4.3.1	Conexão à rede e inicialização da captura	51
4.3.2	Aquisição dos dados CSI	51
4.3.3	Organização por filas e tarefas	52
4.3.4	Formato de transmissão serial	53
4.3.5	Confiabilidade da aquisição embarcada	54
4.4	Ferramentas de aquisição no computador	54
4.4.1	Recepção e interpretação dos quadros	55
4.4.2	Armazenamento dos dados brutos	55
4.4.3	Visualização e diagnóstico da aquisição	56
4.4.4	Organização dos arquivos coletados	56
4.5	Implementação da inferência em tempo real	57
4.5.1	Reutilização dos parâmetros treinados	58
4.5.2	Processamento contínuo das amostras	58
4.5.3	Consolidação temporal das decisões	59
4.5.4	Interface gráfica de operação	59
5	RESULTADOS	61
5.1	Caracterização do conjunto de dados	61
5.2	Seleção da configuração e do classificador	61
5.2.1	Configuração temporal	62
5.2.2	Redução dimensional e escolha do modelo	62
5.3	Desempenho do modelo selecionado	64
5.3.1	Desempenho geral e por classe	64
5.3.2	Influência da zona de coleta	65
5.4	Avaliação do sistema em tempo real	66
5.5	Limitações e dependências da configuração	67
5.5.1	Limitações experimentais	67
5.5.2	Limitações da tecnologia e da configuração utilizada	68
6	CONCLUSÃO	70
	REFERÊNCIAS	74

1 Introdução

O monitoramento da presença e do nível geral de atividade de pessoas em ambientes internos possui aplicações em segurança, automação, eficiência energética, assistência a idosos e acompanhamento de pacientes. Informações relacionadas à ocupação podem auxiliar sistemas responsáveis pela tomada de decisões em edifícios inteligentes e aplicações de saúde (SHAHBAZIAN; TRUBITSYNA, 2023). Em determinados contextos, especialmente quando o ambiente possui caráter privado ou sensível, esse acompanhamento precisa ser realizado sem produzir registros visuais desnecessários das pessoas presentes. A adoção de uma tecnologia de monitoramento deve, portanto, considerar não apenas sua confiabilidade e custo de implantação, mas também o conforto dos usuários e a preservação da privacidade.

As soluções tradicionalmente utilizadas apresentam limitações importantes. Câmeras fornecem informações detalhadas, mas dependem de iluminação, posicionamento e linha de visão, além de produzirem registros visuais que podem ser indesejados em ambientes nos quais a privacidade é um requisito relevante (KANG; KIM; PARK, 2024). Sensores vestíveis exigem contato com o corpo, alimentação e utilização correta pelo usuário, o que pode dificultar o acompanhamento contínuo (MIAO et al., 2025). Sensores infravermelhos passivos possuem menor custo, mas respondem principalmente ao movimento e podem não identificar adequadamente uma pessoa que permaneça imóvel por períodos prolongados (SHAHBAZIAN; TRUBITSYNA, 2023).

O sensoriamento baseado em radiofrequência oferece uma alternativa sem contato físico e sem formação de imagens do ambiente. A presença humana modifica a propagação dos sinais por meio de fenômenos como reflexão, difração, espalhamento e atenuação, permitindo que alterações no ambiente sejam observadas pelo receptor (SHAHBAZIAN; TRUBITSYNA, 2023). Entre as abordagens existentes, o *Wi-Fi Sensing* utiliza os próprios sinais de uma comunicação Wi-Fi como fonte de dados, aproveitando componentes amplamente disponíveis e de baixo custo. Uma das medidas empregadas nesse tipo de sistema é o *Channel State Information*, ou **CSI**, que fornece informações sobre a resposta do canal nas diferentes subportadoras da comunicação. Por preservar um nível de detalhamento superior ao indicador agregado de potência recebida, o **CSI** permite observar alterações que podem estar associadas à ausência de pessoas, à presença estática ou à movimentação no ambiente (MENDEZ et al., 2024; WANG; WU, 2024).

1.1 Objetivos

O objetivo geral deste trabalho é desenvolver e avaliar um sistema passivo de sensoriamento baseado em [CSI](#), utilizando dispositivos ESP32-S3, capaz de distinguir três condições de ocupação em um ambiente interno: ambiente vazio, presença estática e presença em movimento. A proposta busca obter informações gerais sobre a condição de ocupação do ambiente, sem reconhecer ações específicas, identificar indivíduos ou interpretar seus comportamentos. Para alcançar esse objetivo, foram definidos os seguintes objetivos específicos:

- desenvolver uma arquitetura de aquisição composta por um Emissor, responsável pela criação da rede Wi-Fi e pela geração periódica de tráfego, e por um Receptor, responsável pela captura e pelo encaminhamento das amostras [CSI](#) ao computador;
- implementar mecanismos para recepção, validação, diagnóstico e armazenamento dos dados brutos obtidos durante as aquisições;
- formar um conjunto de dados controlado contendo amostras das três condições de ocupação, obtidas em diferentes regiões do ambiente e períodos de coleta;
- desenvolver um pipeline de processamento composto por validação, pré-processamento, segmentação temporal, extração e seleção de características;
- comparar classificadores de aprendizado de máquina por meio de uma estratégia de avaliação que preserve a separação entre os arquivos utilizados no treinamento e no teste;
- integrar o classificador selecionado a uma aplicação de inferência em tempo real e avaliar o efeito da consolidação temporal sobre as previsões produzidas continuamente.

1.2 Justificativa

Em diferentes aplicações de monitoramento interno, a obtenção de informações sobre ocupação e nível geral de atividade precisa ser conciliada com a privacidade e o conforto das pessoas presentes. Essa necessidade é especialmente relevante em ambientes nos quais a formação de imagens pode expor situações, espaços ou indivíduos que não precisam ser visualmente registrados para que a condição de ocupação seja conhecida. O uso de câmeras, embora permita obter informações detalhadas, pode ser inadequado nesses contextos. Dispositivos vestíveis, por sua vez, exigem contato, alimentação e utilização correta, tornando o funcionamento dependente da adesão do usuário. Nesse cenário,

métodos capazes de acompanhar alterações gerais no ambiente sem formar imagens e sem exigir que a pessoa porte sensores apresentam interesse particular.

Aplicações clínicas e assistenciais constituem um exemplo de contexto em que essas características podem ser relevantes. [Ndebele et al. \(2024\)](#), ao avaliarem uma tecnologia de monitoramento sem contato em unidades de internação psiquiátrica, observaram que as informações produzidas contribuíram para a atuação preventiva das equipes. Embora o presente trabalho não se restrinja a esse domínio, o estudo exemplifica uma situação em que informações sobre presença e atividade podem ser úteis sem que seja necessária a formação contínua de imagens das pessoas monitoradas.

O sensoriamento baseado em sinais de radiofrequência constitui uma possibilidade para esse tipo de aplicação, pois permite observar alterações associadas à presença e ao movimento sem produzir imagens do ambiente. A utilização de um enlace Wi-Fi também possibilita empregar uma infraestrutura discreta e formada por componentes amplamente disponíveis. Neste trabalho, a distinção entre ambiente vazio, presença estática e presença em movimento foi escolhida por representar diferentes condições gerais de ocupação sem avançar para o reconhecimento de ações, comportamentos ou indivíduos. A proposta é, portanto, orientada a situações em que informações gerais de ocupação sejam suficientes e em que características como ausência de contato, preservação da privacidade e baixa complexidade de instalação sejam desejáveis.

A utilização de dispositivos da família ESP32 tornou a aquisição de [CSI](#) mais acessível por meio de plataformas compactas e de baixo custo. Essa característica favorece o desenvolvimento de sistemas que possam ser instalados sem equipamentos especializados ou alterações significativas na infraestrutura do ambiente. Entretanto, a viabilidade de uma solução desse tipo não depende apenas do desempenho do classificador. A regularidade das amostras, as perdas de pacotes e as diferenças entre plataformas de aquisição também precisam ser consideradas durante o desenvolvimento ([BISIO et al., 2025](#); [KHALIULLIN; SULIMOV; BUSSE, 2024](#)). Além disso, o [CSI](#) é sensível à posição dos dispositivos, à organização do ambiente, às interferências externas e ao período de coleta ([WANG; WU, 2024](#); [MIAO et al., 2025](#)). Essa dependência das condições de propagação indica que a aplicação do método deve considerar as características específicas de cada instalação e o procedimento utilizado para aquisição e calibração do sistema.

Dessa forma, a investigação proposta contempla não apenas a classificação das condições de ocupação, mas também a integridade da aquisição, a preservação dos dados brutos, a separação adequada entre treinamento e teste e a estabilidade das decisões produzidas continuamente. Esses aspectos permitem avaliar o funcionamento do fluxo completo, desde a recepção das amostras até a apresentação do estado estimado, mantendo uma arquitetura de baixo custo e complexidade limitada.

1.3 Organização do trabalho

Além deste capítulo introdutório, o trabalho está organizado em cinco capítulos, estruturados da seguinte forma:

- **Capítulo 2: Revisão bibliográfica.** Apresenta os fundamentos do sensoriamento humano em ambientes internos, o funcionamento do sensoriamento baseado em [CSI](#), os procedimentos de aquisição e processamento dos dados e os trabalhos relacionados à proposta.
- **Capítulo 3: Metodologia.** Descreve o procedimento experimental, a formação do conjunto de dados e as etapas utilizadas no pré-processamento, na extração e seleção de características e na avaliação dos classificadores.
- **Capítulo 4: Desenvolvimento.** Detalha a implementação da solução, incluindo o Emissor, o Receptor, as ferramentas utilizadas para aquisição e armazenamento dos dados no computador e a aplicação de inferência em tempo real.
- **Capítulo 5: Resultados.** Apresenta a caracterização do conjunto de dados, os experimentos realizados para definir a configuração do pipeline e o classificador, o desempenho do modelo selecionado e sua avaliação em uma sequência contínua.
- **Capítulo 6: Conclusão.** Retoma os principais resultados, discute as limitações observadas e apresenta possibilidades para a continuidade e o aprimoramento do sistema.

2 Revisão bibliográfica

Este capítulo apresenta os fundamentos técnicos relacionados ao sensoriamento humano por CSI. Inicialmente, o problema investigado é situado entre as tarefas de detecção de ocupação e reconhecimento de atividades. Em seguida, são discutidos a representação do canal por meio do CSI, as particularidades de sua aquisição com dispositivos ESP32 e os principais procedimentos de processamento e avaliação encontrados na literatura. Por fim, são analisados trabalhos relacionados e suas contribuições para a definição da abordagem adotada.

2.1 Sensoriamento humano em ambientes internos

Os sistemas de sensoriamento humano podem ser classificados de acordo com o tipo de informação que pretendem obter. A detecção de ocupação busca determinar a ausência ou a presença de pessoas em determinado espaço, enquanto a localização acrescenta uma estimativa da posição ocupada e o reconhecimento de atividades procura diferenciar ações ou comportamentos específicos (SHAHBAZIAN; TRUBITSYNA, 2023). Essas tarefas exigem diferentes níveis de detalhamento, pois a detecção de ocupação pode utilizar alterações gerais observadas no ambiente, ao passo que o reconhecimento de atividades depende da identificação de padrões associados a ações como caminhar, sentar ou deitar.

O problema investigado neste trabalho ocupa uma posição intermediária. Além de distinguir o ambiente vazio daquele ocupado, busca-se separar a presença predominantemente estática da presença em movimento, sem reconhecer ações específicas, identificar indivíduos ou interpretar seu comportamento. Para realizar essa distinção sem contato físico ou formação de imagens, podem ser utilizadas alterações produzidas pelo corpo humano nos sinais de radiofrequência. Entre as abordagens disponíveis, destaca-se o sensoriamento baseado em CSI, apresentado na seção seguinte (SHAHBAZIAN; TRUBITSYNA, 2023; MIAO et al., 2025).

2.2 Sensoriamento baseado em CSI

Em ambientes internos, o sinal transmitido alcança o receptor por diferentes trajetórias. Paredes, móveis, equipamentos e pessoas produzem reflexões, difrações, espalhamentos e atenuações, originando a propagação por múltiplos percursos. Como as componentes recebidas apresentam amplitudes, fases e atrasos distintos, alterações na configuração física do ambiente também modificam a resposta do canal (SHAHBAZIAN; TRUBITSYNA, 2023).

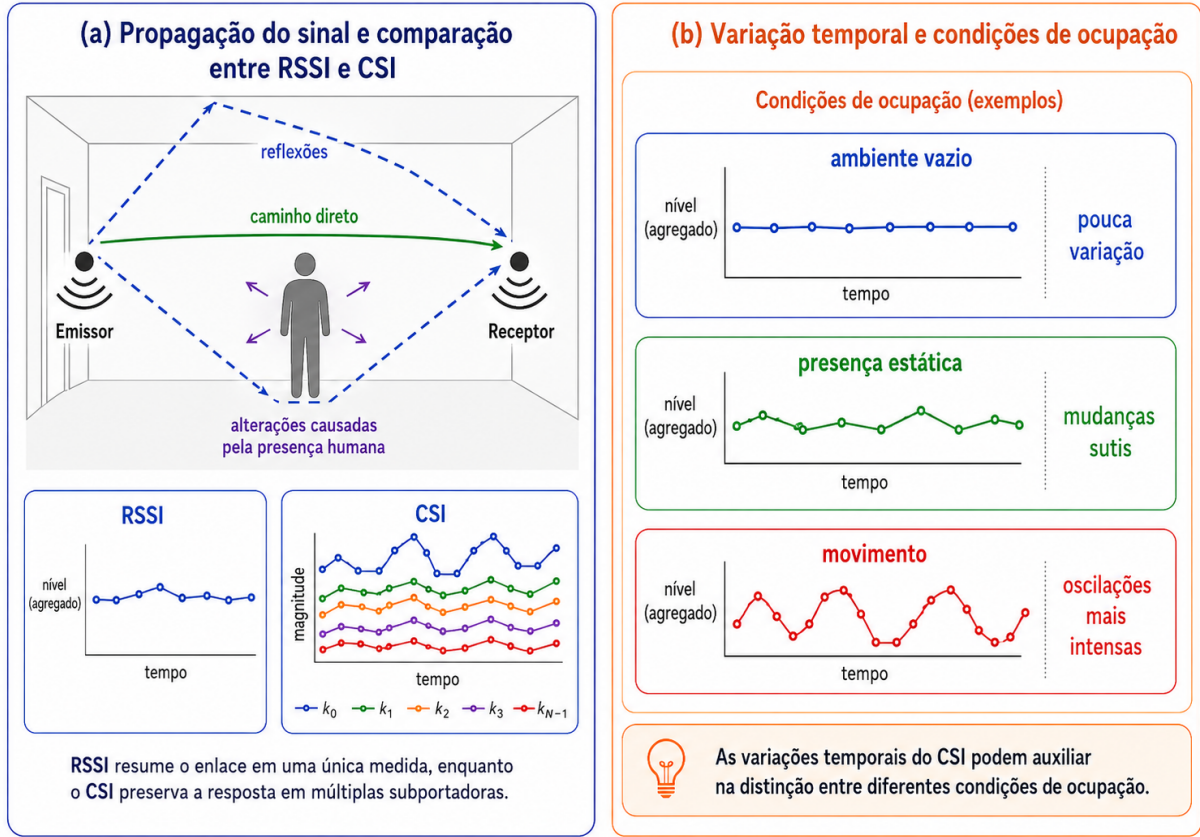
O *Received Signal Strength Indicator*, ou Received Signal Strength Indicator (**RSSI**), representa de forma agregada a intensidade com que um pacote foi recebido. Por resumir o enlace em um único valor, essa medida não preserva individualmente as diferenças existentes entre as subportadoras da comunicação.

O *Channel State Information*, ou **CSI**, fornece uma representação mais detalhada da resposta do canal nas subportadoras empregadas pela comunicação Orthogonal Frequency-Division Multiplexing (**OFDM**). Enquanto o **RSSI** descreve o enlace por meio de uma única medida de potência, o **CSI** preserva respostas associadas às diferentes subportadoras, permitindo acompanhar como distintas regiões do espectro variam ao longo do tempo. A quantidade e a organização dos coeficientes disponibilizados, entretanto, não constituem uma característica fixa para todas as redes Wi-Fi, pois dependem da plataforma de aquisição e da configuração da camada física utilizada. Aspectos como o modo de operação, a largura do canal e os campos empregados na estimação do canal podem modificar a representação obtida, fazendo com que diferentes sistemas disponibilizem vetores **CSI** com dimensionalidades distintas (**BISIO et al., 2025**).

A configuração do enlace também interfere nas condições sob as quais o **CSI** é observado. A escolha do canal determina a região do espectro utilizada pela comunicação e pode modificar tanto as condições de propagação quanto a interferência proveniente de outras redes e equipamentos. A largura do canal, por sua vez, altera a faixa de frequências ocupada e a organização das subportadoras empregadas pela transmissão. Dessa forma, canal, largura de banda e modo de operação devem ser considerados tanto na comparação entre experimentos quanto na reprodução de um sistema de sensoriamento baseado em **CSI**. Neste trabalho, esses parâmetros foram mantidos constantes durante as aquisições, e suas configurações específicas são apresentadas no Capítulo 4.

A Figura 1 apresenta uma representação simplificada desses conceitos. No painel (a), são ilustrados o caminho direto e os percursos refletidos entre o Emissor e o Receptor, bem como as alterações produzidas pela presença humana. O painel também evidencia a diferença entre a medida agregada fornecida pelo **RSSI** e as múltiplas respostas temporais preservadas pelo **CSI**. No painel (b), são exemplificados os comportamentos esperados nas condições de ambiente vazio, presença estática e presença em movimento. Essas representações são conceituais e têm como objetivo facilitar a compreensão das informações utilizadas pelo sistema.

Figura 1 – Representação didática da propagação do sinal, da diferença entre RSSI e CSI e das variações associadas às condições de ocupação.



Fonte: imagem gerada por meio do ChatGPT e revisada pelo autor, 2026.

Para cada subportadora k , o **CSI** pode ser representado pelo coeficiente complexo:

$$H_k = I_k + jQ_k, \quad (2.1)$$

em que I_k e Q_k representam, respectivamente, as componentes real e imaginária. A partir desse coeficiente, podem ser obtidas a amplitude e a fase:

$$|H_k| = \sqrt{I_k^2 + Q_k^2}, \quad (2.2)$$

$$\angle H_k = \text{atan2}(Q_k, I_k). \quad (2.3)$$

A amplitude representa a intensidade da resposta estimada em cada subportadora, enquanto a fase descreve seu deslocamento angular. A fase bruta, contudo, pode apresentar distorções decorrentes de diferenças de sincronização, frequência e temporização entre transmissor e receptor. Por esse motivo, alguns sistemas priorizam a amplitude ou aplicam procedimentos de correção antes de utilizar a fase (SHAHBAZIAN; TRUBITSYNA, 2023).

Em uma tarefa de localização interna, [Mendez et al. \(2024\)](#) observaram desempenho superior dos modelos baseados em CSI em comparação com configurações equivalentes que utilizavam somente RSSI. A vantagem foi mais expressiva nas sequências temporais, indicando que a observação individual das subportadoras preserva informações ausentes em uma medida agregada de potência.

