Universidade de Brasília

Faculdade de Tecnologia

Departamento de Engenharia Elétrica

CNN Few-Shot para Aprendizado Flexível em Ambientes Prediais

Andressa Maria Monteiro Sena

PROJETO FINAL DE CURSO

ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Brasília

2025

Universidade de Brasília

Faculdade de Tecnologia
Departamento de Engenharia Elétrica

CNN Few-Shot para Aprendizado Flexível em Ambientes Prediais

Andressa Maria Monteiro Sena

Projeto Final de Curso submetido como requisito parcial para obtenção do grau de Engenheiro de Controle e Automação.

Orientador: Prof. Dr. Adolfo Bauchspiess

Brasília
2025

FICHA CATALOGRÁFICA

Sena, Andressa Maria Monteiro.
S474c CNN Few-Shot para Aprendizado Flexível em Ambientes Prediais / Andressa Maria Monteiro Sena; orientador Adolfo Bauchspiess. -- Brasília, 2025.
55 p.

Projeto Final de Curso (Engenharia de Controle e Automação) -- Universidade de Brasília, 2025.

1. Automação Predial. 2. Reconhecimento de Atividades Humanas. 3. Aprendizagem de Máquina. 4. Aprendizagem Profunda. I. Bauchspiess, Adolfo, orient. II. Título.

Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Elétrica

CNN Few-Shot para Aprendizado Flexível em Ambientes Prediais

Andressa Maria Monteiro Sena

Projeto Final de Curso submetido como requisito parcial para obtenção do grau de Engenheiro de Controle e Automação.

Trabalho aprovado. Brasília, 21 de fevereiro de 2025:

Prof. Dr. Adolfo Bauchspiess,
UnB/FT/ENE
Orientador

Prof. Dr. Geovany Araújo Borges,
UnB/FT/ENE
Professor Associado

Prof. Dr. Alexandre Ricardo Soares Romariz,
UnB/FT/ENE
Professor Associado

Este trabalho é dedicado ao meu tio Leandro Molina,
que sempre estará no meu coração.

Agradecimentos

Primeiramente, gostaria de agradecer ao Thor e à Athena por terem ressignificado minha vida no momento mais conturbado da humanidade (COVID 19).

Agradeço também à toda minha família por me incentivar no caminho da independência a partir dos estudos desde cedo. Obrigada à minha mãe Sandra que sempre será minha maior referência acadêmica e ao meu pai Sena por ter transferido o seu fascínio pela tecnologia para mim. Aos meus tios Leandro Molina (em memória) e Carolina Genaro por sempre me acolherem. À minha tia Isabel Monteiro e sua família pelos momentos de descontração. E ao meu irmão Álvaro por ser o maior e melhor parceiro de todos.

Gratidão especial a todos os meus amigos que me ajudaram nas situações mais complicadas possíveis e me reergueram pra que eu pudesse continuar minha caminhada. Essa conquista é fruto do suporte de vocês.

Sou imensamente grata aos professores que cruzaram meu caminho e que, de alguma forma, me inspiraram e incentivaram a seguir em frente. Cada aula, conselho e desafio proposto contribuíram para a construção da minha personalidade e trajetória acadêmica.

E por fim, agradeço ao professor Adolfo por ser a paciência e o carisma em pessoa e por ter sido o melhor orientador que eu poderia ter tido.

"Ninguém nasce mulher: torna-se mulher."
(Simone de Beauvoir)

Resumo

A responsabilidade pelo aumento do número de cidades sustentáveis é global. A Automação Predial pode contribuir nesse cenário incentivando o uso eficiente de energia por meio de ferramentas tecnológicas. Definir qual atividade uma pessoa executa em determinado ambientes permite que decisões acerca de termorregulação e controle de iluminação sejam tomadas. Para o reconhecimento de atividades humanas (HAR), implementa-se uma rede CSI de sensores Wi-Fi com microcontroladores ESP32 para criação de banco de dados de treinamento para redes neurais artificiais. Com intuito de comparação, avalia-se o desempenho de dois métodos distintos de aprendizado profundo de máquina: rede neural convolucional (aprendizado em um ambiente e uso no mesmo) e *Few-Shot Learning* (aprendizado em um ambiente e uso em outros ambientes similares - *transfer learning*). Códigos disponíveis em https://github.com/dekomonte/csi-har-few_shot-unb.git.