A entrada de uma pessoa no ambiente modifica os percursos de propagação entre transmissor e receptor. Quando ela permanece predominantemente estática, a configuração do canal pode diferir daquela observada no ambiente vazio, ainda que apresente pouca variação temporal. Durante o movimento, as alterações sucessivas na geometria do canal tendem a produzir oscilações mais acentuadas nas subportadoras. Essa diferença permite utilizar medidas de distribuição, energia e variação temporal para representar as condições de ocupação ([WANG; WU, 2024](#)).

A presença estática constitui uma condição particularmente difícil de separar do ambiente vazio, pois sua identificação depende de alterações sutis que podem ser encobertas por ruídos, interferências ou mudanças não relacionadas à ocupação. Além disso, o CSI é sensível à posição dos dispositivos, à organização física do ambiente e ao período de coleta. Reduções de desempenho entre períodos distintos foram observadas por [Wang e Wu \(2024\)](#), enquanto [Miao et al. \(2025\)](#) apontam diferenças entre sessões, pessoas e ambientes como um desafio recorrente dos sistemas de sensoriamento por Wi-Fi. Essa dependência das condições de propagação indica que um modelo obtido em determinada configuração não deve ser considerado diretamente transferível para outra instalação, sendo necessário considerar procedimentos de aquisição e calibração compatíveis com o ambiente em que o sistema será utilizado.

2.3 Aquisição e processamento dos dados

O CSI é estimado internamente pelos dispositivos Wi-Fi para auxiliar na compensação dos efeitos do canal durante a recepção. O acesso a essas informações, entretanto, depende de plataformas, controladores ou firmwares capazes de disponibilizá-las ao usuário.

A placa de rede Intel 5300 tornou-se uma referência nos primeiros estudos de sensoriamento por Wi-Fi, mas sua utilização exige computadores e configurações específicas de hardware e software. Os dispositivos da família ESP32 oferecem uma alternativa compacta e de menor custo, além de permitirem maior controle sobre a geração do tráfego e a captura dos dados. As informações disponibilizadas pelas duas plataformas, contudo, não são equivalentes. [Bisio et al. \(2025\)](#) identificaram diferenças na quantidade, na resolução e na representação das amostras CSI. Além da plataforma empregada, parâmetros da comunicação, como modo de operação e largura de canal, fazem parte das condições sob as quais os coeficientes são obtidos. Dessa forma, a quantidade de posições do vetor CSI

não deve ser tratada isoladamente como uma característica universal da tecnologia, e a comparação entre trabalhos deve considerar também as configurações utilizadas durante a aquisição.

A qualidade das séries temporais também depende da regularidade com que os pacotes são recebidos. Perdas durante a comunicação reduzem a taxa efetiva de amostragem e podem produzir intervalos irregulares entre os registros. [Khaliullin, Sulimov e Busse \(2024\)](#) destacam esse problema na aquisição realizada com dispositivos ESP32, indicando a necessidade de acompanhar a continuidade e a integridade dos dados, além de armazenar os próprios vetores [CSI](#).

Antes da classificação, as amostras precisam ser validadas e organizadas. Entre as verificações possíveis estão a compatibilidade entre as componentes reais e imaginárias, o comprimento dos vetores e a identificação de registros incompletos. A origem de cada aquisição também deve ser preservada, pois arquivos obtidos em diferentes sessões, posições ou períodos podem apresentar características distintas.

Após a validação, as componentes complexas são convertidas para a representação utilizada no processamento. Nos sistemas baseados em amplitude, o cálculo é realizado a partir das componentes real e imaginária de cada subportadora, produzindo uma matriz temporal cujas linhas representam os pacotes e cujas colunas correspondem às subportadoras observadas.

Os sinais adquiridos podem conter valores atípicos, ruídos e oscilações não relacionadas às condições investigadas. Por esse motivo, a literatura apresenta diferentes combinações de filtragem e suavização. [Wang e Wu \(2024\)](#) empregaram remoção de valores anômalos e filtragem por *wavelets*, enquanto [R et al. \(2025\)](#) utilizaram um filtro passa-baixas associado a uma média móvel. Independentemente do método adotado, o processamento deve reduzir perturbações sem eliminar as variações produzidas pela presença ou pelo movimento.

A normalização permite reduzir diferenças de escala entre as subportadoras. Para evitar que informações do conjunto avaliado influenciem o processamento, seus parâmetros devem ser estimados somente com os dados destinados ao treinamento e posteriormente reutilizados nos demais subconjuntos. O mesmo cuidado se aplica a outras operações dependentes da distribuição dos dados, como seleção de subportadoras e características.

As sequências processadas podem ser divididas em janelas temporais, cada uma composta por um intervalo consecutivo de pacotes. Janelas sobrepostas permitem produzir decisões em intervalos menores do que sua duração total, mas também criam dependência entre amostras consecutivas. Essa relação deve ser considerada durante a divisão dos dados experimentais.

As janelas podem ser fornecidas diretamente a modelos que processam sequências ou

convertidas em vetores de características. Entre as medidas empregadas na literatura estão média, desvio padrão, valores extremos, energia, amplitude pico a pico e diferenças entre amostras consecutivas (WANG; WU, 2024). Redes Long Short-Term Memory (LSTM) e Bidirectional Long Short-Term Memory (BiLSTM) modelam diretamente as dependências temporais, mas geralmente apresentam maior custo computacional. Modelos tradicionais, como regressão logística, máquinas de vetores de suporte, árvores de decisão e conjuntos de árvores, operam sobre características previamente extraídas e podem facilitar a análise da complexidade e a execução em plataformas com recursos limitados.

A estratégia de avaliação exerce influência direta sobre os resultados. Quando janelas provenientes de uma mesma aquisição são distribuídas entre treinamento e teste, os dois subconjuntos podem conter segmentos fortemente relacionados, produzindo vazamento de dados e estimativas excessivamente otimistas. Para reduzir esse risco, a separação pode utilizar arquivos, sessões, pessoas ou períodos como unidades independentes. Gaiba et al. (2024) realizaram a divisão antes do processamento, enquanto Wang e Wu (2024) utilizaram dados de dias distintos para avaliar a generalização temporal.

2.4 Trabalhos relacionados

Em 2024, Gaiba et al. (2024) investigaram a identificação humana com dispositivos ESP32. Pessoas com diferentes características físicas permaneceram entre o transmissor e o receptor, e o sistema alcançou aproximadamente 95% de acurácia para dois indivíduos. Com a inclusão de uma terceira pessoa, o desempenho caiu para cerca de 74%. Embora a tarefa seja distinta da detecção de ocupação, o estudo demonstra que o CSI preserva informações associadas à interação entre o corpo e o canal e evidencia a dificuldade de separar classes com padrões semelhantes.

Também em 2024, no contexto de localização interna e execução em dispositivos de borda, Mendez et al. (2024) compararam modelos baseados em RSSI e CSI utilizando amostras isoladas e sequências temporais. O CSI apresentou desempenho superior em todas as configurações avaliadas, com vantagem mais acentuada nas séries temporais. Apesar de investigar uma tarefa diferente, o trabalho reforça a utilidade das informações preservadas por subportadora e indica a possibilidade de utilizar modelos baseados em CSI em plataformas com recursos limitados.

Wang e Wu (2024) utilizaram CSI obtido por um dispositivo ESP32 para investigar condições de ambiente vazio, pessoa estática e pessoa em movimento, com ênfase na robustez temporal da detecção. Os dados foram coletados em cinco dias, utilizando o primeiro para treinamento e validação e os quatro seguintes para avaliar o comportamento do método em períodos não utilizados durante o ajuste. Após a remoção de valores atípicos, filtragem por *wavelets* e extração de características temporais e estatísticas, os autores

combinaram máquinas de vetores de suporte, florestas aleatórias e *Extreme Gradient Boosting* por meio de votação ponderada, alcançando acurácia média superior a 98% na avaliação proposta.

Embora existam condições experimentais semelhantes às utilizadas nesta monografia, os trabalhos possuem objetivos e arquiteturas distintos. Wang e Wu (2024) concentram sua análise na robustez da detecção de movimento diante de variações entre diferentes dias e utilizam um roteador comercial como transmissor e um ESP32 como receptor. Neste trabalho, mantém-se como problema explícito a classificação entre ambiente vazio, presença estática e presença em movimento, incluindo a análise das confusões entre essas três classes. Além disso, a proposta contempla o desenvolvimento da cadeia completa de aquisição com dois ESP32-S3, a preservação e verificação dos dados brutos, a avaliação de diferentes configurações de pré-processamento e classificadores e a integração do modelo selecionado a um fluxo de inferência contínua.

Entre os trabalhos publicados em 2025, Bisio et al. (2025) compararam as capacidades de sensoriamento da placa Intel 5300 e do ESP32 em tarefas de detecção de presença e reconhecimento de atividades, com e sem linha de visão. Foram avaliadas diferentes representações do CSI utilizando florestas aleatórias e redes LSTM. Em determinadas configurações, o ESP32 apresentou desempenho próximo ao da plataforma de referência, mas os resultados variaram de acordo com a representação utilizada, o classificador e as condições de propagação.

No trabalho de R et al. (2025), módulos ESP32 foram empregados na detecção de presença em quartos, laboratórios e escritórios. Segmentos temporais de CSI foram classificados por uma rede BiLSTM e posteriormente submetidos a uma regra de consolidação aplicada sobre decisões consecutivas. A estratégia elevou a acurácia de 95,77% para 97,30% no conjunto de quartos e alcançou 96,83% em ambientes não utilizados no treinamento. Os resultados indicam que a consolidação temporal pode reduzir oscilações entre previsões, ainda que a solução dependa de uma arquitetura neural de maior complexidade.

Em conjunto, os estudos demonstram a viabilidade do CSI em tarefas de detecção de ocupação, reconhecimento de atividades, identificação e localização. Entretanto, as soluções diferem quanto à plataforma utilizada, à configuração da comunicação, às condições experimentais, às representações do CSI, aos procedimentos de processamento e às estratégias de avaliação. Essas diferenças dificultam a comparação direta dos resultados numéricos e reforçam a necessidade de descrever de forma explícita tanto as condições de aquisição quanto o protocolo utilizado para treinamento e teste.

Com base nesses aspectos, este trabalho adota uma arquitetura formada por dois dispositivos ESP32-S3, com geração controlada de tráfego, preservação dos dados brutos, mecanismos de verificação da aquisição e separação por arquivo durante a avaliação. A proposta concentra-se na distinção entre ambiente vazio, presença estática e presença em

movimento, utilizando características de baixa complexidade e comparando diferentes classificadores antes da integração do modelo selecionado a um fluxo de inferência contínua.

2.5 Síntese da revisão

A literatura demonstra que o CSI permite observar, ao longo do tempo, alterações no canal associadas à presença e ao movimento humano. A utilização de dispositivos ESP32 reduz o custo e o tamanho da infraestrutura, mas exige atenção à regularidade da aquisição, às perdas de pacotes e às particularidades da representação disponibilizada pela plataforma. A quantidade e a organização dos coeficientes também estão associadas às condições de comunicação empregadas, de modo que plataforma, modo de operação, largura de canal e demais parâmetros da aquisição devem ser considerados na interpretação e na comparação dos resultados.

Os estudos analisados também evidenciam a influência do pré-processamento, da representação das janelas, do classificador e do protocolo de avaliação sobre os resultados. Em particular, a separação entre treinamento e teste deve preservar a origem das aquisições, evitando que segmentos relacionados sejam distribuídos entre os dois subconjuntos. Diferentemente de abordagens voltadas principalmente à robustez temporal da detecção de movimento, este trabalho mantém como problema explícito a distinção entre ambiente vazio, presença estática e presença em movimento. Para isso, adota uma aquisição controlada com dois dispositivos ESP32-S3, preservação dos dados antes do processamento, divisão experimental por arquivo, características de baixa complexidade e integração do classificador selecionado a um sistema de inferência contínua.

3 Metodologia

Este capítulo apresenta os procedimentos experimentais e computacionais adotados para a formação do conjunto de dados, o processamento das aquisições CSI e a avaliação dos classificadores. A metodologia foi organizada de modo a preservar a origem de cada aquisição e impedir que informações pertencentes ao conjunto de teste participassem do ajuste do pré-processamento, da seleção de características ou do treinamento dos modelos.

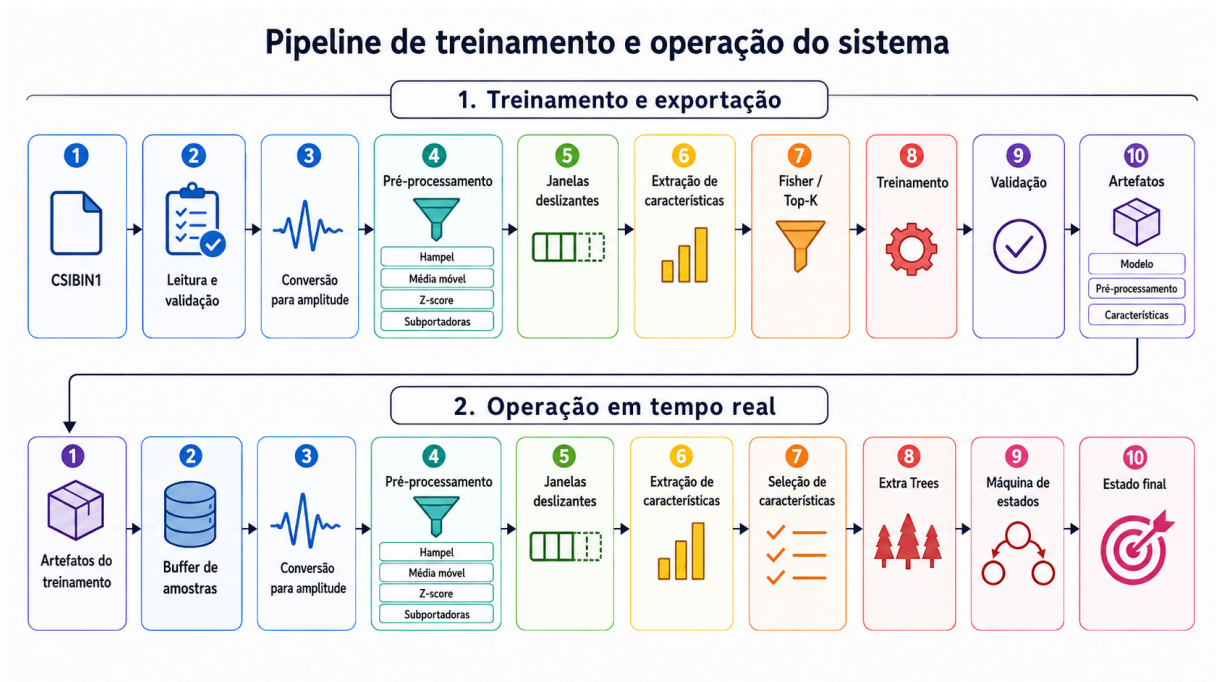
Inicialmente, são descritos o ambiente experimental, as condições representadas pelas classes, a organização das sessões de coleta e os critérios utilizados para aceitar ou rejeitar as aquisições. Em seguida, são apresentadas as etapas de conversão, validação, filtragem, normalização, redução de redundância e segmentação temporal dos dados.

Após o pré-processamento, são descritos os procedimentos de extração e seleção das características utilizadas como entrada dos classificadores. Por fim, são apresentados o treinamento dos modelos, a divisão dos arquivos entre treinamento e teste, as métricas de avaliação e o registro dos experimentos realizados.

A relação entre as etapas utilizadas na formação do modelo e aquelas reproduzidas durante a operação do sistema é sintetizada na Figura 2. As aquisições validadas foram armazenadas em um formato binário próprio, denominado CSIBIN1, utilizado para preservar as componentes reais e imaginárias do CSI, os números sequenciais, os marcadores temporais e os metadados associados a cada recepção. No fluxo de treinamento, esses arquivos são lidos, convertidos em amplitude, pré-processados e segmentados em janelas temporais. Em seguida, são extraídas e selecionadas as características utilizadas no treinamento e na avaliação dos classificadores. Ao final desse processo, são exportados o modelo, os parâmetros do pré-processamento e os índices das características selecionadas.

No fluxo de operação em tempo real, os artefatos produzidos durante o treinamento são carregados e aplicados às novas amostras. A conversão, o pré-processamento, a segmentação e a extração de características reproduzem a configuração definida anteriormente. O vetor resultante é reduzido pelos índices já armazenados, classificado pelo modelo Extra Trees e encaminhado à máquina de estados responsável pela consolidação temporal das previsões.

Figura 2 – Visão geral dos fluxos de treinamento, exportação e operação em tempo real



Fonte: imagem gerada por meio do ChatGPT e revisada pelo autor, 2026.

3.1 Procedimento experimental e formação do conjunto de dados

Para a formação do conjunto de dados utilizado no treinamento e na avaliação dos classificadores, foi empregado um sistema de aquisição composto por dois microcontroladores ESP32-S3 e um computador. O Emissor criou uma rede Wi-Fi experimental e transmitiu pacotes User Datagram Protocol (UDP) periodicamente, enquanto o Receptor se conectou a essa rede, capturou as amostras CSI associadas aos pacotes recebidos e as encaminhou ao computador por comunicação serial. No computador, os quadros foram interpretados, verificados quanto à integridade e armazenados em arquivos binários antes da aplicação de qualquer etapa de pré-processamento. Cada coleta foi associada à classe representada, à zona do ambiente, à sessão e ao arquivo de origem, preservando as informações necessárias à organização e à divisão do conjunto de dados.

Durante os experimentos, foram mantidos constantes os parâmetros de comunicação e de processamento embarcado utilizados pelo sistema de aquisição. O Emissor e o Receptor permaneceram nas mesmas posições e orientações, enquanto a condição de ocupação e a zona ocupada foram alteradas conforme o procedimento definido para cada coleta. Também foram mantidos o canal Wi-Fi, a largura de canal e a periodicidade nominal de transmissão, reduzindo a introdução de variações decorrentes da própria configuração do enlace.

A identificação da classe, da sessão, da posição e do arquivo de origem foi preservada junto às aquisições. Essas informações não são fornecidas ao classificador, mas permitem organizar o conjunto de dados, realizar divisões sem compartilhamento de janelas entre

treinamento e teste e analisar posteriormente a influência das condições de coleta.

3.1.1 Ambiente de aquisição

As aquisições foram realizadas em um ambiente interno com dimensões aproximadas de 3×3 metros. O espaço foi utilizado em sua configuração habitual, mantendo os móveis, objetos e equipamentos normalmente presentes no local durante as coletas. Dessa forma, o cenário não foi esvaziado para os experimentos, preservando elementos capazes de produzir reflexões e múltiplos percursos do sinal Wi-Fi. A disposição desses elementos foi mantida durante as aquisições, de modo que as principais alterações introduzidas no ambiente fossem aquelas relacionadas à condição de ocupação e à região ocupada pelo participante.

O Emissor e o Receptor foram posicionados em regiões opostas do ambiente, formando um enlace Wi-Fi aproximadamente diagonal em relação à área monitorada. O Emissor foi instalado a uma altura aproximada de 1,0 metro, enquanto o Receptor foi posicionado a aproximadamente 1,5 metro. Essa disposição fez com que o enlace atravessasse uma parcela significativa da região utilizada pelo participante durante as aquisições. Os dispositivos permaneceram nas mesmas posições e orientações ao longo do procedimento experimental, evitando que alterações na geometria do próprio enlace fossem introduzidas entre as coletas.

Essa disposição permitiu que diferentes regiões da sala interferissem de maneira distinta no enlace. A presença ou o movimento do participante poderia modificar os percursos de propagação entre os dispositivos, mesmo quando ele não estivesse posicionado diretamente sobre a linha que os separava.

A Figura 3 apresenta uma representação simplificada do ambiente, contendo a posição do Emissor, do Receptor e das zonas utilizadas nas aquisições.

A divisão em zonas não representa uma separação física rígida do ambiente. Cada zona corresponde a uma região aproximada na qual o participante deveria permanecer ou realizar os movimentos previstos para a classe coletada. Pequenas variações de posição dentro da mesma região foram permitidas, evitando que todas as aquisições representassem um único ponto espacial. Dessa forma, as zonas foram utilizadas como referência para distribuir as coletas pelo ambiente e representar diferentes relações geométricas entre a posição do participante e o enlace formado pelo Emissor e pelo Receptor.

A utilização de diferentes zonas permite analisar se a resposta do sistema varia de acordo com a posição do participante em relação ao Emissor e ao Receptor. Regiões próximas ao percurso direto entre os dispositivos podem produzir alterações distintas daquelas observadas em posições mais afastadas, devido à modificação dos diferentes percursos de propagação existentes no ambiente. A zona associada a cada aquisição é preservada como metadado durante o processamento. Essa informação não é fornecida ao

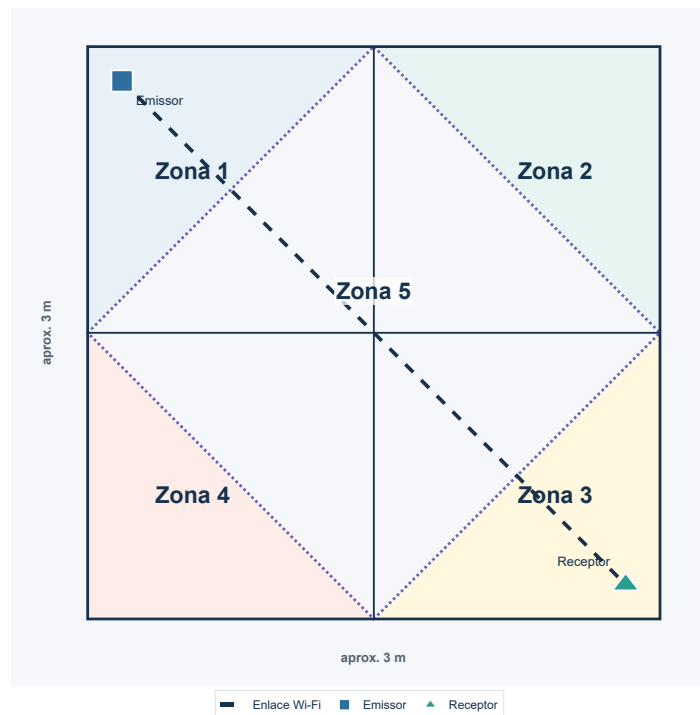


Figura 3 – Representação do ambiente e do posicionamento dos dispositivos durante as aquisições

classificador, sendo utilizada apenas para organizar o conjunto de dados, realizar a divisão entre treinamento e teste e analisar posteriormente o desempenho em diferentes regiões do ambiente.

3.1.2 Definição das classes

As aquisições foram organizadas em três classes: ambiente vazio, presença estática e presença em movimento. Todas as coletas envolvendo presença humana foram realizadas com um único participante, mantendo o mesmo indivíduo ao longo de todo o conjunto de dados. Cada classe foi definida de forma operacional, de modo que as condições fossem reproduzidas de maneira semelhante nas diferentes zonas e sessões de coleta.

A classe de ambiente vazio corresponde às aquisições realizadas sem pessoas no interior do cômodo. Durante essas coletas, o ambiente permaneceu nas condições normais de utilização, mantendo os móveis, objetos e equipamentos presentes na sala.

Na classe de presença estática, o participante permanecia dentro da zona indicada durante todo o intervalo de aquisição. Procurou-se evitar deslocamentos e movimentos amplos, embora pequenas alterações de postura e movimentos involuntários pudessem ocorrer. Essa condição representa situações em que existe ocupação, mas não há atividade corporal suficiente para produzir variações intensas no canal.

A classe de presença em movimento corresponde às aquisições em que o participante realizava movimentos dentro da zona definida. Foram permitidos deslocamentos curtos, alterações de posição e movimentação dos membros dentro da região correspondente à coleta. Não foi estabelecida uma sequência específica de ações a ser repetida, pois o objetivo dessa classe não era reconhecer uma atividade determinada, mas representar uma condição de ocupação com variações corporais mais evidentes do que aquelas observadas na presença estática.

A delimitação das classes não busca reconhecer comportamentos individuais, como caminhar, sentar ou levantar. O objetivo é distinguir a ausência de pessoas, a presença com pouca movimentação e a presença acompanhada de variações corporais mais evidentes. As mesmas definições foram utilizadas durante todas as sessões. A classe associada a cada aquisição foi registrada no nome e no conteúdo do arquivo, permitindo verificar sua correspondência durante o carregamento e o processamento dos dados.

3.1.3 Sessões de coleta

As aquisições foram distribuídas em sessões distintas. Cada sessão reúne um conjunto de coletas realizado em um mesmo período, mantendo a configuração dos dispositivos, os parâmetros de comunicação e o procedimento experimental previamente definido.

A separação em sessões permite distinguir aquisições realizadas em momentos diferentes. Mesmo quando a disposição do ambiente e o posicionamento dos dispositivos são preservados, pequenas variações podem ocorrer entre os períodos de coleta, como mudanças nas condições de propagação, interferências externas e alterações discretas no ambiente. Durante cada uma das duas sessões, foram realizadas aquisições das três classes nas cinco zonas previstas. Para cada combinação entre sessão, zona e classe foram planejadas dez aquisições independentes, mantendo o mesmo participante nas condições de presença estática e presença em movimento. A ordem das coletas foi definida pela ferramenta de aquisição e pelo operador, mantendo-se o mesmo procedimento de preparação, gravação e verificação dos arquivos.

A sessão associada a cada aquisição foi preservada como metadado durante o carregamento e o processamento dos dados. Essa informação não é fornecida ao classificador, sendo utilizada para organizar o conjunto de dados e permitir avaliações em que arquivos pertencentes a períodos distintos sejam identificados e analisados separadamente.

3.1.4 Critérios de aceitação das aquisições

Antes do processamento, os arquivos eram verificados quanto à presença de amostras válidas, à quantidade suficiente de pacotes, à compatibilidade entre as componentes reais e imaginárias e à identificação correta da classe, da sessão e da zona. Também eram

observados os indicadores de integridade e continuidade produzidos pelo sistema de aquisição. Dessa forma, arquivos vazios, interrompidos, estruturalmente incompatíveis ou incapazes de formar ao menos uma janela completa não eram utilizados nos experimentos. Os critérios aplicados e a quantidade de arquivos preservados no conjunto final são apresentados no capítulo de resultados.

3.2 Pré-processamento dos dados

O pipeline de pré-processamento recebe os arquivos **CSIBIN1** produzidos durante as aquisições e os transforma em matrizes temporais adequadas à extração de características. O rótulo preservado em cada arquivo, juntamente com seu nome e sua localização na estrutura de diretórios, permite identificar a classe, a sessão, a zona e a aquisição de origem. Os números sequenciais e os marcadores temporais, por sua vez, permitem verificar a continuidade dos registros e reconstruir a ordem das amostras.

Durante a leitura, as componentes reais I e imaginárias Q são recuperadas e verificadas quanto à presença, ao comprimento e à compatibilidade estrutural. Os registros válidos são convertidos em amplitude e organizados em uma matriz na qual as linhas representam os pacotes recebidos e as colunas correspondem às posições do vetor **CSI**. A partir dessa representação, são aplicadas a remoção de valores atípicos, a suavização, a normalização, a redução de redundância e a segmentação em janelas. A estrutura completa do formato binário e seu processo de gravação são apresentados na Subseção 4.4.2.

As operações dependentes da distribuição dos dados foram separadas entre ajuste e transformação. Durante o ajuste, as médias, os desvios padrão e os índices das subportadoras preservadas são determinados exclusivamente a partir dos arquivos de treinamento. Na transformação, esses mesmos parâmetros são aplicados aos arquivos de treinamento, teste e às novas aquisições, impedindo que o conjunto avaliado participe da definição do pré-processamento.

Cada etapa possui parâmetros próprios, como o tamanho da vizinhança do filtro, o comprimento da suavização, o limiar de correlação, o tamanho das janelas e o passo entre elas. Os valores avaliados e a configuração selecionada são apresentados no capítulo de resultados, junto às comparações que fundamentaram sua escolha.

3.2.1 Conversão e validação dos dados

As componentes reais e imaginárias preservadas nos arquivos são inicialmente convertidas em amplitude. Para cada posição k do vetor **CSI** e instante t , o cálculo é realizado por:

$$A_k(t) = \sqrt{I_k(t)^2 + Q_k(t)^2}. \quad (3.1)$$

A conversão somente é realizada quando os vetores real e imaginário estão presentes e possuem o mesmo comprimento. Os pacotes com estrutura incompatível ou quantidade de elementos diferente da esperada para o arquivo são removidos, sem preenchimento com zeros ou truncamento automático.

Após a conversão, cada arquivo é representado por uma matriz $\mathbf{X} \in \mathbb{R}^{N \times M}$, dada por:

$$\mathbf{X} = \begin{bmatrix} A_0(t_0) & A_1(t_0) & \cdots & A_{M-1}(t_0) \\ A_0(t_1) & A_1(t_1) & \cdots & A_{M-1}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ A_0(t_{N-1}) & A_1(t_{N-1}) & \cdots & A_{M-1}(t_{N-1}) \end{bmatrix}, \quad (3.2)$$

em que N corresponde à quantidade de pacotes do arquivo e M à quantidade de posições do vetor [CSI](#). Dessa forma, cada linha representa um pacote recebido e cada coluna corresponde a uma posição do vetor [CSI](#). A correspondência entre cada coluna e seu índice original é mantida durante todo o processamento, pois os parâmetros de normalização e redução de redundância dependem dessa associação.

A amplitude foi utilizada como representação de entrada do pipeline. As componentes originais permanecem preservadas nos arquivos, mas a fase não foi incorporada ao processamento implementado.

3.2.2 Remoção de valores atípicos

O filtro de Hampel é aplicado separadamente à sequência temporal de cada coluna da matriz [CSI](#). Para cada amostra, são calculadas a mediana e a dispersão absoluta em uma vizinhança temporal local.

Para a amostra $X_{i,k}$, correspondente ao instante i e à coluna k , considera-se uma vizinhança temporal local $\mathcal{V}_{i,k}$. A mediana dessa vizinhança é dada por:

$$m_{i,k} = \text{mediana}(\mathcal{V}_{i,k}). \quad (3.3)$$

A dispersão absoluta local é calculada por:

$$d_{i,k} = \text{mediana}(|X_{j,k} - m_{i,k}|), \quad X_{j,k} \in \mathcal{V}_{i,k}. \quad (3.4)$$

Assim, o valor filtrado é definido por:

$$\widetilde{X}_{i,k} = \begin{cases} m_{i,k}, & \text{se } |X_{i,k} - m_{i,k}| > \lambda d_{i,k}, \\ X_{i,k}, & \text{caso contrário,} \end{cases} \quad (3.5)$$

em que λ representa o limiar aplicado à dispersão local.

Quando a diferença entre o valor analisado e a mediana da vizinhança ultrapassa um limiar definido a partir da dispersão local, a amostra é classificada como atípica e substituída pela mediana. Quando a dispersão local é nula, o valor é mantido, pois não existe variação suficiente para estabelecer um limite diferente de zero.

A substituição ocorre apenas na posição identificada como atípica. Pacotes completos não são removidos, preservando a quantidade de linhas da matriz e o alinhamento temporal entre as subportadoras. O tamanho da vizinhança e o limiar utilizados na configuração final são apresentados junto aos resultados da definição do pipeline.

3.2.3 Suavização temporal

Após o filtro de Hampel, é aplicada uma média móvel centrada à sequência temporal de cada subportadora. Em cada posição, o valor suavizado é obtido pela média das amostras pertencentes a uma vizinhança local. Para a amostra localizada no instante i e na subportadora k , considera-se o conjunto de índices disponíveis na vizinhança centrada:

$$\mathcal{J}_i = \{j \mid \max(0, i - h) \leq j \leq \min(N - 1, i + h)\}, \quad (3.6)$$

em que h representa o raio da janela e N a quantidade de pacotes da sequência. O valor suavizado é calculado por:

$$\overline{X}_{i,k} = \frac{1}{|\mathcal{J}_i|} \sum_{j \in \mathcal{J}_i} \widetilde{X}_{j,k}, \quad (3.7)$$

em que $\widetilde{X}_{j,k}$ corresponde ao valor resultante do filtro de Hampel e $|\mathcal{J}_i|$ à quantidade de amostras disponíveis na vizinhança.

Nas extremidades da sequência, são utilizadas apenas as posições disponíveis. Dessa forma, a operação mantém a mesma quantidade de pacotes e não altera a correspondência temporal entre as subportadoras. A suavização busca reduzir oscilações isoladas sem eliminar completamente as variações utilizadas posteriormente na classificação. O comprimento escolhido para a média móvel é apresentado no capítulo de resultados.

3.2.4 Normalização das subportadoras

As colunas resultantes são normalizadas individualmente pelo método Z-score. Para cada coluna k , são calculados a média μ_k e o desvio padrão σ_k , utilizando somente os arquivos pertencentes ao conjunto de treinamento:

$$z_{i,k} = \frac{\bar{X}_{i,k} - \mu_k}{\sigma_k}, \quad (3.8)$$

em que $\bar{X}_{i,k}$ corresponde ao valor suavizado da subportadora k no instante i .

As referências não são recalculadas para cada arquivo ou janela. Os valores obtidos durante o ajuste são armazenados e aplicados sem alterações aos demais dados da mesma divisão experimental e às novas aquisições processadas pelo modelo correspondente. Antes da normalização, as colunas com desvio padrão inferior a um limite mínimo são classificadas como não informativas e removidas. Essa verificação impede divisões por valores próximos de zero e evita a manutenção de posições sem variação temporal significativa.

A estimativa dos parâmetros somente com os arquivos de treinamento impede que informações estatísticas do conjunto de teste sejam incorporadas ao pipeline antes da avaliação. O limite numérico adotado na configuração final é apresentado junto aos resultados do pré-processamento.

3.2.5 Redução de redundância

Após a remoção das colunas não informativas, as posições restantes são comparadas pela correlação de Pearson. Duas colunas são consideradas redundantes quando o valor absoluto de sua correlação ultrapassa um limiar previamente definido. A seleção é realizada sequencialmente. A primeira coluna válida é preservada, e cada nova candidata é comparada com as colunas já mantidas. Quando a candidata apresenta correlação acima do limiar com qualquer uma delas, é descartada. Caso contrário, passa a integrar o conjunto reduzido.

Como o procedimento depende da ordem das colunas, os índices originais são mantidos durante o ajuste. A lista resultante é armazenada e reutilizada para filtrar os arquivos de treinamento, teste e as novas aquisições. Essa etapa busca remover comportamentos fortemente redundantes, enquanto a capacidade de cada característica para diferenciar as classes é avaliada posteriormente, durante a seleção de características. O limiar escolhido e a quantidade de posições preservadas são apresentados no capítulo de resultados.

3.2.6 Segmentação em janelas temporais

Após o pré-processamento, cada arquivo permanece representado por uma sequência temporal de pacotes. Essa sequência é dividida em janelas deslizantes com tamanho W e passo S .

Para um arquivo contendo N pacotes processados, a quantidade de janelas completas é dada por:

$$J = \left\lfloor \frac{N - W}{S} \right\rfloor + 1. \quad (3.9)$$

Janelas incompletas ao final do arquivo não são utilizadas. Cada janela preserva o rótulo, a sessão, a posição e o arquivo de origem da aquisição. Essa associação permite realizar posteriormente a divisão entre treinamento e teste sem distribuir janelas provenientes do mesmo arquivo entre os dois conjuntos.

A sobreposição aumenta a quantidade de exemplos e permite produzir decisões em intervalos menores do que o tamanho total da janela. Entretanto, as janelas de uma mesma aquisição permanecem relacionadas e não devem ser interpretadas como arquivos independentes. O procedimento geral de segmentação é apresentado na Figura 4.

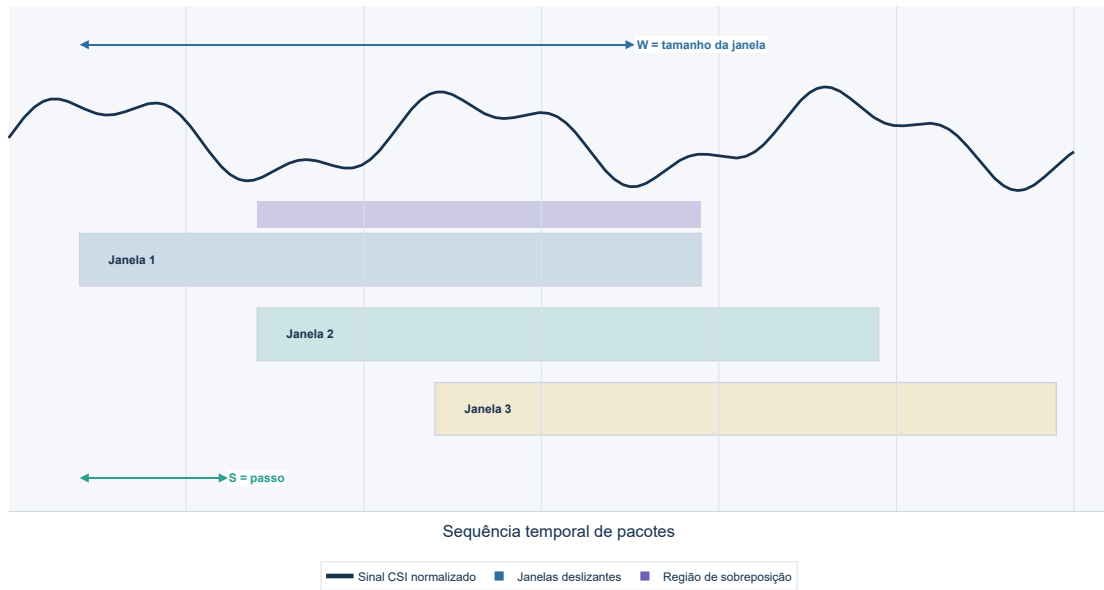


Figura 4 – Formação de janelas temporais deslizantes com tamanho W e passo S

Como os parâmetros podem ser definidos em quantidade de pacotes, a duração temporal correspondente depende da taxa efetiva de aquisição. Os valores avaliados para W e S , bem como a configuração selecionada, são apresentados no capítulo de resultados.

3.3 Extração e seleção de características

Após o pré-processamento e a segmentação, cada janela é representada por uma matriz composta por W pacotes e M subportadoras. Para produzir a entrada dos classificadores, cada sequência temporal é convertida em características estatísticas e temporais de baixa complexidade.

A implementação utiliza onze características por subportadora. Os valores são extraídos sempre na mesma ordem e posteriormente concatenados em um único vetor. Essa ordem deve permanecer associada aos parâmetros do pré-processamento e ao modelo treinado, pois cada posição possui um significado específico.