Palavras-chave: Automação Predial. Reconhecimento de Atividades Humanas. Aprendizagem de Máquina. Aprendizagem Profunda.

Abstract

Responsibility for increasing the number of sustainable cities is global. Building automation can contribute to this scenario by encouraging the efficient use of energy through technological tools. Defining which activity a person performs in a given environment allows decisions to be made about thermoregulation and lighting control. For human activity recognition (HAR), a CSI network of Wi-Fi sensors is implemented with ESP32 microcontrollers to create a training database for artificial neural networks. For comparison purposes, the performance of two different deep machine learning methods is evaluated: convolutional neural networks (learning in one environment and using it in the same environment) and few-shot learning (learning in one environment and using it in other similar environments - transfer learning). Codes are available at https://github.com/dekomonte/csi-har-few_shot-unb.git.

Keywords: Building Automation. Human Activit Recognition. Machine Learning. Deep Learning.

Lista de Figuras

Figura 2.1 17 Objetivos de Desenvolvimento Sustentável promovidos pela ONU. O 11º se refere a Cidades e Comunidades Sustentáveis. 17
Figura 2.2 Atividades humanas típicas em escritórios/laboratórios estimadas neste trabalho: "em pé", "sentado", "andando"e "vazio". 18
Figura 2.3 Representação de uma estrutura MIMO. Múltiplos transmissores enviam o mesmo sinal para múltiplos receptores. 19
Figura 2.4 Sinal OFDM no domínio de frequências. Representação esquemática das subportadoras. 20
Figura 2.5 Sinal CSI representado em forma de imagem. Parte real e parte imaginárias divididas em duas subfiguras. 20
Figura 2.6 Tensor CSI 4D, expande a matriz de CSI para uma quarta dimensão temporal. 21
Figura 2.7 Representação de uma Rede Neural Convolucional arbitrária. Sua estrutura multicamadas é capaz de compartilhar informações. 22
Figura 2.8 Representação de convolução 2D. Versão gráfica do somatório de convolução. 23
Figura 2.9 Representação simplificada de uma rede prototípicas. As características das amostras são separadas em um espaço vetorial. 24
Figura 2.10 Representação de uma rede Few-Shot. Mostrando a etapa de treinamento com CNN, a entrada dividida em conjunto de suporte e de consulta e a fase linear com a função de similaridade. 24
Figura 2.11 Modelo de Zonas de Fresnel. 25
Figura 3.1 Foto do ESP32-DevKitC. 26
Figura 3.2 Desenvolvimento de aplicações para Esp32. 27
Figura 3.3 Antena Externa Articulada. 28
Figura 4.1 Visão abstrata das salas do LARA que foram utilizadas no experimento. 30
Figura 4.2 Primeira topologia de distribuição dos módulos Esp32 nas duas salas. . 31
Figura 4.3 Links entre transmissores e receptores da primeira topologia da rede de sensores - Zonas de Fresnel. 1ª zona em linha grossa. 31
Figura 4.4 Segunda topologia de distribuição dos módulos Esp32 nas duas salas. . 32
Figura 4.5 Links entre transmissores e receptores, rede com 2 transmissores e 6 receptores - Zonas de Fresnel. 32
Figura 4.6 Fluxo Temporal da Aquisição de Sinal CSI241. Neste projeto $NS = 4$, $CYCLE = 400m$ e $\Delta = 90ms$. 34
Figura 4.7 Distribuição módulos Esp32 - Sala de Reunião. 35