3.3.1 Extração das características

Para uma subportadora k , a sequência contida em uma janela é representada por:

$$\mathbf{x}_k = [x_0, x_1, \dots, x_{W-1}], \quad (3.10)$$

em que x_i corresponde à amplitude normalizada no pacote i . As características são calculadas separadamente para cada subportadora, conforme apresentado na Tabela 1.

Para as medidas baseadas na variação entre pacotes, é utilizada a diferença de primeira ordem:

$$d_i = x_i - x_{i-1}, \quad 1 \leq i < W. \quad (3.11)$$

Tabela 1 – Características extraídas de cada subportadora

Característica	Definição
Média	Média aritmética dos valores da janela.
Desvio padrão	Dispersão dos valores em relação à média.
Mínimo	Menor valor observado na janela.
Máximo	Maior valor observado na janela.
Amplitude pico a pico	Diferença entre os valores máximo e mínimo.
Energia	Soma dos quadrados dos valores da sequência.
Média das diferenças absolutas	Média de $ d_i $ entre pacotes.
Máxima diferença absoluta	Maior valor de $ d_i $ observado na janela.
Energia das diferenças	Soma dos quadrados das diferenças de primeira ordem.
Desvio padrão das diferenças	Dispersão das diferenças de primeira ordem.
Inclinação linear	Coefficiente angular da reta ajustada à sequência temporal.

As seis primeiras características representam o nível e a distribuição dos valores da janela. As quatro características baseadas nas diferenças descrevem as alterações entre pacotes consecutivos, enquanto a inclinação representa a tendência geral da sequência. Do ponto de vista computacional, todas as operações possuem custo linear ou constante em

relação ao tamanho da janela. Como a quantidade de características é fixa, o custo total da extração para M subportadoras e W pacotes é da ordem de:

$$O(MW). \quad (3.12)$$

3.3.2 Construção do vetor de características

Para cada subportadora preservada pelo pré-processamento, as onze características são calculadas na ordem apresentada na Tabela 1. O vetor parcial correspondente à subportadora s_j é dado por:

$$\mathbf{f}_{s_j} = [\mu, \sigma, x_{\min}, x_{\max}, x_{\text{pp}}, E_x, D_{\text{média}}, D_{\max}, E_d, \sigma_d, a]_{s_j}. \quad (3.13)$$

O vetor completo da janela é formado pela concatenação dos vetores parciais:

$$\mathbf{f} = [\mathbf{f}_{s_0}, \mathbf{f}_{s_1}, \dots, \mathbf{f}_{s_{M-1}}]. \quad (3.14)$$

Consequentemente, sua dimensão inicial é:

$$D = 11M, \quad (3.15)$$

em que M corresponde à quantidade de subportadoras mantidas após a redução de redundância. A dimensão D resulta, portanto, das onze características extraídas de cada uma das M subportadoras.

Cada posição do vetor permanece associada ao tipo da característica e ao índice original da subportadora, permitindo preservar sua interpretação ao longo das etapas de seleção, treinamento e inferência.

Além do vetor numérico, são preservados o rótulo, a zona, o arquivo de origem e o índice da janela dentro da aquisição. Esses metadados não são fornecidos ao classificador, mas permitem realizar a separação por arquivos, calcular métricas por posição e retornar às aquisições associadas às classificações analisadas. Os vetores de características podem ser armazenados em arquivos intermediários para fins de diagnóstico e reutilização em experimentos que empreguem exatamente a mesma configuração de processamento.

Nas avaliações finais, entretanto, a divisão por arquivos ocorre antes do ajuste das operações dependentes dos dados. Dessa forma, cada repetição experimental produz suas próprias médias, desvios, subportadoras preservadas e pontuações de Fisher a partir do subconjunto de treinamento.

3.3.3 Seleção pelo Fisher Score

Nem todas as características do vetor completo apresentam a mesma capacidade de diferenciar as classes. Para ordenar os atributos individualmente, foi utilizado o Fisher Score, que compara a separação entre as médias das classes com a dispersão interna de cada uma. Para a característica de índice j , a pontuação é calculada por:

$$F_j = \frac{\sum_c n_c (\bar{x}_{j,c} - \bar{x}_j)^2}{\sum_c n_c s_{j,c}^2 + \varepsilon}, \quad (3.16)$$

em que n_c representa a quantidade de amostras da classe c , $\bar{x}_{j,c}$ e $s_{j,c}^2$ correspondem, respectivamente, à média e à variância da característica j nessa classe, e $\bar{x}_j$ representa sua média global. O termo ε é uma constante positiva de pequeno valor utilizada para evitar a divisão por zero.

Pontuações maiores indicam atributos cujas médias apresentam maior separação entre as classes e menor dispersão interna. Após o cálculo, os índices são ordenados em ordem decrescente, e as primeiras K posições formam o subconjunto fornecido ao classificador.

O Fisher Score é calculado exclusivamente com as janelas pertencentes aos arquivos de treinamento. Os mesmos índices, na mesma ordem, são aplicados posteriormente aos vetores de treinamento e teste. Sempre que a divisão entre os arquivos é alterada, o ranking deve ser ajustado novamente a partir do novo conjunto de treinamento. Os índices selecionados e suas pontuações são armazenados junto aos parâmetros do experimento. Durante a inferência, o ranking não é recalculado. O sistema extrai as características na ordem original e aplica os índices definidos durante o treinamento.

A seleção permanece vinculada à lista de subportadoras, à ordem das onze características, aos parâmetros de normalização, ao tamanho da janela e ao modelo treinado. Uma alteração em qualquer um desses elementos modifica a interpretação das posições e exige a geração de novos parâmetros.

3.4 Treinamento e validação dos classificadores

A infraestrutura de treinamento e validação reúne as etapas necessárias para transformar as aquisições em conjuntos processados e avaliar os classificadores sob condições controladas. O fluxo contempla a geração das janelas e dos vetores de características, a seleção dos atributos, o treinamento dos modelos, a produção das previsões e o cálculo das métricas.

Cada execução registra os parâmetros empregados, como o tamanho da janela, a

quantidade de características, a configuração do classificador e a semente utilizada na divisão dos dados. Esse registro permite comparar diferentes experimentos e reproduzir suas condições de avaliação sem modificar os arquivos brutos.

3.4.1 Geração do conjunto de características

As aquisições reconhecidas pelo índice do conjunto de dados são processadas para produzir as janelas e seus respectivos vetores de características. Arquivos inválidos, vazios ou com quantidade insuficiente de pacotes para formar uma janela completa não geram amostras. O fluxo de geração e armazenamento das amostras processadas é apresentado na Figura 5.

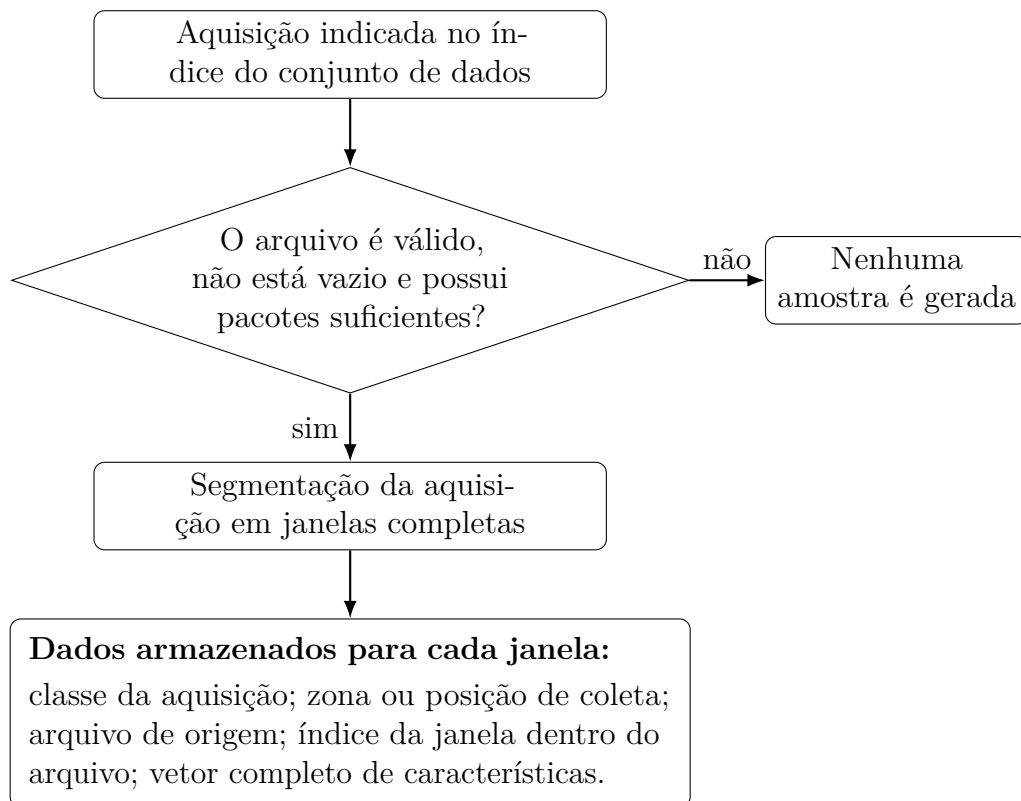


Figura 5 – Fluxo de geração e armazenamento das amostras processadas

A associação com o arquivo de origem é preservada para permitir a divisão por aquisições e a análise posterior das previsões, embora os metadados não sejam utilizados como entradas do classificador. Além disso, os conjuntos processados podem ser serializados para fins de diagnóstico ou reutilização quando empregam uma configuração fixa. Os parâmetros correspondentes, como médias, desvios padrão, subportadoras preservadas, tamanho da janela e limiar de correlação, são registrados separadamente.

Nas avaliações utilizadas para estimar o desempenho dos classificadores, a separação dos arquivos ocorre antes da estimativa das médias, dos desvios, das correlações e das

pontuações de Fisher. Os arquivos de teste são transformados somente com os parâmetros obtidos a partir do treinamento.

3.4.2 Classificadores avaliados

A avaliação foi organizada de forma a permitir a comparação de diferentes classificadores utilizando a mesma representação de entrada, as mesmas divisões entre treinamento e teste e o mesmo conjunto de métricas. Uma árvore de decisão foi inicialmente implementada em Python sem dependência de bibliotecas de aprendizado de máquina. Cada nó interno utiliza uma regra da forma:

$$f_j \leq \theta, \quad (3.17)$$

em que f_j representa uma característica e θ o limiar da divisão. A estrutura foi utilizada como referência de baixa complexidade e permitiu observar diretamente as regras formadas pelo treinamento.

Além da implementação própria, foram avaliados classificadores disponibilizados por bibliotecas de aprendizado de máquina, incluindo modelos lineares, métodos baseados em distância, máquinas de vetores de suporte, árvores e conjuntos de árvores. Os modelos foram adaptados a uma interface comum de treinamento, predição e obtenção das probabilidades das classes. Para os classificadores sensíveis à escala dos atributos, foi aplicada uma padronização adicional, ajustada somente com os vetores de treinamento e reutilizada no conjunto de teste. Os parâmetros configuráveis incluem, entre outros, profundidade, quantidade mínima de amostras, número de estimadores, taxa de aprendizado, número de vizinhos, tipo de núcleo, balanceamento das classes e quantidade de características selecionadas.

Além das métricas de classificação, a infraestrutura registra medidas de tempo, complexidade e tamanho das estruturas produzidas. A escolha do classificador utilizado na aplicação final considera tanto o desempenho preditivo quanto o custo de execução e a possibilidade de integração com o sistema em tempo real. As comparações entre os modelos, as configurações avaliadas e a escolha do candidato final são apresentadas no capítulo de resultados.

3.4.3 Validação separada por arquivo

As janelas provenientes de uma mesma aquisição são reunidas pelo campo que identifica o arquivo de origem e tratadas como um único grupo. Antes da divisão, o programa verifica se todas as janelas pertencentes ao grupo possuem o mesmo rótulo. Uma inconsistência provoca a interrupção da execução.

Foi adotada a estratégia de *holdout* estratificado separado por arquivo, destinando aproximadamente 80% das aquisições ao treinamento e 20% ao teste. Os estratos são definidos considerando a classe, a sessão e a zona de coleta, buscando preservar a representação desses grupos nos dois subconjuntos. Todas as janelas provenientes de um mesmo arquivo permanecem exclusivamente no treinamento ou no teste. Como arquivos diferentes podem produzir quantidades distintas de janelas, a proporção entre janelas não corresponde necessariamente à proporção entre arquivos. A divisão é realizada antes de qualquer operação dependente da distribuição dos dados. A sequência aplicada em cada execução é apresentada na Figura 6.

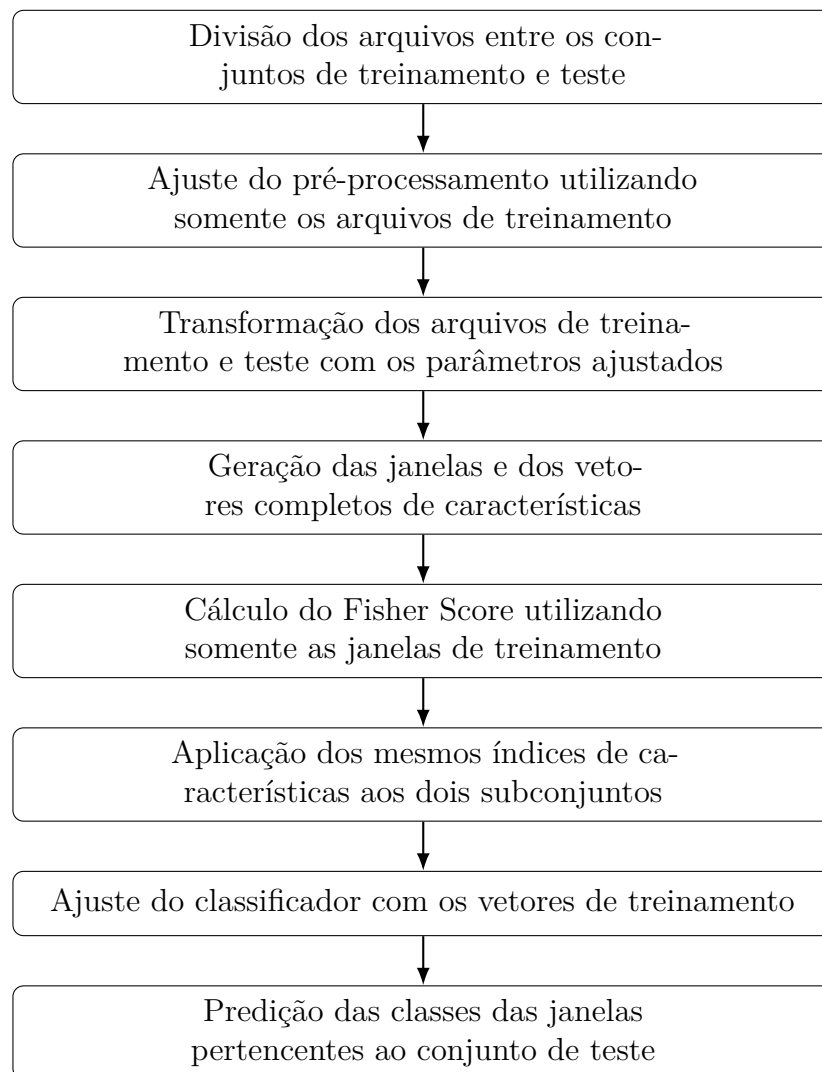


Figura 6 – Fluxo da validação separada por arquivo

O conjunto de teste não participa do cálculo da normalização, da redução por correlação, do ranking das características, da padronização adicional nem do ajuste do modelo. A infraestrutura permite repetir esse procedimento com diferentes sementes de aleatoriedade. Foram definidos previamente os valores 7, 13, 21, 42, 84, 126, 168, 210, 336 e 512, os quais permanecem registrados na configuração experimental. Esses valores não

possuem significado associado às classes ou às aquisições, sendo utilizados apenas para inicializar de forma reprodutível o embaralhamento dos arquivos.

Cada semente produz uma nova divisão entre treinamento e teste, um novo ajuste do pré-processamento, um novo ranking de características e um novo modelo. Dentro de uma mesma repetição, todos os classificadores recebem exatamente a mesma divisão, permitindo que seus resultados sejam comparados sob condições equivalentes. Também foram previstas avaliações nas quais grupos completos, como sessões ou posições de coleta, são mantidos fora do treinamento. Esses protocolos permitem observar a capacidade de generalização para condições não representadas diretamente durante o ajuste. As métricas são calculadas individualmente em cada repetição e posteriormente consolidadas. Os valores resultantes e a quantidade de repetições utilizadas são apresentados no capítulo de resultados.

Considerando L classes e N janelas de teste, sejam VP_c , FP_c , FN_c e VN_c , respectivamente, os verdadeiros positivos, falsos positivos, falsos negativos e verdadeiros negativos da classe c . Para o cálculo das métricas por classe, a classe analisada é considerada positiva, enquanto todas as demais são reunidas como negativas. A acurácia corresponde à proporção de previsões corretas entre todas as janelas avaliadas:

$$\text{Acc} = \frac{\sum_{c=1}^L VP_c}{N}. \quad (3.18)$$

Essa métrica fornece uma visão geral do desempenho do modelo. Por exemplo, uma acurácia de 0,80 indica que 80% das janelas receberam o rótulo correto. Entretanto, seu valor deve ser analisado com cautela quando as classes possuem quantidades diferentes de amostras, pois um bom desempenho na classe mais frequente pode ocultar dificuldades nas demais. A precisão e a revocação de cada classe são definidas por:

$$P_c = \frac{VP_c}{VP_c + FP_c}, \quad R_c = \frac{VP_c}{VP_c + FN_c}. \quad (3.19)$$

A precisão responde à seguinte questão: entre todas as janelas classificadas pelo modelo como pertencentes à classe c , quantas realmente pertenciam a essa classe? Um valor reduzido indica a ocorrência de muitos falsos positivos.

A revocação, também denominada sensibilidade, indica a proporção das janelas que realmente pertenciam à classe c e foram corretamente identificadas. Um valor reduzido indica que o modelo deixou de reconhecer muitas amostras dessa classe, produzindo falsos negativos. O F1-score da classe corresponde à média harmônica entre precisão e revocação:

$$F1_c = 2 \frac{P_c R_c}{P_c + R_c}. \quad (3.20)$$

Essa métrica resume o equilíbrio entre as duas medidas. O F1-score somente assume um valor elevado quando tanto a precisão quanto a revocação são elevadas. Dessa forma, ele penaliza modelos que identificam grande parte das amostras de uma classe, mas também produzem muitas classificações incorretas, ou que apresentam alta precisão à custa de deixar de identificar diversas amostras reais. O Macro F1 corresponde à média aritmética dos valores de F1-score das classes:

$$F1_{\text{macro}} = \frac{1}{L} \sum_{c=1}^L F1_c. \quad (3.21)$$

Como todas as classes recebem o mesmo peso, essa métrica permite observar se o classificador mantém desempenho equilibrado entre ambiente vazio, presença estática e presença em movimento. Um resultado reduzido pode indicar que ao menos uma das classes apresenta desempenho inferior, mesmo quando a acurácia global permanece elevada. O suporte n_c corresponde à quantidade de janelas reais pertencentes à classe c . Ele não constitui uma medida de desempenho, mas informa quantas amostras contribuíram para o cálculo das métricas daquela classe.