Figura 4.8 Representação da rede prototípica de (Bahadori; Ashdown; Restuccia, 2022). 37
Figura 4.9 Visão geral da rede Few-Shot adaptada ao trabalho na UnB. 39
Figura 5.1 Atalho do terminal CMD. 40
Figura 5.2 Terminal CMD com as configurações de ambiente sendo carregadas. . . 40
Figura 5.3 Gravação dos Módulos Esp32 41
Figura 5.4 Opção Wi-Fi CSI(Channel State Information) habilitada no menu de configuração. 42
Figura 5.5 Sala de Estudos do LARA com indicação do posicionamento dos módulos. 43
Figura 5.6 Sala de Reunião do LARA com indicação do posicionamento dos módulos. 43
Figura 5.7 Parte de um arquivo gerado para uma rotina na sala de estudos antes de passar pelo algoritmo de inserções. 44
Figura 5.8 Parte de um arquivo gerado para uma rotina na sala de reuniões antes de passar pelo algoritmo de inserções. 44
Figura 5.9 Sinal CSI de um rotina completa de aquisição de dados em forma de imagem. 46
Figura 5.10 Sinal CSI representado em forma de imagem. 47
Figura 5.11 Subportadoras CSI consecutivas, sensor 6. 48
Figura 5.12 Matriz de Confusão da CNN. 305 imagens de Treinamento e 132 de Testes (30%), 19 experimentos de 4 min. 49
Figura 5.13 Superfícies 3D com valores de amplitude e fase de uma imagem. A subfigura da esquerda mostra valores condizentes, a da direita representa um outlier. 51
Figura 5.14 Demonstrativo de quantas linhas dos arquivos estão sendo utilizada ao longo de duas janelas de 10s. A linha verde representa o tempo e as marcações em azul e laranja indicam o quanto do arquivo foi utilizado na geração das imagens. 51

Lista de Tabelas

Tabela 3.1 Características principais do ESP32 destacadas pela Espressif. 26
Tabela 4.1 Comparação entre topologia de redes de sensores utilizadas em estudos. . 33
Tabela 4.2 Valores de Δ e T_{Folga} no ciclo de transmissão. 35
Tabela 4.3 Distribuição de atividades da rotina fixa no tempo, 4 minutos/experimento. 35
Tabela 4.4 Redução de dimensionalidade no Encoder. 38
Tabela 5.1 Quantidade de inserções de linhas no pré-processamento (Dados Sala de Estudos). 45
Tabela 5.2 Quantidade de inserções de linhas no pré-processamento(Dados Sala de Reunião). 45
Tabela 5.3 Tabela de Acurácia das Configurações 50

Lista de Abreviaturas e Siglas

ABNT	Associação Brasileira de Normas Técnicas
ACP	Análise de Componentes Principais
CNN	Convolutional Neural Network
CSI	Channel State Information
IA	Inteligência Artificial
IoT	Internet of Things
LARA	Laboratório de Automação e Robótica
LGPD	Lei Geral de Proteção de Dados
MIMO	Multiple Input Multiple Output
OFDM	Orthogonal Frequency Division Multiplexing
PCA	Principal Component Analysis
RNA	Redes Neurais Artificiais
UnB	Universidade de Brasília

Sumário

1 Introdução 15
1.1 Motivação 15
1.2 Objetivos 16
1.2.1 Objetivos Gerais 16
1.2.2 Objetivos Específicos 16
1.3 Organização do Manuscrito 16

2 Revisão Bibliográfica 17
2.1 Eficiência Energética 17
2.2 Human Activity Recognition 18
2.3 Sistemas MIMO e OFDM 19
2.4 Wi-Fi Sensing e Channel State Information (CSI) 19
2.5 Redes Neurais Artificiais 21
2.5.1 Redes Neurais Convolucionais 21
2.5.2 Few-Shot Learning 23
2.6 Zonas de Fresnel 25

3 Materiais e Ferramentas 26
3.1 ESP32 26
3.1.1 ESP-IDF 27
3.1.2 ESP32 ESP-NOW 27
3.1.3 ESP32 ESP-NOW OFDM-MIMO 28
3.1.4 ESP32 ESP-NOW CSI 28
3.1.5 Antenas 28
3.2 Python 29

4 Métodos 30
4.1 Aquisição de Dados - Banco da Sala de Estudos 30
4.1.1 Topologia da Rede de Sensores 30
4.1.2 Rotina de Aquisição 33
4.1.3 Atividades da Rotina 35
4.2 Aquisição de Dados - Banco da Sala de Reunião 35
4.3 Pré-processamento de Dados 36
4.3.1 Correção da Sequência de Sensores 36
4.3.2 Filtro Passa-baixas 36
4.3.3 Sanitização 36
4.4 Few-Shot Learning 37