O F1 ponderado considera o suporte ao combinar os resultados:

$$F1_{\text{ponderado}} = \frac{1}{N} \sum_{c=1}^L n_c F1_c. \quad (3.22)$$

Nesse caso, as classes com maior quantidade de janelas exercem maior influência sobre o resultado final. O F1 ponderado representa melhor o desempenho médio considerando a distribuição observada no conjunto de teste, enquanto o Macro F1 atribui a mesma importância a todas as classes.

Além das métricas calculadas a partir dos rótulos finais, são utilizadas as curvas ROC e precisão-revocação, que consideram as probabilidades produzidas pelo classificador. Para cada classe c , ela é tratada como positiva e as demais como negativas, enquanto o limiar de decisão é variado. Na curva ROC, a taxa de verdadeiros positivos e a taxa de falsos positivos são definidas por:

$$\text{TPR}_c = \frac{VP_c}{VP_c + FN_c}, \quad \text{FPR}_c = \frac{FP_c}{FP_c + VN_c}. \quad (3.23)$$

A curva ROC representa TPR_c em função de FPR_c para diferentes limiares. A área sob essa curva, denominada Area Under the Receiver Operating Characteristic Curve (**AUC-ROC**), pode ser expressa por:

$$\text{AUC-ROC}_c = \int_0^1 \text{TPR}_c(\text{FPR}_c) d\text{FPR}_c. \quad (3.24)$$

A **AUC-ROC** mede a capacidade do modelo de atribuir probabilidades maiores às amostras da classe analisada do que às amostras das demais classes. Valores próximos de 1 indicam melhor separação, enquanto um valor próximo de 0,5 corresponde ao comportamento esperado de uma classificação aleatória.

A curva precisão-revocação representa a precisão em função da revocação obtida para diferentes limiares. A área sob essa curva é definida por:

$$\text{AUC-PR}_c = \int_0^1 P_c(R_c) dR_c. \quad (3.25)$$

A Area Under the Precision-Recall Curve (**AUC-PR**) evidencia a relação entre a capacidade de encontrar as amostras positivas e a quantidade de falsos positivos produzidos. Essa métrica é especialmente útil quando existe diferença entre as quantidades de amostras das classes, pois seu valor é mais sensível ao desempenho na classe considerada positiva.

No problema multiclasse, a **AUC-ROC** e a **AUC-PR** são calculadas separadamente para cada classe pela estratégia um contra as demais. Os resultados por classe permitem identificar em quais categorias as probabilidades apresentam maior ou menor capacidade de separação. Os valores também podem ser consolidados por média macro, atribuindo o mesmo peso às classes, ou por média ponderada, considerando seus respectivos suportes.

A matriz de confusão **C** é formada pelos elementos C_{ij} , que indicam a quantidade de amostras da classe real i classificadas como pertencentes à classe j . Os valores da diagonal principal representam os acertos, enquanto os demais elementos mostram quais classes foram confundidas entre si. Diferentemente das métricas resumidas em um único valor, essa matriz permite identificar diretamente, por exemplo, se as amostras de presença estática são classificadas incorretamente como presença em movimento ou como ambiente vazio.

As previsões também podem ser agrupadas por arquivo e zona, permitindo identificar aquisições ou posições que apresentem maior dificuldade.

A frequência de seleção das características é obtida pela proporção de repetições nas quais cada atributo aparece entre as primeiras posições do ranking. Essa medida permite observar se o subconjunto utilizado pelo modelo permanece estável quando os arquivos de treinamento são alterados.

3.4.4 Registro e organização dos experimentos

Cada execução experimental produz um conjunto de artefatos destinado a registrar a configuração utilizada, a composição dos subconjuntos, as previsões e as métricas obtidas. Os resultados são organizados em diretórios separados para tabelas, figuras e relatórios. Entre os dados armazenados encontram-se:

- configuração do pipeline;
- arquivos utilizados no treinamento e no teste;
- características selecionadas;
- parâmetros do classificador;
- previsões das janelas de teste;
- métricas globais e por classe;
- matrizes de confusão;
- desempenho por arquivo, sessão e posição;
- medidas de tempo e complexidade;
- resultados individuais e consolidados das repetições.

O resumo da divisão registra a classe, a sessão, a posição, o caminho e a quantidade de janelas produzidas por cada arquivo. Assim, a composição dos subconjuntos pode ser recuperada sem depender apenas da semente utilizada. Após a escolha da configuração final, o classificador treinado é serializado juntamente com os parâmetros necessários ao pré-processamento. Também são preservados os índices das características selecionadas, a ordem das classes e as informações necessárias para verificar a compatibilidade entre o modelo e a configuração.

Um índice geral mantém uma linha resumida para cada grupo de experimentos, permitindo localizar as tabelas, figuras, relatórios e configurações correspondentes. Esse índice facilita a comparação das execuções sem substituir os arquivos detalhados. Portanto, a reprodução de um experimento depende da preservação conjunta das aquisições utilizadas, da divisão entre os subconjuntos, dos parâmetros do pipeline, da seleção das características, da configuração do classificador e das versões do código e das bibliotecas.

4 Desenvolvimento

Este capítulo apresenta a implementação dos componentes de hardware e software que constituem o sistema desenvolvido. São descritos o Emissor, o Receptor, os mecanismos de transmissão e armazenamento das amostras, as ferramentas executadas no computador e a aplicação utilizada para a inferência contínua. Os procedimentos experimentais, as etapas de pré-processamento, a extração e seleção de características e o protocolo de avaliação dos classificadores são apresentados no Capítulo 3.

4.1 Visão geral da solução desenvolvida

A solução desenvolvida é composta por uma cadeia de aquisição, armazenamento e processamento de dados CSI obtidos em um enlace Wi-Fi entre dois microcontroladores ESP32-S3. O sistema foi dividido em componentes independentes para geração de tráfego, captura das amostras e processamento no computador, o que facilita a verificação de cada etapa e a repetição dos experimentos.

O primeiro ESP32-S3 opera como ponto de acesso e transmite pacotes periodicamente. O segundo atua como estação, recebe esse tráfego, captura os dados CSI associados e os encaminha ao computador pela comunicação serial. Após o armazenamento, os dados são submetidos ao pré-processamento, à extração e seleção de características e à classificação.

O sistema busca distinguir três estados do ambiente: vazio, presença estática e presença em movimento. Não se pretende reconhecer ações específicas, mas estimar a condição de ocupação a partir das alterações provocadas no canal de comunicação.

4.1.1 Requisitos e decisões de projeto

O sistema foi projetado para realizar o sensoriamento sem câmeras e sem exigir que a pessoa monitorada portasse qualquer dispositivo. Essa decisão evita a produção de registros visuais e permite observar a ocupação do ambiente de forma passiva, com base nas alterações provocadas no enlace Wi-Fi.

Também foi priorizado o uso de componentes de baixo custo e disponíveis comercialmente. Foram utilizados dois dispositivos ESP32-S3, que possuem comunicação Wi-Fi integrada e permitem o acesso aos dados CSI por meio do Espressif IoT Development Framework (ESP-IDF). O emprego de dispositivos idênticos simplificou a configuração do enlace e dispensou placas de rede específicas ou equipamentos dedicados à aquisição.

A arquitetura utiliza um único enlace Wi-Fi atravessando a região monitorada. Essa configuração reduz a quantidade de equipamentos e elimina a necessidade de sincronização entre múltiplos transmissores ou receptores. Para limitar variações introduzidas pela própria comunicação, foram mantidos constantes o canal Wi-Fi, a largura de banda, a posição dos dispositivos e a periodicidade de transmissão. Durante as aquisições, o enlace operou no canal 6, no modo High Throughput com largura de canal de 20 MHz (HT20), com largura de banda de 20 MHz, enquanto o Emissor gerou tráfego UDP com intervalo nominal de 20 ms. Esses valores definem a configuração experimental adotada e foram mantidos fixos para que alterações nesses parâmetros não constituíssem uma variável adicional durante a formação do conjunto de dados.

Foram consideradas as classes ambiente vazio, presença estática e presença em movimento. A separação entre as duas condições de presença foi mantida porque uma pessoa pode permanecer no ambiente sem realizar deslocamentos ou movimentos amplos. Nesse caso, uma classificação apenas entre vazio e ocupado não representaria adequadamente as diferenças temporais observadas no CSI.

Os dados foram preservados antes de qualquer etapa de pré-processamento. Em cada recepção aceita, o Receptor captura o vetor CSI disponibilizado pelo ESP-IDF, composto pelas componentes real e imaginária dispostas de forma intercalada, além dos metadados associados à recepção, como RSSI, taxa física, canal, comprimento do vetor e indicadores da camada física. Em seguida, o firmware acrescenta a cada amostra um número sequencial e um marcador temporal, organiza os dados em quadros binários, aplica mecanismos de verificação de integridade e os transmite pela Universal Asynchronous Receiver-Transmitter (UART).

No computador, os quadros recebidos são validados e armazenados antes da aplicação de filtros, normalizações ou transformações matemáticas. Esse procedimento preserva os dados necessários para repetir o processamento com diferentes configurações, sem realizar novas aquisições. Para possibilitar essa flexibilidade, as tarefas dos microcontroladores foram limitadas à comunicação, à captura e ao encaminhamento dos dados, enquanto o processamento foi executado no computador. Essa divisão permitiu modificar filtros, janelas, características e classificadores sem alterar os firmwares de aquisição. Ao mesmo tempo, a escolha de características estatísticas e modelos de menor complexidade considerou a possibilidade de uma futura execução embarcada.

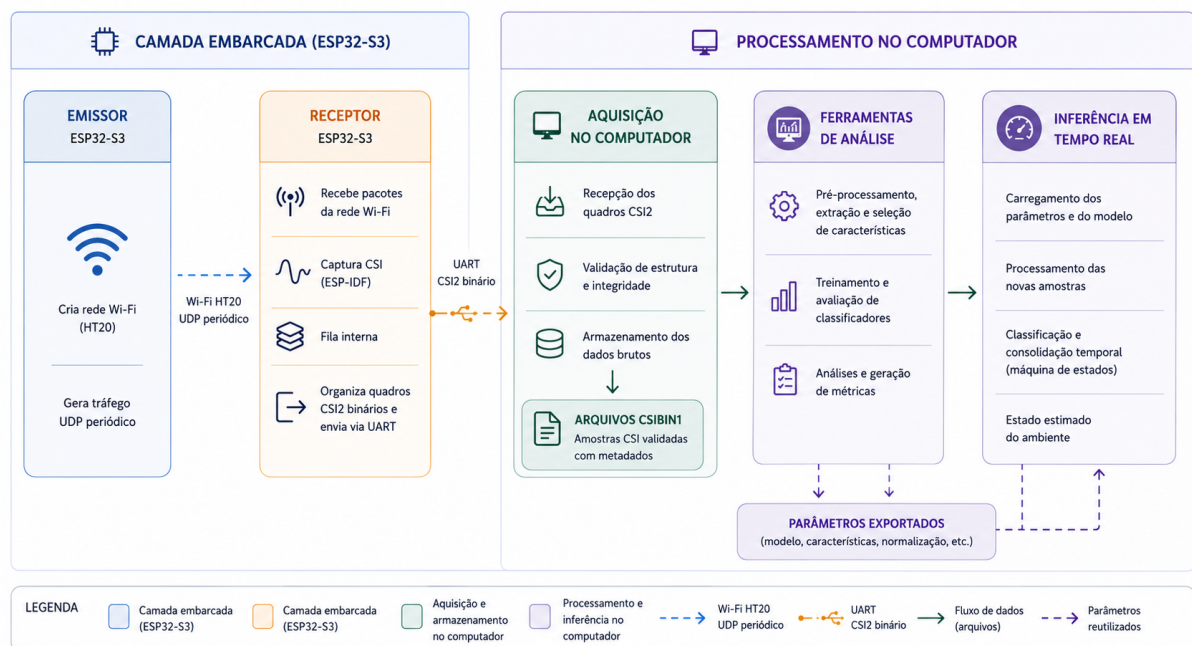
4.1.2 Arquitetura de hardware e software

A arquitetura desenvolvida é formada por três blocos principais: o Emissor, o Receptor e o Processamento no computador. Os dois primeiros são executados em microcontroladores ESP32-S3, enquanto o terceiro reúne as ferramentas desenvolvidas em Python para aquisição, tratamento e análise dos dados.

O Emissor é responsável por criar a rede Wi-Fi experimental e transmitir pacotes UDP periodicamente ao Receptor. O conteúdo desses pacotes não é utilizado diretamente na classificação, pois sua função é manter um fluxo controlado de comunicação e permitir a obtenção contínua de amostras CSI. Por sua vez, o Receptor conecta-se à rede criada pelo Emissor, recebe os pacotes transmitidos e captura o CSI disponibilizado pelo ESP-IDF por meio de uma função de retorno registrada no controlador Wi-Fi. Para evitar bloqueios durante a recepção, essa função apenas copia os dados para uma fila interna. Em seguida, uma tarefa separada organiza as informações em quadros binários e as transmite ao computador pela interface UART.

O bloco de Processamento no computador recebe e valida os quadros transmitidos, armazena as aquisições e executa as etapas de pré-processamento, geração de janelas, extração e seleção de características, treinamento e validação dos classificadores. Na operação em tempo real, as previsões também são encaminhadas a uma máquina de estados responsável pela consolidação temporal das decisões.

Figura 7 – Arquitetura geral do sistema



Fonte: imagem gerada por meio do ChatGPT e revisada pelo autor, 2026.

A separação entre os blocos permite alterar o processamento ou o classificador sem interferir na aquisição embarcada. Da mesma forma, mudanças nas ferramentas de análise não modificam os arquivos brutos, preservando a rastreabilidade dos experimentos.

4.1.3 Fluxo geral dos dados

O fluxo inicia-se com a transmissão periódica de pacotes pelo Emissor. Após a recepção, o Receptor captura o vetor CSI e os metadados associados, copia a amostra para

uma fila interna e a transmite ao computador por meio do protocolo binário CSI2. Esse protocolo organiza cada amostra e seus metadados em quadros próprios para a transmissão serial, incluindo mecanismos de identificação e verificação de integridade.

No computador, a ferramenta de aquisição identifica os quadros CSI2, verifica sua integridade e continuidade e armazena as componentes reais e imaginárias sem processamento prévio. As amostras validadas são gravadas no formato binário CSIBIN1, utilizado para preservar os dados CSI e os metadados necessários ao processamento posterior. A partir desses arquivos, são calculadas as amplitudes e aplicadas as etapas de validação, remoção de valores atípicos, suavização, normalização e redução de redundância.

Os sinais processados são segmentados em janelas temporais, das quais são extraídas características estatísticas e temporais. Parte dessas características é selecionada e fornecida ao classificador, que produz uma estimativa para cada uma das classes consideradas. Os procedimentos utilizados para ajustar essas operações exclusivamente com os arquivos de treinamento são apresentados no Capítulo 3.

Na execução em tempo real, as amostras recebidas são acumuladas em um buffer até que exista informação suficiente para formar uma janela completa. Cada nova janela é processada com os parâmetros exportados durante o treinamento e fornecida ao classificador. As probabilidades resultantes são então encaminhadas a uma camada de decisão temporal, responsável por produzir o estado final apresentado ao usuário.

4.2 Desenvolvimento do ponto de acesso

O firmware do Emissor, executado no primeiro ESP32-S3, é responsável por criar a rede Wi-Fi experimental e gerar o tráfego periódico recebido pelo Receptor. O dispositivo foi configurado no modo ponto de acesso, permitindo estabelecer o enlace sem a utilização de um roteador externo. Os principais parâmetros da rede foram definidos por meio do mecanismo Kconfig do ESP-IDF e permanecem configuráveis sem alterações no código-fonte principal.

A rede utiliza, por padrão, o identificador TCC_CSI_AP e autenticação WPA/Wi-Fi Protected Access 2 with Pre-Shared Key (WPA2-PSK). O canal foi fixado em 6, a largura de banda em HT20, correspondente a 20 MHz, e a quantidade máxima de estações conectadas foi limitada a uma. O canal e a largura de banda permaneceram inalterados durante todas as aquisições, de modo que o conjunto de dados fosse obtido sob uma mesma configuração da camada física. Os principais parâmetros empregados nas aquisições são apresentados na Tabela 2.

Tabela 2 – Principais parâmetros da rede e do tráfego gerado pelo ponto de acesso

Parâmetro	Configuração utilizada
Modo de operação	Ponto de acesso
Identificador padrão	TCC_CSI_AP
Autenticação	WPA/WPA2-PSK
Canal Wi-Fi	6
Largura de banda	HT20, 20 MHz
Máximo de estações	1
Protocolo de transporte	UDP
Endereço de destino	192.168.4.2
Porta de destino	3333
Intervalo de transmissão	20 ms
Taxa nominal	50 pacotes por segundo

A escolha do canal 6 não resultou de uma busca pelo canal com melhor desempenho para o sensoriamento, mas de sua adoção como configuração fixa durante o procedimento experimental. Dessa forma, o canal utilizado deixou de variar entre as aquisições. De maneira semelhante, a largura de 20 MHz foi mantida por meio do modo HT20, estabelecendo uma única configuração de largura de canal para a formação do conjunto de dados. Não foram realizados experimentos sistemáticos variando canal ou largura de banda; portanto, essas escolhas não devem ser interpretadas como configurações ótimas para a tarefa investigada, mas como condições controladas sob as quais o sistema foi desenvolvido e avaliado.

Após a inicialização da rede, o Emissor cria um *socket* UDP e transmite datagramas periodicamente para o endereço Internet Protocol (IP)v4 192.168.4.2, atribuído ao Receptor, utilizando a porta 3333. O intervalo nominal de 20 ms estabelece uma taxa de geração de 50 pacotes por segundo, produzindo uma sequência temporal regular de recepções para a obtenção das amostras CSI. Esse intervalo foi mantido constante em todas as aquisições e não foi tratado como parâmetro de otimização do classificador. A taxa efetivamente armazenada pode diferir do valor nominal devido a perdas no enlace, ao processamento realizado pelo Receptor ou à transmissão serial.

Cada datagrama contém um identificador, um número sequencial e o instante de envio registrado pelo temporizador do ESP32-S3. Esses campos possuem finalidade diagnóstica e não são utilizados diretamente pelo classificador. Também são mantidos contadores de pacotes enviados e erros, permitindo acompanhar a regularidade da transmissão. A periodicidade é controlada com base no instante previsto para cada ciclo, reduzindo o acúmulo do tempo gasto na construção e no envio dos datagramas. A tarefa de transmissão permanece ativa mesmo quando o Receptor está desconectado, permitindo que o tráfego seja retomado após o restabelecimento do enlace, sem necessidade de reinicializar

o Emissor.

4.3 Desenvolvimento do receptor CSI

O firmware do Receptor, executado no segundo ESP32-S3, é responsável por conectar-se ao ponto de acesso, capturar as informações CSI associadas aos pacotes recebidos e encaminhar as amostras ao computador. Para evitar que a transmissão serial interfira na recepção Wi-Fi, a captura e o envio dos dados são realizados separadamente.

4.3.1 Conexão à rede e inicialização da captura

O ESP32-S3 é configurado no modo estação e se conecta à rede utilizando o identificador e a senha definidos no menuconfig. Após a obtenção do endereço IP, o firmware identifica o Basic Service Set Identifier (BSSID) do ponto de acesso e o endereço Media Access Control (MAC) da estação, cria a fila de amostras e a tarefa de saída serial, registra a função de retorno CSI e habilita a captura.

A configuração da captura utiliza os campos de treinamento Legacy Long Training Field (L-LTF) e High Throughput Long Training Field (HT-LTF) e mantém desabilitada a aplicação manual de escala. Os campos habilitados fazem parte da configuração utilizada pelo controlador Wi-Fi para disponibilizar as informações CSI e, juntamente com o modo de operação e a largura de canal, influenciam a composição do vetor obtido em cada recepção. Por esse motivo, a dimensionalidade do vetor CSI observada neste trabalho deve ser associada à configuração empregada e não interpretada como uma quantidade fixa para qualquer sistema Wi-Fi. A escala manual foi mantida desabilitada para preservar os valores segundo a representação fornecida pelo controlador antes do processamento realizado no computador. O receptor UDP permanece aguardando, na porta 3333, os pacotes enviados pelo Emissor. O conteúdo dos datagramas não é encaminhado pela serial, pois sua função é produzir as recepções utilizadas na captura. A economia de energia da interface Wi-Fi também foi desativada, priorizando a regularidade da aquisição.

Quando ocorre uma desconexão, a captura CSI é desabilitada, o filtro de endereços é invalidado e a fila é limpa. Após a reconexão, o processo de inicialização é repetido, permitindo restabelecer o enlace sem reinicializar manualmente o dispositivo.

4.3.2 Aquisição dos dados CSI

A aquisição ocorre por meio de uma função de retorno registrada no controlador Wi-Fi do ESP-IDF. Para cada recepção compatível, essa função recebe o vetor CSI, seu comprimento, os endereços do quadro e os metadados da recepção. O vetor é fornecido

como uma sequência de inteiros sinalizados de 8 bits, com as componentes imaginária e real intercaladas:

$$[Q_0, I_0, Q_1, I_1, \dots, Q_{N-1}, I_{N-1}], \quad (4.1)$$

em que I_k e Q_k correspondem às componentes real e imaginária de cada coeficiente complexo. Como cada coeficiente é representado por dois valores escalares intercalados, o comprimento do vetor bruto não corresponde diretamente à quantidade de coeficientes complexos disponíveis. A estrutura interna do firmware foi dimensionada para admitir até 384 valores escalares, constituindo apenas a capacidade máxima reservada para a cópia de uma amostra e não uma indicação de que 384 subportadoras tenham sido utilizadas. Nas aquisições que compõem o conjunto de dados deste trabalho, foram observados 128 coeficientes complexos por pacote, correspondentes a 256 valores intercalados entre componentes imaginárias e reais. Essa dimensionalidade está associada à configuração de captura empregada durante os experimentos.

Antes da cópia, são verificados a validade dos dados, o comprimento do vetor e sua compatibilidade com a capacidade reservada. Também é aplicado um filtro pelos endereços **MAC**, aceitando apenas quadros originados no ponto de acesso experimental e destinados diretamente à estação.

As amostras aceitas recebem um número sequencial e um marcador temporal em microssegundos. Também são preservados o **RSSI**, a taxa física, o canal e indicadores relacionados à largura de banda, ao modo do sinal, ao Space-Time Block Coding (**STBC**), ao Modulation and Coding Scheme (**MCS**) e à possível invalidade do primeiro valor **CSI**. Esses metadados permitem registrar as condições associadas a cada recepção e verificar sua compatibilidade com a configuração esperada. O vetor é copiado sem cálculo de amplitude, filtragem, normalização ou seleção de posições, e sua compatibilidade estrutural é verificada posteriormente antes do processamento.

4.3.3 Organização por filas e tarefas

A função de retorno **CSI** pertence ao fluxo interno do controlador Wi-Fi e deve executar apenas operações curtas. Por esse motivo, os dados são validados e copiados para uma fila do Free Real-Time Operating System (**FreeRTOS**) com capacidade para 64 amostras. Dessa forma, a inserção ocorre sem bloqueio: quando há espaço, a amostra é enfileirada; quando a fila está cheia, ela é descartada e o evento é registrado, evitando que a função de retorno interrompa o processamento de novos quadros.

Uma tarefa separada consome as estruturas da fila, constrói os quadros do protocolo **CSI2** e realiza a escrita na **UART**. Aproximadamente uma vez por segundo, também é transmitido um quadro de estatísticas contendo os registros de amostras aceitas, enfileiradas,

enviadas, descartadas e inválidas, além da ocupação da fila. Esses contadores permitem identificar perdas ocorridas entre a aceitação da amostra e a transmissão serial, mas não contabilizam pacotes perdidos no próprio enlace Wi-Fi.

Figura 8 – Organização da aquisição CSI por meio de fila e tarefa de saída



Fonte: imagem gerada por meio do ChatGPT e revisada pelo autor, 2026.

4.3.4 Formato de transmissão serial

As amostras são transmitidas pelo protocolo binário CSI2, evitando a conversão individual dos valores para texto. A comunicação utiliza a UART a 921600 baud, com oito bits de dados, sem paridade, um bit de parada e sem controle de fluxo por hardware. Os campos com mais de um byte são organizados em formato *little-endian*.

Cada quadro começa com a identificação ASCII CSI2, seguida pela versão do protocolo, pelo tipo de mensagem e pelo tamanho total. O tipo 1 representa uma amostra CSI, enquanto o tipo 2 transporta as estatísticas periódicas do Receptor. A estrutura do quadro de amostra é apresentada na Tabela 3.

O campo flags reúne a largura de banda, o modo do sinal, a indicação de STBC, a condição do primeiro valor CSI e parte do índice MCS. A integridade é verificada pelo CRC-16/CCITT-FALSE, calculado a partir do campo de versão até o fim dos dados.

Tabela 3 – Organização do quadro de amostra do protocolo CSI2

Campo	Tamanho	Descrição
Identificação	4 bytes	Sequência fixa CSI2, utilizada para localizar o início do quadro.
Versão	1 byte	Versão da estrutura do protocolo.
Tipo	1 byte	Valor 1 para amostras CSI e valor 2 para estatísticas.
Tamanho total	2 bytes	Quantidade total de bytes da mensagem.
Sequencia	4 bytes	Número sequencial da amostra aceita.
Timestamp	8 bytes	Marcador temporal expresso em microssegundos.
Metadados	4 bytes	RSSI, taxa física, canal e campo flags.
Comprimento	2 bytes	Quantidade de valores presentes no vetor CSI.
Vetor CSI	csi_len bytes	Componentes imaginárias e reais intercaladas.
CRC	2 bytes	Código CRC-16/CCITT-FALSE para verificação de integridade.

A identificação CSI2 permite recuperar o alinhamento do fluxo após um quadro inválido ou incompleto. O número sequencial complementa essa verificação ao indicar descontinuidades entre os quadros válidos. O Cyclic Redundancy Check (CRC) identifica alterações no conteúdo recebido, enquanto a sequência permite detectar a ausência de amostras já numeradas pelo firmware.

4.3.5 Confiabilidade da aquisição embarcada

A confiabilidade do Receptor foi tratada por meio da reconexão automática, da reinicialização da captura após a perda do enlace, da desativação da economia de energia, do filtro por endereços, da validação do comprimento, da fila não bloqueante, da transmissão binária e dos mecanismos de CRC, sequência e estatísticas. Esses mecanismos permitem acompanhar diferentes pontos do fluxo, mas não eliminam as perdas inerentes à comunicação sem fio. A avaliação da aquisição considera conjuntamente a taxa de amostras recebidas, os descartes da fila, os erros de CRC e os saltos de sequência.

4.4 Ferramentas de aquisição no computador

Os dados transmitidos pelo Receptor são recebidos, validados e armazenados por ferramentas desenvolvidas em Python. O computador interpreta o fluxo serial, acompanha os diagnósticos da aquisição e grava as amostras que serão utilizadas nas etapas posteriores.

4.4.1 Recepção e interpretação dos quadros

A leitura da porta serial é executada em uma *thread* dedicada, criada separadamente da linha de execução da interface gráfica. Essa divisão impede que a espera por novos dados bloqueie a atualização dos controles e gráficos. A comunicação utiliza a biblioteca `pySerial`, com taxa de 921600 baud, correspondente à configuração do Receptor. Os sinais Data Terminal Ready (DTR) e Request to Send (RTS) permanecem desativados para evitar reinicializações involuntárias do ESP32-S3 durante a abertura da porta.

Os bytes são lidos continuamente e encaminhados ao módulo responsável pela interpretação dos quadros. Como uma leitura pode conter apenas parte de um quadro ou vários quadros consecutivos, esse módulo mantém um buffer interno e realiza a interpretação de forma incremental. A assinatura CSI2 é utilizada para localizar o início das mensagens, após o que são verificados a versão, o tipo e o tamanho informado no cabeçalho.

Quando todos os bytes esperados estão disponíveis, o código CRC recebido é comparado ao valor calculado no computador. Quadros com estrutura ou integridade inválida são rejeitados, e a busca pela próxima assinatura permite recuperar o alinhamento sem interromper a leitura serial. Após essa validação, os quadros do tipo amostra fornecem o número sequencial, os marcadores temporais, os metadados da recepção e o vetor CSI. Durante a interpretação, as componentes imaginárias e reais originalmente intercaladas são separadas, sem aplicação de filtragem, normalização ou cálculo de amplitude. Os quadros de estatísticas, por sua vez, transportam os contadores produzidos pelo Receptor embarcado.

Além do marcador temporal do ESP32-S3, o programa registra o instante de processamento no computador e estima o horário de captura pela associação entre os dois relógios. Essa informação permite organizar temporalmente as amostras sem utilizar apenas o momento em que elas chegaram à interface. A continuidade também é verificada pelo número sequencial. Saltos entre amostras válidas são acumulados como lacunas, enquanto retornos da sequência provocados por uma reinicialização do Receptor são tratados separadamente. Os eventos interpretados são encaminhados à interface por uma fila limitada, cujo preenchimento e eventuais descartes permanecem disponíveis nos diagnósticos.

4.4.2 Armazenamento dos dados brutos

As amostras pertencentes ao intervalo de coleta são mantidas temporariamente em memória e, ao término da aquisição, gravadas no formato binário CSIBIN1. Neste trabalho, dados brutos correspondem aos valores CSI preservados antes de filtragem, normalização, suavização, cálculo de amplitude ou extração de características. O arquivo não representa uma cópia integral da porta serial, pois somente os quadros CSI2 validados

são armazenados.

A versão final do formato utiliza a identificação CSIBIN1 e um campo de versão no início do arquivo. Em seguida, as amostras são gravadas sequencialmente, preservando o rótulo, os marcadores temporais, o número sequencial, o RSSI, a taxa física, o canal, os indicadores da recepção e as componentes imaginárias e reais. Os vetores CSI são armazenados como inteiros sinalizados de 16 bits, mantendo o formato utilizado pelas ferramentas de leitura. O rótulo permanece no conteúdo do arquivo, além de aparecer em seu nome. Essa redundância permite verificar a correspondência entre a classe indicada externamente e os registros armazenados. A preservação das duas componentes também permite alterar o processamento posteriormente sem repetir a aquisição.

Ao final do intervalo configurado, o programa processa os eventos que já se encontram na fila antes de concluir a gravação. Caso nenhuma amostra válida tenha sido recebida, o arquivo não é criado. Os arquivos CSIBIN1 constituem a fonte preservada do conjunto de dados, a partir da qual podem ser reconstruídos os sinais processados, as janelas e as características.

4.4.3 Visualização e diagnóstico da aquisição

A ferramenta principal de aquisição possui uma interface gráfica desenvolvida com PyQt5 e pyqtgraph. A interface permite selecionar a porta serial, iniciar ou interromper a leitura, definir o rótulo, configurar o atraso anterior à coleta, informar sua duração e escolher a pasta de destino.

O atraso permite que o operador inicie o procedimento e se desloque até a posição prevista antes que as amostras sejam incluídas no arquivo. Durante a coleta, uma barra de progresso indica o tempo decorrido e o encerramento do intervalo configurado. Para inspeção, a interface calcula a amplitude de algumas posições do vetor CSI e apresenta sua evolução junto ao RSSI. Esses valores são utilizados somente na visualização e não substituem as componentes armazenadas. Os gráficos permitem perceber interrupções, mudanças abruptas e variações inesperadas durante o procedimento, sem participar da escolha definitiva das características do classificador.

Também são exibidos a taxa de recepção, a ocupação da fila no computador, as lacunas de sequência, os erros de CRC e os descartes registrados no computador e no Receptor. A leitura conjunta desses indicadores permite identificar se uma anormalidade está relacionada ao enlace, à transmissão serial ou ao processamento dos eventos.

4.4.4 Organização dos arquivos coletados

Os arquivos são organizados de forma a preservar a classe, a sessão e a posição utilizadas em cada aquisição. Cada coleta é armazenada em um arquivo binário independente,

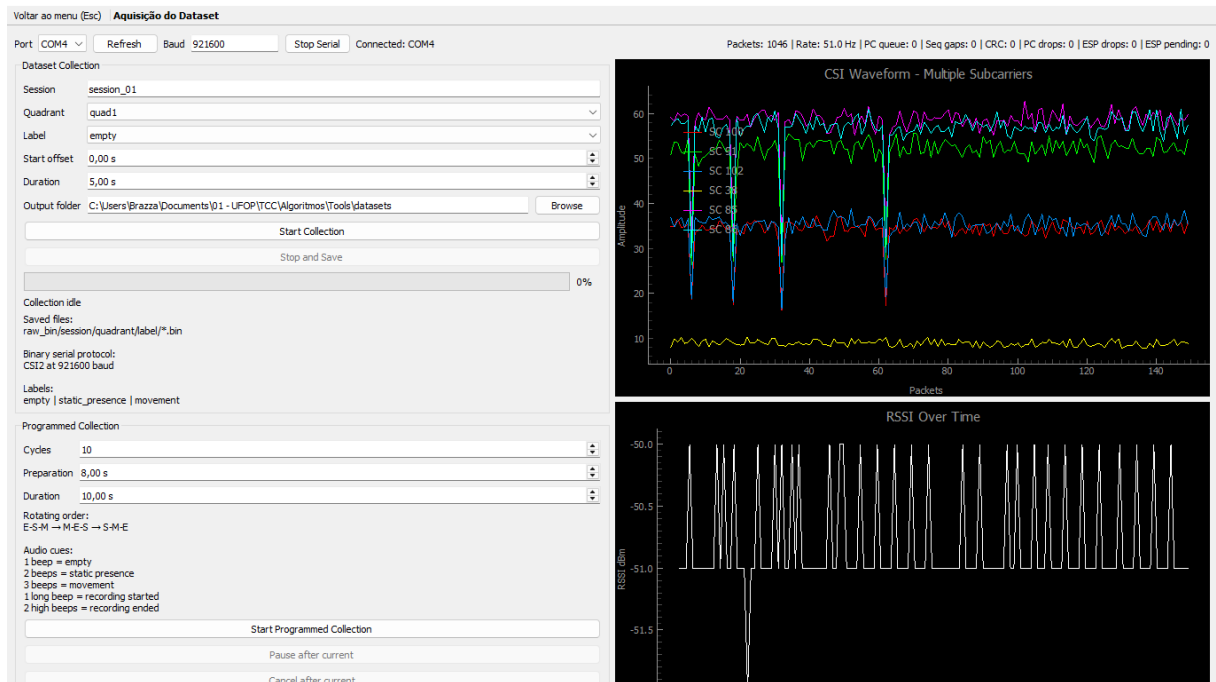


Figura 9 – Interface utilizada para visualização, diagnóstico e armazenamento das aquisições CSI

cujos nomes contêm a classe e o instante de criação:

`<classe>_AAAAMDD_HHMMSS.bin`

As classes são representadas internamente por `<empty>`, `<static_presence>` e `<movement>`. A sessão e a posição são identificadas pela estrutura de diretórios utilizada durante as coletas. A separação por sessão permite distinguir grupos de aquisições realizados em momentos ou condições experimentais diferentes. A identificação da posição, por sua vez, permite analisar a influência da localização do participante dentro do ambiente monitorado.

Para localizar os dados, a ferramenta de carregamento percorre recursivamente o diretório de aquisições e constrói um índice lógico. Cada entrada mantém o caminho do arquivo, a classe, a sessão, a posição e os demais metadados necessários ao processamento. A associação com o arquivo de origem é preservada em todas as etapas seguintes. Dessa forma, é possível contabilizar as aquisições por grupo, retornar ao arquivo correspondente durante a análise e impedir que janelas derivadas da mesma coleta sejam distribuídas entre treinamento e teste.

4.5 Implementação da inferência em tempo real

Após a definição do classificador, foi implementado um fluxo de inferência capaz de processar continuamente as amostras recebidas do Receptor. Essa etapa reutiliza o

mesmo pré-processamento e a mesma organização das características empregados durante o treinamento, evitando diferenças entre os dados utilizados nos experimentos e aqueles classificados durante a execução. Além da classificação de cada janela, o sistema mantém um estado consolidado, apresenta os diagnósticos da aquisição e registra os dados necessários para análise posterior.

4.5.1 Reutilização dos parâmetros treinados

Ao final do treinamento, o classificador é armazenado juntamente com os parâmetros necessários para reproduzir o pipeline. Entre essas informações estão as médias e os desvios utilizados na normalização, os índices das posições do vetor CSI preservadas após a redução de redundância, a ordem das características e as posições selecionadas para compor a entrada do modelo.

Durante a inferência, esses parâmetros são carregados e aplicados diretamente às novas amostras. Dessa forma, a normalização, a redução de redundância e a seleção das características permanecem compatíveis com o conjunto utilizado no treinamento. Ou seja, nenhum desses valores é recalculado quando a execução em tempo real é iniciada. Porém, uma alteração no pré-processamento, na seleção das características ou no classificador exige a geração de um novo conjunto de artefatos, preservando a correspondência entre o modelo e os dados fornecidos à sua entrada.

4.5.2 Processamento contínuo das amostras

As amostras válidas recebidas pela interface serial são convertidas em amplitude e adicionadas a um buffer temporal. A classificação somente é realizada quando existe uma quantidade suficiente de pacotes para formar uma janela completa e disponibilizar o contexto adicional necessário às operações centradas de filtragem. Esse contexto permite aplicar a remoção de valores atípicos e a média móvel de maneira compatível com o processamento realizado sobre os arquivos completos. Para cada janela, são executadas as mesmas etapas utilizadas durante o treinamento: remoção de valores atípicos, suavização, normalização, aplicação dos índices das posições CSI preservadas, extração das características e aplicação dos índices definidos pelo Fisher Score. O vetor resultante é então fornecido ao classificador.

Após a primeira inferência, o buffer é atualizado continuamente. Novas decisões são produzidas de acordo com o passo definido para as janelas, permitindo reutilizar parte das amostras anteriores sem repetir o processamento de todo o fluxo recebido. A saída de cada inferência contém a classe estimada e as probabilidades associadas aos três estados considerados. Essa previsão corresponde ao estado bruto do sistema e ainda pode apresentar alterações entre janelas consecutivas, sendo posteriormente analisada pela etapa

de consolidação temporal.

4.5.3 Consolidação temporal das decisões

As classificações produzidas para janelas consecutivas podem apresentar oscilações, principalmente quando as probabilidades das classes são próximas. Para reduzir mudanças isoladas, foi implementada uma máquina de estados responsável por consolidar temporalmente as previsões do classificador.

A máquina de estados é inicializada na condição de ambiente vazio. A cada nova inferência, a classe com maior probabilidade é considerada como estado candidato. Quando ela coincide com o estado atual, a saída é mantida. Quando é diferente, a previsão é submetida aos critérios de confirmação e, quando aplicável, de confiança.