4.4.1 ProtoNet . . . 37
4.4.2 Rede Few-Shot Adaptada . . . 37
4.5 Desempenho das Redes Neurais . . . 38

5 Desenvolvimento . . . 40
5.1 Instalação e Configuração de Ambiente . . . 40
5.2 Gravação Módulos . . . 41
5.2.1 Gravação M1/M11 . . . 41
5.2.2 Gravação M2/M12 e S . . . 42
5.3 Procedimento Experimental . . . 42
5.4 Resultados . . . 44
5.4.1 Aquisição de Dados e Pré-processamento . . . 44
5.4.2 CSI . . . 46
5.4.3 CNN . . . 48
5.4.4 Few-Shot . . . 48
5.5 Discussão . . . 49
5.5.1 Interferências Eletromagnéticas . . . 49
5.5.2 Esp32 . . . 49
5.5.3 Qualidade dos Dados . . . 50
5.5.4 Uso Parcial dos Dados . . . 51
5.5.5 Desempenho das Redes Neurais . . . 52

6 Conclusões . . . 53
6.1 Projetos Futuros . . . 53

Referências . . . 54

1 Introdução

1.1 Motivação

Uma vantagem para a Automação Predial é ter uma forma confiável de se determinar atividades de pessoas (*Human Activity Recognition* - HAR) em ambientes. Tal aspecto contribui diretamente nas tomadas de decisões assertivas relacionadas à eficiência energética como, por exemplo, o controle de iluminação e/ou termorregulação. Aliado a isso, tecnologias de Inteligência Artificial (IA) e Internet das Coisas (IoT) vêm ganhando cada vez mais espaço nas implementações de sistemas inteligentes devido às suas facilidades de integração.

Estudos recentes, realizados na Universidade de Brasília (UnB), obtiveram resultados relevantes, aproximadamente 25% de "economia"de energia, estimando ocupação/identificação de pessoas e de atividades por meio de câmeras e controle antecipativo (Rodrigues; Silva, 2018). Porém, esse método vai de encontro às jurisdições de privacidade, como a Lei Geral de Proteção de Dados Pessoais - LGPD (BRASIL, 2018), quando expandido para contextos diferentes. A utilização das formas de onda do sinal Wi-Fi como alternativa passiva de monitoramento se mostra, nesse sentido, uma proposta de bastante potencial visto que as redes sem fio estão presentes na maioria dos lugares e não comprometem a privacidade das pessoas (Bauchspiess, 2023).

A estimação de atividades em ambientes prediais via CSI (*Channel State Information*) com aprendizado de máquina, utilizando grandes conjuntos de dados rotulados, tem mostrado resultados positivos (Monteiro (2024) obteve 48,33% de acurácia com dados de treinamento e validação de um mesmo ambiente). No entanto, suas aplicações são limitadas aos locais de sua coleta e exigem uma vasta quantidade de dados para o treinamento. Em contrapartida, técnicas de *Few-Shot Learning* são investigadas como alternativa, possibilitando o uso de uma quantidade reduzida de dados rotulados para determinar os parâmetros em ambientes diferentes (*Cross Domain*).

Dentre os principais benefícios potenciais do uso dessa tecnologia, destacam-se citar: a promoção do conforto térmico, o aumento da produtividade em ambientes comerciais e residenciais, e o monitoramento de pessoas em situação de vulnerabilidade, como idosos, pessoas com deficiência e crianças.

Espera-se, como consequência desse estudo, uma redução teórica de $\approx 20\%$ do consumo de energia do laboratório em comparação com atual tecnologia implementada no ambiente.

1.2 Objetivos

1.2.1 Objetivos Gerais

O objetivo geral deste trabalho é, investigar métodos para estimar atividades humanas - HAR - em novos ambientes prediais empregando aprendizado de máquina com poucas amostras (*Few-Shot Learning*) no contexto de eficiência energética em automação predial.

1.2.2 Objetivos Específicos

Os objetivos específicos do projeto incluem os itens a seguir.

- Construir um banco de dados para treinamento de aprendizado de máquina de atividades típicas em escritórios/laboratórios: e.g., “Vazio”, ”Caminhar”, ”Em pé”, “Sentar” (*Empty*, *Walk*, *Stand*, *Sit*). Essas rotinas são rotuladas pelo tempo de execução das atividades;
- Automatizar a Sala de Reunião com esse banco de dados por meio de sistema CNN Few-Shot;
- Validar o sistema CNN-Few-Shot em outros ambientes;

As salas do Laboratório de Automação e Robótica (LARA) foram escolhidas como cenário de treinamento e validação dos experimentos.

1.3 Organização do Manuscrito

Este projeto encontra-se distribuído ao longo de 6 capítulos resumidos a seguir.