O critério de confirmação corresponde à quantidade de previsões consecutivas necessárias para aceitar um novo estado. Na configuração final, esse valor foi definido como uma previsão, de modo que a transição pode ocorrer na primeira janela que indique uma classe diferente. Apesar disso, as transições diretas entre ambiente vazio e presença em movimento somente são aceitas quando a probabilidade atribuída ao estado candidato é igual ou superior a 0,60. Caso esse limite não seja atingido, o estado anterior é mantido. Para as transições que envolvem a presença estática, não é aplicado esse limiar adicional de confiança.

Essa etapa não modifica as probabilidades nem realiza um novo treinamento. Sua função é apenas organizar a sequência das previsões e produzir uma saída mais estável para a operação em tempo real. As configurações avaliadas e os efeitos da consolidação são apresentados no capítulo de resultados.

4.5.4 Interface gráfica de operação

A operação do sistema foi reunida em uma aplicação gráfica desenvolvida em Python com a biblioteca PyQt5. A interface permite acessar as ferramentas de aquisição, treinamento e inferência em tempo real sem a necessidade de executar individualmente os programas que compõem o projeto.

Durante o monitoramento, são apresentados o estado bruto produzido pelo classificador, o estado consolidado pela máquina de estados, as probabilidades das classes, a taxa de recepção, o preenchimento do buffer, o tempo de inferência e os contadores de diagnóstico. A comunicação com o Receptor é realizada por meio da biblioteca pySerial, enquanto os gráficos e as informações atualizadas durante a execução utilizam a biblioteca pyqtgraph.

Quando a máquina de estados aceita uma mudança, a aplicação pode anunciar o novo estado por síntese de voz. Também podem ser armazenadas as amostras, as previsões

e os metadados produzidos durante o monitoramento, possibilitando a análise posterior das execuções. A interface atua apenas como meio de operação e visualização, sem modificar o processamento ou as decisões produzidas pelo classificador. A Figura 10 apresenta a interface utilizada durante a operação contínua do sistema.

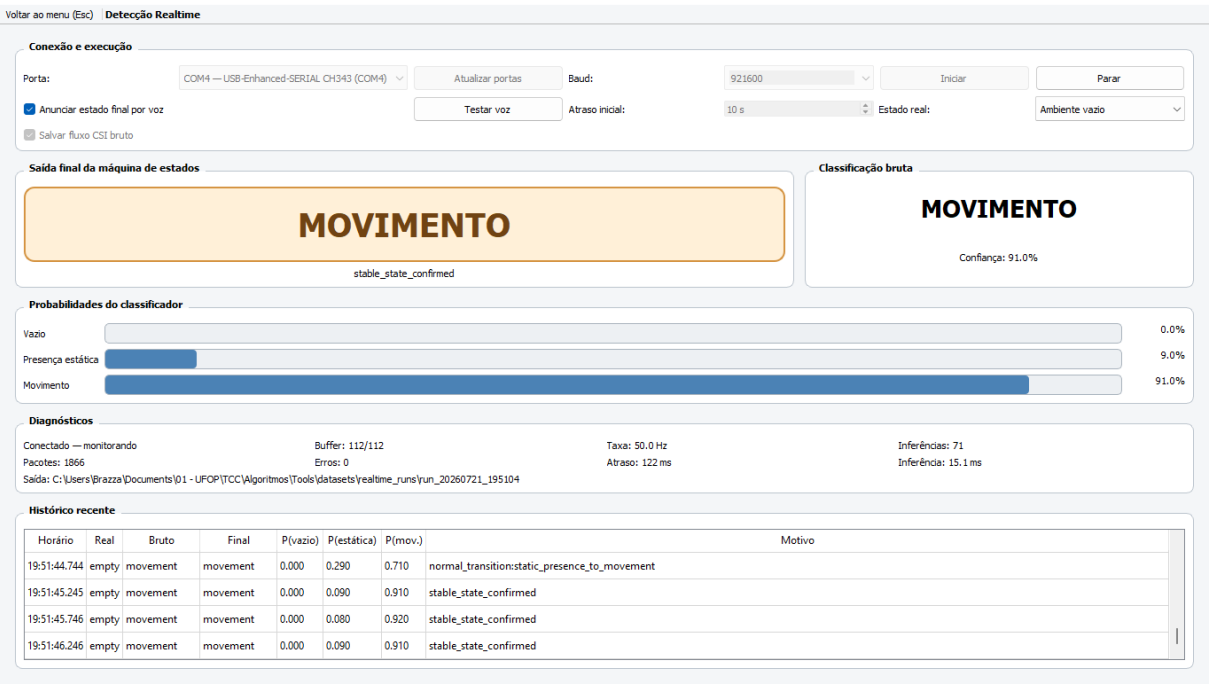


Figura 10 – Interface gráfica utilizada na inferência em tempo real

5 Resultados

Este capítulo apresenta os resultados obtidos na formação do conjunto de dados, na definição da configuração do pipeline, na comparação dos classificadores e na avaliação do sistema em tempo real. As divisões entre treinamento e teste foram realizadas por arquivo, impedindo que janelas provenientes de uma mesma aquisição fossem distribuídas entre os dois subconjuntos.

Além do desempenho global, foram analisadas as diferenças entre as classes e as zonas de coleta. Também foram realizados testes nos quais uma zona completa foi mantida fora do treinamento, permitindo observar o comportamento do modelo em posições não representadas diretamente durante o ajuste.

5.1 Caracterização do conjunto de dados

Após a verificação estrutural das aquisições, foram preservados 300 arquivos válidos, distribuídos igualmente entre ambiente vazio, presença estática e presença em movimento. O conjunto foi organizado em duas sessões e cinco zonas, com dez aquisições para cada combinação entre sessão, zona e classe. Todas as aquisições das classes de presença estática e presença em movimento foram realizadas com o mesmo participante.

Ao todo, foram armazenados 149.141 pacotes CSI válidos. As aquisições apresentaram duração média de 9,98 segundos e mantiveram 128 coeficientes complexos por pacote. Essa dimensionalidade corresponde à configuração de comunicação e captura utilizada durante os experimentos e não deve ser interpretada como uma quantidade fixa de coeficientes para qualquer configuração Wi-Fi. A Tabela 4 reúne as principais características do conjunto final.

A quantidade de pacotes apresentou pequenas diferenças entre as classes, sem comprometer o equilíbrio do conjunto. Todos os arquivos mantiveram a mesma dimensionalidade, evitando a necessidade de completar ou truncar os vetores antes do processamento.

5.2 Seleção da configuração e do classificador

Esta seção apresenta os experimentos utilizados para definir a configuração temporal, o grau de redução dimensional e o classificador empregado na avaliação final e na operação em tempo real.

Tabela 4 – Resumo do conjunto de dados utilizado nos experimentos

Indicador	Resultado
Arquivos válidos	300
Sessões de coleta	2
Zonas por sessão	5
Classes	3
Arquivos por classe	100
Arquivos por combinação entre sessão, zona e classe	10
Pacotes CSI válidos	149.141
Duração média por arquivo	9,98 s
Coefficientes complexos por pacote	128
Taxa mediana de aquisição	50,00 Hz

5.2.1 Configuração temporal

Foram avaliadas oito combinações entre o tamanho da janela e o passo utilizado para gerar novas amostras. A comparação foi realizada em três divisões por arquivos, mantendo todas as janelas de uma aquisição no mesmo subconjunto. O Macro F1 foi utilizado como principal critério, pois atribui a mesma importância às três classes.

A Figura 11 apresenta os resultados obtidos. A configuração mais curta, com janela de 0,25 segundo e passo de 0,10 segundo, alcançou Macro F1 de 0,7071. O desempenho aumentou com o acréscimo de contexto temporal, chegando a 0,8249 com janela de 1,5 segundo e a 0,8294 com janela de 2 segundos.

A janela de 2 segundos foi selecionada por apresentar o maior Macro F1 médio. Com a taxa nominal de 50 Hz, ela corresponde a 100 pacotes. O passo de 0,5 segundo equivale a 25 pacotes e permite produzir novas decisões antes do término de uma janela completa, sem aumentar excessivamente a sobreposição e o custo de processamento. Essa configuração gerou 5.039 janelas no conjunto completo.

5.2.2 Redução dimensional e escolha do modelo

Foram avaliados limiares de correlação entre 0,85 e 0,99. Os valores entre 0,85 e 0,95 apresentaram desempenho próximo, enquanto limiares superiores não produziram melhora. Embora o limiar de 0,85 tenha apresentado o maior Macro F1 médio nessa comparação, sua diferença em relação ao valor de 0,95 foi de apenas 0,0038. O valor de 0,95 foi mantido por remover parte da redundância sem restringir excessivamente o conjunto de posições do vetor CSI fornecido à seleção de características.

Após a ordenação pelo Fisher Score, foram avaliados subconjuntos contendo entre 30 e 160 características, além do vetor completo, cuja dimensão média era de 1.078 posições.

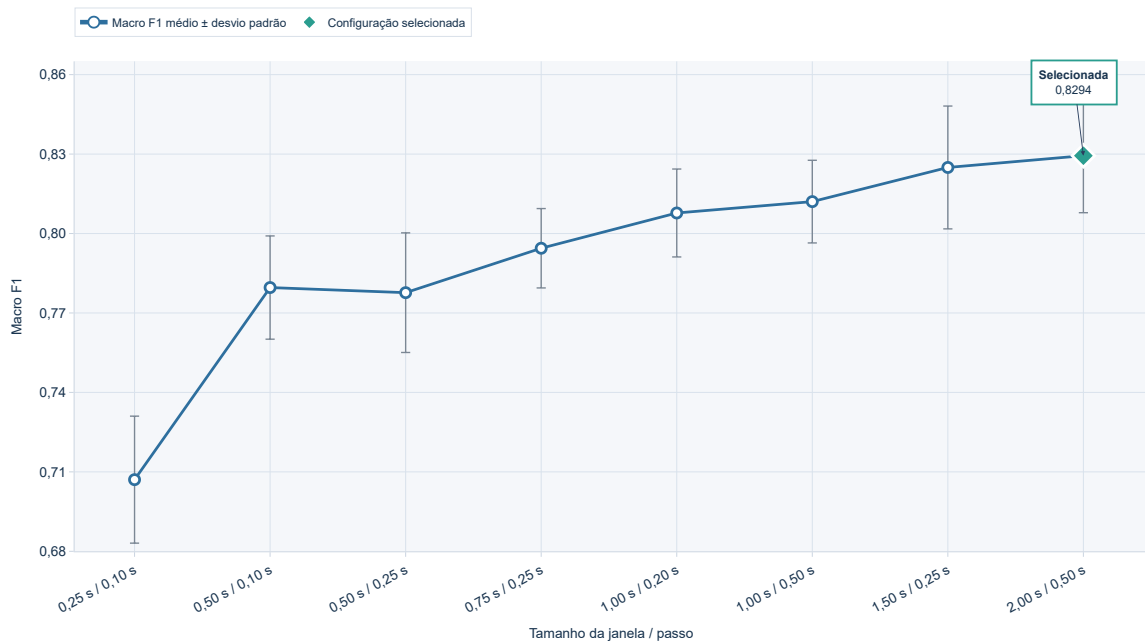


Figura 11 – Macro F1 obtido para as combinações de tamanho de janela e passo

O desempenho aumentou até aproximadamente 100 características e passou a apresentar pouca variação acima desse valor, com resultados próximos entre os subconjuntos de 126 e 160 atributos. A utilização de 160 características manteve cerca de 15% das posições do vetor completo, correspondendo a uma redução dimensional aproximada de 85%. Nos modelos utilizados nessa comparação, essa redução resultou em uma diminuição de Macro F1 entre 0,01 e 0,02 em relação ao uso de todas as características. Em seguida, foram comparados classificadores com diferentes estruturas. A Tabela 5 apresenta os modelos mais representativos e os três candidatos que alcançaram os melhores resultados.

Tabela 5 – Comparação entre os principais classificadores avaliados

Classificador	Caract.	Macro F1	F1 pond.	AUC-ROC	AUC-PR	F1 movimento
Árvore de decisão	126	0,8035	0,8036	0,8867	0,7903	0,9008
Regressão logística	≈ 1080	0,8089	0,8090	0,9219	0,8508	0,9268
<i>Gradient Boosting</i>	≈ 1080	0,8409	0,8410	0,9459	0,9001	0,9247
XGBoost	≈ 1080	0,8384	0,8385	0,9457	0,9001	0,9235
<i>Extra Trees</i>	160	0,8402	0,8403	0,9477	0,9017	0,9318

A árvore de decisão e a regressão logística apresentaram desempenho inferior ao dos métodos baseados em conjuntos de árvores. Entre estes, o *Gradient Boosting*, o Extreme Gradient Boosting (**XGBoost**) e o *Extra Trees* obtiveram resultados muito próximos, com diferença de apenas 0,0007 entre o maior Macro F1 observado e o valor alcançado pelo modelo posteriormente selecionado.

Embora não tenha liderado essa métrica, o modelo escolhido apresentou os maiores valores de **AUC-ROC** e **AUC-PR**, além do melhor F1 para a classe de presença em movi-

mento. Também utilizou apenas 160 características, enquanto as melhores configurações dos outros dois candidatos empregaram o vetor completo. Seu tempo médio de treinamento foi de aproximadamente 1,12 segundo, inferior aos 5,03 segundos registrados pelo [XGBoost](#) e aos 83,56 segundos do *Gradient Boosting*.

Diante desse equilíbrio entre desempenho preditivo, redução dimensional e custo computacional, foi selecionado o *Extra Trees* com 100 árvores, balanceamento das classes e 160 características para a avaliação final e para a integração com o sistema em tempo real.

5.3 Desempenho do modelo selecionado

Esta seção apresenta o desempenho do modelo selecionado nas diferentes divisões por arquivo, considerando as métricas gerais e por classe, a matriz de confusão e a influência da zona de coleta.

5.3.1 Desempenho geral e por classe

O modelo selecionado foi avaliado em dez divisões distintas por arquivos. Em cada repetição, o pré-processamento, a seleção das características e o treinamento foram novamente ajustados apenas com as aquisições destinadas ao treinamento.

A acurácia média foi de $0,8433 \pm 0,0237$, o Macro F1 de $0,8402 \pm 0,0242$ e o F1 ponderado de $0,8403 \pm 0,0242$. Também foram obtidos [AUC-ROC](#) macro de $0,9477 \pm 0,0133$ e [AUC-PR](#) macro de $0,9017 \pm 0,0254$. O Macro F1 variou entre 0,8024 e 0,8892, mostrando que a composição dos arquivos de teste influenciou o resultado, embora todas as divisões tenham permanecido acima de 0,80. As métricas agregadas para cada classe são apresentadas na Tabela 6.

Tabela 6 – Desempenho do classificador para cada classe

Classe	Precisão	Revocação	F1
Ambiente vazio	0,7926	0,9019	0,8432
Presença estática	0,8224	0,6830	0,7455
Presença em movimento	0,9206	0,9439	0,9318

A presença em movimento apresentou o melhor resultado, com F1 de 0,9318. A presença estática obteve o menor valor, igual a 0,7455, principalmente pela redução de sua revocação. A matriz de confusão normalizada, apresentada na Figura 12, permite observar a distribuição dos erros.

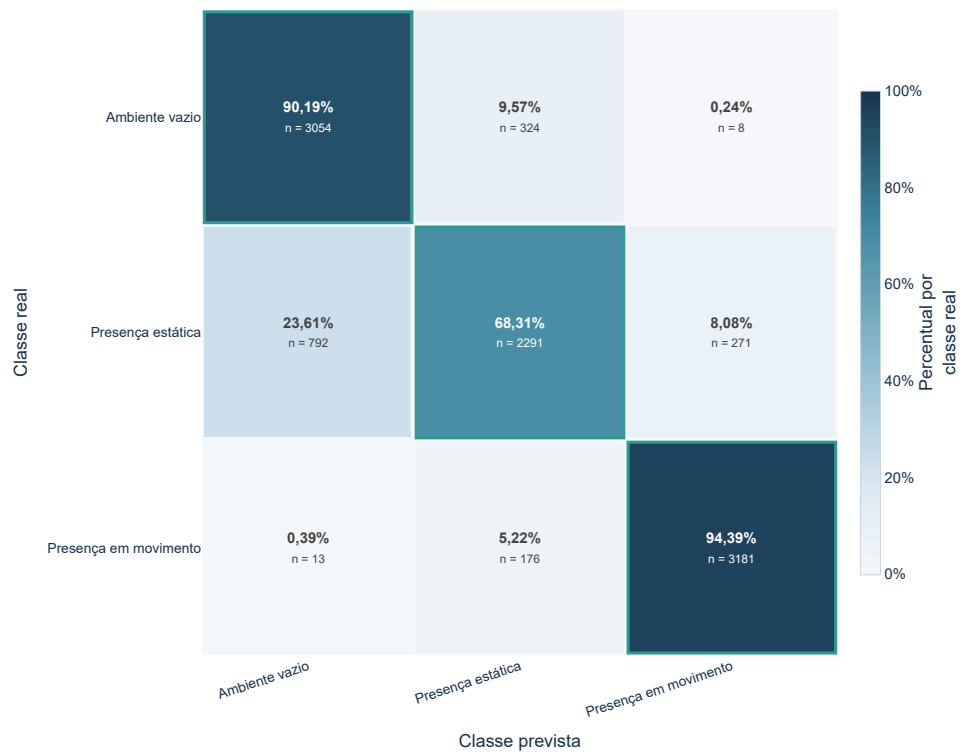


Figura 12 – Matriz de confusão normalizada do classificador selecionado

O classificador identificou corretamente 90,19% das janelas de ambiente vazio, 68,31% das janelas de presença estática e 94,39% das janelas de presença em movimento. O principal erro ocorreu na presença estática, pois 23,61% de suas janelas foram classificadas como ambiente vazio. A confusão direta entre vazio e movimento permaneceu abaixo de 0,4% nos dois sentidos. Os diagnósticos binários reforçaram esse comportamento. A separação entre presença estática e movimento alcançou Macro F1 de 0,9301, enquanto a distinção entre ambiente vazio e presença obteve 0,8732. Também foi testada uma arquitetura hierárquica em duas etapas, mas seu Macro F1 de 0,8371 ficou ligeiramente abaixo da classificação multiclasse direta, que foi mantida.

5.3.2 Influência da zona de coleta

A posição do participante em relação ao Emissor e ao Receptor pode alterar sua influência sobre os percursos de propagação do sinal. Por esse motivo, o desempenho foi analisado separadamente para as cinco zonas, tanto nas divisões convencionais quanto em testes nos quais uma zona completa era mantida fora do treinamento.

Tabela 7 – Desempenho do classificador em função da zona de coleta

Zona	Acurácia média	Macro F1 com zona não vista
Zona 1	$0,8662 \pm 0,0886$	0,8550
Zona 2	$0,8313 \pm 0,0502$	0,7935
Zona 3	$0,8797 \pm 0,0655$	0,8665
Zona 4	$0,7610 \pm 0,0479$	0,7464
Zona 5	$0,8782 \pm 0,0448$	0,8447

As zonas 3 e 5 apresentaram as maiores acurácias nas divisões convencionais, enquanto a zona 4 obteve o menor resultado. O mesmo comportamento apareceu quando cada região foi completamente retirada do treinamento: a zona 3 alcançou Macro F1 de 0,8665 e a zona 4 permaneceu com o menor valor, igual a 0,7464. A presença estática apresentou o menor desempenho em todas as zonas, enquanto a zona 4 reduziu os resultados das três classes. Esses resultados mostram que o desempenho em regiões não representadas diretamente durante o treinamento não foi uniforme e que a relação entre a posição do participante e a geometria do enlace influencia a representação obtida pelo sistema.