O Capítulo 1 apresenta a contextualização e os objetivos da proposta. Nos Capítulos 2 e 3 são abordadas a fundamentação teórica de referência do trabalho e a descrição dos materiais usados durante seus procedimentos, respectivamente. Em seguida, o Capítulo 4 descreve os métodos empregados no planejamento das atividades práticas e na sua execução. O capítulo 5 descreve os experimentos e testes realizados, incluindo suas descrições e resultados obtidos. E, por fim, o Capítulo 6 expõe conclusões acerca do projeto.

2 Revisão Bibliográfica

2.1 Eficiência Energética

Estima-se que, até 2050, mais da metade da população mundial viverá em áreas urbanas; a falta de planejamento durante o processo de urbanização em muitas cidades resultou em um grande volume de habitações precárias e aumento da poluição (Economic; Affairs, 2023). Segundo a Organização das Nações Unidas (ONU), as cidades estão diretamente relacionadas às mudanças climáticas, pois consomem quase 80% da energia mundial produzida e são responsáveis por mais de 60% dos gases de efeito estufa emitidos na atmosfera (UN-Habitat, 2024).

Figura 2.1 – 17 Objetivos de Desenvolvimento Sustentável promovidos pela ONU. O 11º se refere a Cidades e Comunidades Sustentáveis.

Fonte: https://sdgs.un.org/goals.

Considerando-se esse e outros problemas, a ONU frequentemente propõe planos de ações globais para seus países membros com o intuito de promover sustentabilidade em escala mundial. Em 2015, a Agenda 2030 estabeleceu 17 Objetivos de Desenvolvimento Sustentável (ODS), dos quais o 11º, especificamente, foca em "Tornar as cidades inclusivas,

seguras, resilientes e sustentáveis". Evidenciando-se o grau de importância do assunto.

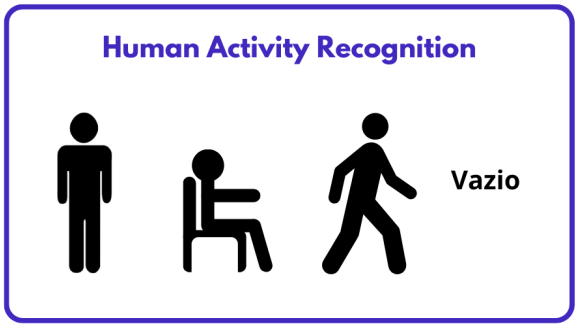
Alinhando-se com essas metas, as **Cidades Inteligentes** aplicam conceitos tecnológicos em áreas urbanas com o intuito de melhorar a gestão de recursos e o fornecimento de serviços à população. Na atual realidade de Indústria 4.0, os sistemas digital e físico estão fortemente conectados e podem se unir com as mais diversas modernidades. Destaca-se, aqui, o uso de IA como forma de melhoria de processos humanos (saúde, transporte, moradia, educação etc) a partir de Análise de Dados. Em paralelo, os serviços de Internet das Coisas (em inglês *Internet of Things*, IoT) trazem como vantagem o acesso a um volume gigantesco de informações e dados.

Ainda pensando-se em sustentabilidade, há duas abordagens importantes para se promover **Eficiência Energética**: reduzir o consumo de energia, ou seja, utilizá-la de forma mais eficaz, e ampliar o uso de energias renováveis. Em **Automação Predial** - área que estuda as tecnologias relacionadas à automação de edifícios comerciais como hospitais, supermercados, hotéis, lojas, shoppings, escolas, entre outros -, o controle de iluminação e a termorregulação feitos de forma inteligente permitem economia de energia maior que 15% (Prudente, 2011).

2.2 Human Activity Recognition

Estudos sobre HAR avaliam formas de se determinar atividades (figura 2.2) exercidas por pessoas em determinado ambiente por meio de sensores. Neste trabalho, avalia-se utilizar dados de CSI aplicados em HAR. Com base nas informações geradas, decisões são tomadas para incentivar eficiência energética.

Figura 2.2 – Atividades humanas típicas em escritórios/laboratórios estimadas neste trabalho: "em pé", "sentado", "andando"e "vazio".