5.4 Avaliação do sistema em tempo real

O funcionamento contínuo foi avaliado sobre uma gravação contendo mudanças sucessivas entre ambiente vazio, presença estática e presença em movimento. O arquivo foi processado pelo mesmo pipeline exportado para a aplicação em tempo real, produzindo 286 janelas consecutivas. Para o cálculo das acurácias, foram consideradas somente as 221 janelas associadas a uma das três classes e com pureza mínima de 0,99. Os intervalos que continham transições entre estados não participaram da comparação.

Tabela 8 – Efeito da consolidação temporal sobre o cenário contínuo

Indicador	Resultado
Janelas utilizadas na avaliação	221
Acurácia das previsões brutas	0,8914
Acurácia após a consolidação	0,9140
Ganho absoluto de acurácia	0,0226

A consolidação temporal elevou a acurácia de 0,8914 para 0,9140. A principal alteração ocorreu nas regiões de ambiente vazio, nas quais a quantidade de acertos aumentou de 57 para 62 entre 73 janelas. Os resultados de presença estática e movimento foram preservados, com 66 acertos em 74 janelas e 74 acertos em 74 janelas, respectivamente. A Figura 13 apresenta a sequência temporal do estado real, da classificação bruta e do estado consolidado, permitindo observar os pontos em que a máquina de estados manteve ou corrigiu a decisão.

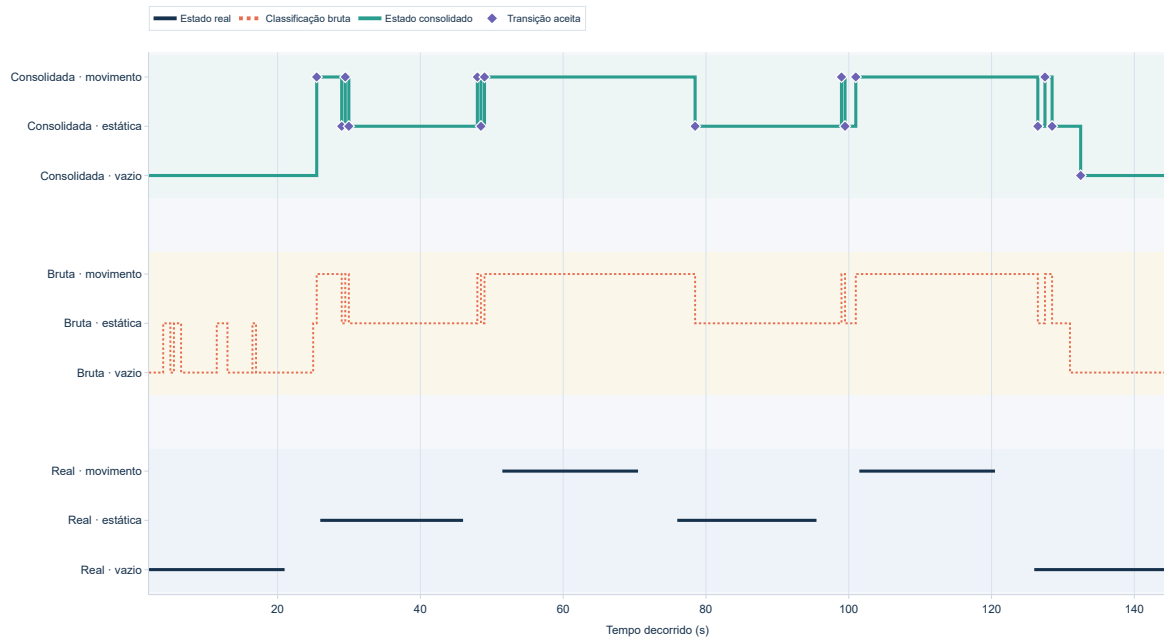


Figura 13 – Comparação temporal entre o estado real, a classificação bruta e o estado consolidado

5.5 Limitações e dependências da configuração

Esta seção apresenta as principais limitações identificadas durante o desenvolvimento e a avaliação do sistema. Inicialmente, são discutidas as limitações relacionadas ao procedimento experimental e ao conjunto de dados utilizado. Em seguida, são apresentadas as dependências associadas à tecnologia e à configuração de hardware e comunicação adotadas nos experimentos.

5.5.1 Limitações experimentais

A principal dificuldade observada na classificação foi a diferenciação entre ambiente vazio e presença estática. Quando o participante permanecia com pouca movimentação, as alterações temporais produzidas no canal podiam se aproximar daquelas observadas no ambiente vazio. Esse comportamento aparece na revocação inferior da presença estática e na sua confusão predominante com a classe de ambiente vazio. A variação de desempenho entre as zonas, principalmente na zona 4, também evidencia a influência da posição ocupada em relação à geometria do enlace.

O conjunto de dados foi obtido em um único ambiente e com um único participante nas classes que envolviam presença humana. Dessa forma, os resultados não permitem determinar em que medida diferenças entre indivíduos, como características físicas ou padrões naturais de movimento, afetariam o desempenho do modelo. As posições do Emissor e do Receptor também permaneceram fixas durante as aquisições. A proposta

considera que uma instalação do sistema possa ser calibrada para as condições específicas do ambiente em que será utilizada; portanto, a utilização de um único cenário não pressupõe a transferência direta do mesmo modelo para outros locais. Entretanto, este trabalho não definiu nem avaliou um procedimento prático de calibração que estabeleça a quantidade mínima de dados necessária ou os critérios para uma nova calibração após alterações no posicionamento dos dispositivos, na organização do ambiente ou nas condições de propagação.

A avaliação contínua também foi realizada sobre uma única sequência contendo transições entre as três condições consideradas. Embora essa análise permita verificar o funcionamento integrado do pipeline e da consolidação temporal, outras sequências contínuas são necessárias para caracterizar sua estabilidade em períodos mais longos e diante de diferentes transições. Além disso, o sistema não foi avaliado em ambientes reais de aplicação nos quais a formação de imagens seja indesejada ou sensível. Os resultados obtidos demonstram o funcionamento técnico da abordagem no cenário experimental adotado, mas não estabelecem sua adequação operacional para um domínio específico.

5.5.2 Limitações da tecnologia e da configuração utilizada

Os resultados obtidos estão associados à configuração de hardware e comunicação utilizada nos experimentos: dois dispositivos ESP32-S3 formando um único enlace Wi-Fi, canal 6, modo HT20 com largura de 20 MHz, captura utilizando os campos L-LTF e HT-LTF, intervalo nominal de transmissão de 20 ms e posições fixas para o Emissor e o Receptor. Nessas condições, as aquisições apresentaram 128 coeficientes complexos CSI por pacote. Alterações nesses parâmetros podem modificar as condições de propagação, a interferência observada, a regularidade temporal das amostras e a própria composição do vetor CSI.

Não foram realizados experimentos sistemáticos variando o canal Wi-Fi, a largura de banda, os campos utilizados na obtenção do CSI, a periodicidade de transmissão ou a quantidade e o posicionamento dos enlaces. Dessa forma, os resultados não permitem determinar o efeito individual desses parâmetros sobre a classificação nem identificar uma configuração ótima para o sensoriamento. As escolhas adotadas devem ser interpretadas como condições controladas utilizadas para desenvolver e avaliar o sistema, e não como parâmetros necessariamente superiores a outras configurações possíveis.

Como o modelo é treinado a partir da representação produzida sob essas condições, alterações relevantes na configuração da comunicação ou na geometria do enlace podem modificar a distribuição dos dados fornecidos ao pipeline. Nesses casos, a reutilização direta dos parâmetros de normalização, das posições preservadas, das características selecionadas e do classificador não deve ser assumida sem nova avaliação. Essa dependência reforça a necessidade de tratar a configuração de aquisição como parte integrante da calibração do

sistema.

6 Conclusão

Este trabalho investigou a utilização de informações do estado do canal de uma comunicação Wi-Fi para diferenciar três condições de ocupação em um ambiente interno: ambiente vazio, presença estática e presença em movimento. Para isso, foi desenvolvido um sistema completo de aquisição e processamento baseado em dois microcontroladores ESP32-S3, sem o emprego de câmeras ou a necessidade de que a pessoa monitorada portasse algum dispositivo. A proposta foi orientada à obtenção de informações gerais de ocupação em situações nas quais o monitoramento sem contato e sem formação de imagens possa ser desejável.

A solução desenvolvida compreendeu a criação de um enlace Wi-Fi controlado, a geração periódica de tráfego, a captura das amostras CSI, a transmissão binária dos dados ao computador e seu armazenamento antes da aplicação de qualquer transformação. Também foram implementadas ferramentas para visualização e diagnóstico das aquisições, pré-processamento dos sinais, formação de janelas temporais, extração e seleção de características, treinamento dos classificadores e inferência contínua. Essa integração permitiu avaliar não apenas o modelo de classificação, mas o fluxo completo entre a recepção dos pacotes e a apresentação do estado estimado.

O conjunto de dados final foi formado por 300 aquisições válidas, distribuídas igualmente entre as três classes, duas sessões e cinco zonas do ambiente, sendo todas as condições com presença humana realizadas com um único participante. Foram armazenados 149.141 pacotes CSI válidos, com taxa mediana de 50 Hz e sem lacunas detectadas na sequência das amostras aceitas pelo Receptor. Sob a configuração de comunicação e captura adotada, cada pacote manteve 128 coeficientes complexos CSI. Esses resultados indicam que a arquitetura implementada foi capaz de manter uma aquisição suficientemente regular para a formação das séries temporais utilizadas nos experimentos.

A avaliação dos classificadores foi realizada com separação por arquivo, mantendo todas as janelas provenientes de uma mesma aquisição no mesmo subconjunto. Esse procedimento evitou que trechos temporalmente relacionados fossem distribuídos entre treinamento e teste e produzissem uma estimativa excessivamente otimista do desempenho. Em cada repetição, as operações dependentes da distribuição dos dados também foram ajustadas somente com os arquivos de treinamento. Entre os modelos avaliados, o *Extra Trees* com 100 árvores e 160 características foi selecionado para a configuração final. Embora seu Macro F1 tenha permanecido próximo ao obtido pelo *Gradient Boosting*, o modelo escolhido apresentou as maiores AUC-ROC e AUC-PR, o melhor F1 para presença em movimento e utilizou uma quantidade de características consideravelmente menor do que

o vetor completo.

Nas dez divisões por arquivos, o modelo selecionado apresentou acurácia média de $0,8433 \pm 0,0237$ e Macro F1 de $0,8402 \pm 0,0242$. A presença em movimento foi a classe identificada com maior consistência, alcançando F1 de 0,9318. A presença estática apresentou o menor resultado, com F1 de 0,7455, principalmente porque parte dessas janelas foi classificada como ambiente vazio. Dessa forma, o sistema apresentou maior facilidade para reconhecer alterações dinâmicas no canal do que para distinguir a influência produzida por um participante que permanecesse com pouca movimentação. O desempenho também variou de acordo com a zona ocupada. A zona 4 apresentou os menores resultados tanto nas divisões convencionais quanto na avaliação em que seus arquivos foram mantidos fora do treinamento. Essa diferença mostra que a posição do participante em relação ao Emissor e ao Receptor influencia as alterações observadas no canal e, consequentemente, a representação fornecida ao classificador. Os experimentos com zonas não representadas diretamente durante o treinamento também mostraram que essa influência não ocorre de maneira uniforme em toda a região monitorada.

A integração com a aplicação em tempo real confirmou que as amostras recebidas pela interface serial podiam ser acumuladas, processadas e classificadas continuamente. No cenário temporal avaliado, a consolidação das previsões aumentou a acurácia de 0,8914 para 0,9140, principalmente pela correção de oscilações relacionadas ao ambiente vazio. O sistema também foi executado com dados recebidos diretamente do Receptor, mantendo o processamento e a apresentação dos estados sem interromper o fluxo de aquisição.

Diante desses resultados, considera-se que o objetivo proposto foi alcançado dentro do escopo experimental adotado. O trabalho demonstrou que uma estrutura formada por dois dispositivos ESP32-S3 e um único enlace Wi-Fi pode ser utilizada para obter informações relacionadas à ocupação e ao nível geral de atividade de um ambiente sem produzir imagens e sem exigir o uso de sensores pela pessoa monitorada. A solução apresentou melhor desempenho na identificação de presença em movimento, mas também demonstrou capacidade de distinguir as três condições consideradas, ainda que a separação entre ambiente vazio e presença estática permaneça como o principal desafio observado.

Os resultados permanecem condicionados ao escopo experimental e à configuração utilizada. O conjunto de dados foi obtido em um único ambiente e com um único participante nas classes envolvendo presença humana, enquanto o Emissor e o Receptor permaneceram em posições fixas. Além disso, foram utilizados dois ESP32-S3 formando um único enlace, canal 6, modo HT20 com largura de 20 MHz, captura com L-LTF e HT-LTF e intervalo nominal de transmissão de 20 ms. Não foram avaliadas sistematicamente outras combinações desses parâmetros, de modo que os resultados não permitem determinar o efeito individual de mudanças no canal, na largura de banda, na configuração da captura ou na geometria do enlace.

A proposta considera que uma instalação do sistema possa ser calibrada para as condições específicas do ambiente em que será utilizada, não pressupondo que um único modelo treinado seja diretamente transferível entre diferentes locais ou configurações. Entretanto, o trabalho não estabeleceu um procedimento formal de calibração ou recalibração que determine a quantidade mínima de aquisições necessária, a influência da participação de diferentes indivíduos ou os critérios para repetir o treinamento após alterações no ambiente, no posicionamento dos dispositivos ou na configuração da comunicação. A avaliação contínua também foi realizada sobre uma única sequência de transições, limitando as conclusões sobre a estabilidade da camada de decisão em períodos prolongados.

Como trabalhos futuros, recomenda-se avaliar de forma sistemática o procedimento de calibração do sistema. Essa investigação deve considerar diferentes participantes, quantidades de aquisições utilizadas no treinamento e alterações controladas no posicionamento dos dispositivos e na organização física do ambiente, permitindo determinar em quais condições uma nova calibração se torna necessária. Também é relevante repetir as avaliações em outros ambientes e períodos, não com o objetivo de pressupor a transferência direta de um único modelo, mas para verificar se o procedimento de aquisição, treinamento e calibração mantém comportamento consistente em diferentes instalações.

Também devem ser investigadas diferentes configurações da comunicação Wi-Fi. Experimentos controlados variando o canal, a largura de banda, os campos utilizados na captura do CSI, a periodicidade de transmissão e a quantidade ou o posicionamento dos enlaces permitiriam quantificar a influência desses parâmetros sobre a representação obtida e sobre o desempenho da classificação. Essa análise poderia identificar configurações mais adequadas à tarefa ou determinar quais parâmetros precisam permanecer constantes durante a calibração e a operação do sistema.

Outra possibilidade consiste em investigar representações adicionais do CSI e características mais adequadas à diferenciação entre ambiente vazio e presença estática, principal fonte de erro observada na classificação. A incorporação de informações de fase, quando suficientemente estáveis, e a avaliação de métodos específicos para detectar alterações de baixa intensidade podem auxiliar nessa distinção. Também deve ser realizada uma validação independente da camada de decisão temporal, utilizando gravações contínuas que não tenham participado da definição de seus parâmetros. Nessa etapa, recomenda-se avaliar a incorporação de regras de negócio fundamentadas nas transições fisicamente possíveis entre os estados monitorados. Em particular, a passagem direta de presença estática para ambiente vazio não representa uma sequência fisicamente coerente, pois a saída da pessoa do ambiente necessariamente produz algum nível de movimento antes que a condição de ausência seja estabelecida. Dessa forma, a máquina de estados pode adicionar camadas de verificação a essa transição direta e exigir a identificação intermediária de presença em movimento, além de critérios mínimos de confiança e persistência temporal.

Essa restrição pode reduzir falsas indicações de desocupação causadas por períodos nos quais uma pessoa presente permanece imóvel e produz alterações pouco perceptíveis no canal.

Por fim, a redução do pipeline e do modelo para execução embarcada representa uma continuidade possível deste trabalho. Embora o processamento tenha sido realizado no computador, a configuração final utiliza um subconjunto reduzido de características e um classificador com tempo de inferência compatível com operação contínua. A implementação dessas etapas no próprio dispositivo ou em uma unidade de processamento local permitiria avaliar o funcionamento do sistema sem dependência permanente de um computador e aproximaria o protótipo de aplicações práticas de monitoramento.

Referências

- BISIO, I. et al. Comparison of sensing capabilities using different csi detection tools. In: *ICC 2025 - IEEE International Conference on Communications*. [S.l.: s.n.], 2025. p. 6426–6431. Citado 4 vezes nas páginas 16, 19, 21 e 24.
- GAIBA, F. et al. Wi-fi sensing for human identification through esp32 devices: An experimental study. In: *2024 IEEE 21st Consumer Communications Networking Conference (CCNC)*. [S.l.: s.n.], 2024. p. 206–209. Citado na página 23.
- KANG, D.; KIM, M.; PARK, S. A study on fusion acquisition of image and csi data for posture estimation in indoor environments. In: *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*. [S.l.: s.n.], 2024. p. 1312–1314. Citado na página 14.
- KHALIULLIN, R. F.; SULIMOV, A. I.; BUSSE, A. A. Experimental derivation of channel correlation functions using the esp32-csi-tool. In: *2024 IEEE 9th All-Russian Microwave Conference (RMC)*. [S.l.: s.n.], 2024. p. 135–140. Citado 2 vezes nas páginas 16 e 22.
- MENDEZ, D. et al. On tinyml wifi fingerprinting-based indoor localization: Comparing rssi vs. csi utilization. In: *2024 IEEE 21st Consumer Communications Networking Conference (CCNC)*. [S.l.: s.n.], 2024. p. 1–6. Citado 3 vezes nas páginas 14, 21 e 23.
- MIAO, F. et al. Wi-fi sensing techniques for human activity recognition: Brief survey, potential challenges, and research directions. *ACM Comput. Surv.*, Association for Computing Machinery, New York, NY, USA, v. 57, n. 5, jan. 2025. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/3705893>>. Citado 4 vezes nas páginas 14, 16, 18 e 21.
- NDEBELE, F. et al. Non-contact monitoring to support care in acute inpatient mental health. *Journal of Mental Health*, v. 33, n. 3, p. 320–325, 2024. Citado na página 16.
- R, A. V. et al. Privacy-free human sensing using bilstm and overlapping segment scoring based decision rule. In: *2025 International Conference on Communication, Computing, Networking, and Control in Cyber-Physical Systems (CCNCPS)*. [S.l.: s.n.], 2025. p. 169–174. Citado 2 vezes nas páginas 22 e 24.
- SHAHBAZIAN, R.; TRUBITSYNA, I. Human sensing by using radio frequency signals: A survey on occupancy and activity detection. *IEEE Access*, v. 11, p. 40878–40904, 2023. Citado 3 vezes nas páginas 14, 18 e 20.
- WANG, Q.; WU, Z. Wifi channel state information-based motion detection across time-domain. In: *2024 7th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI)*. [S.l.: s.n.], 2024. p. 1–5. Citado 6 vezes nas páginas 14, 16, 21, 22, 23 e 24.