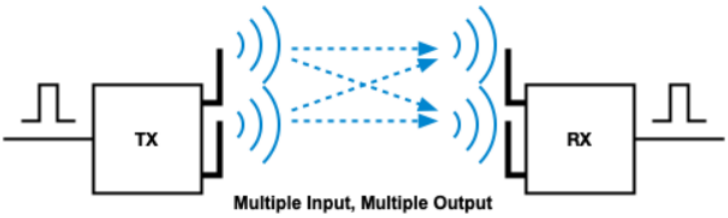


Segundo Gallo e Ribeiro (2007), conforto térmico está relacionado com princípios quantificáveis (temperatura, umidade etc) e não quantificáveis (saúde, hábitos etc). Além disso, as atividades metabólicas exercidas pelas pessoas influenciam diretamente: atividades físicas vigorosas tendem a gerar conforto em temperaturas mais baixas, por exemplo.

2.3 Sistemas MIMO e OFDM

Sistemas de múltiplas entradas e múltiplas saídas (*Multiple-Input and Multiple-Output*, *MIMO*) utilizam múltiplas antenas tanto nos transmissores quanto nos receptores. Essa tecnologia oferece uma solução eficaz para aumentar a qualidade em sistemas modernos de comunicação, melhorando sua confiabilidade, ampliando a cobertura e otimizando o consumo de energia (Rawat; Kaushik; Tiwari, 2021).

Figura 2.3 – Representação de uma estrutura MIMO. Múltiplos transmissores enviam o mesmo sinal para múltiplos receptores.



Fonte: https://www.elektronik-kompendium.de/sites/net/1004251.htm.

A Multiplexação de Sinal por Divisão Ortogonal na Frequência (*Orthogonal Frequency Division Multiplexing, OFDM*) é uma técnica de modulação de dados, utilizada em sistemas de comunicação sem fio como a IEEE 802.11 (Wi-Fi). Ao contrário de sistemas convencionais, no sistema OFDM a transmissão de dados é feita paralelamente em diversas portadoras ortogonais (subportadoras) com frequências diferentes. As subportadoras são divididas de forma que a interferência entre espectros seja evitada. Os vários sinais compartilham a banda de um mesmo canal (Lathi, 2006).

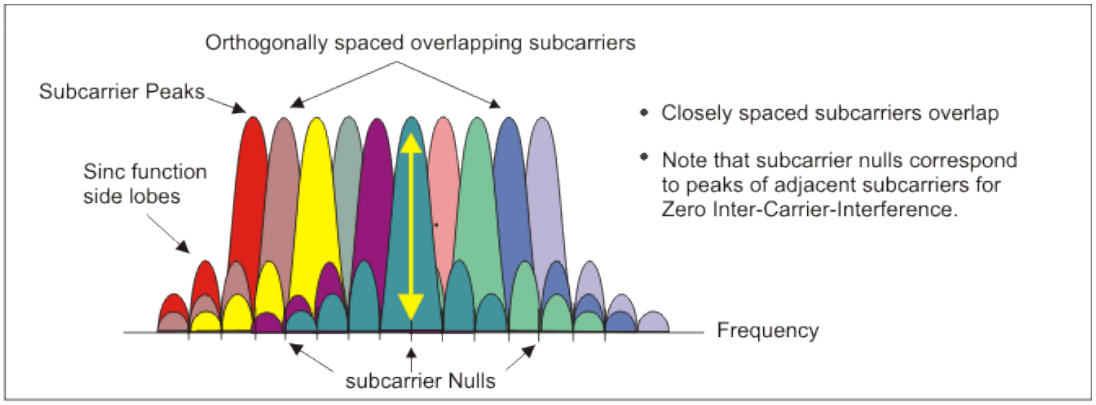
A combinação das arquiteturas MIMO e OFDM destaca-se como uma solução eficaz para aumentar a proteção e confiabilidade dos dados transmitidos, combinando as vantagens de ambas as técnicas.

2.4 Wi-Fi Sensing e Channel State Information (CSI)

Wi-Fi Sensing é uma técnica que consiste em monitorar as mudanças que a disposição e a movimentação de objetos causam nos parâmetros de verificação de transmissão do sinal Wi-Fi entre transmissores e receptores. CSI é o principal indicador dessa variação, seus valores de amplitude e fase permitem a obtenção de funções de transferência (Bahadori; Ashdown; Restuccia, 2022) (Ma; Zhou; Wang, 2019).

Sinais de Wi-Fi IEEE 802.11 utilizam, atualmente, tecnologia OFDM, que transmite os pacotes de dados por várias subportadoras ortogonais (tipicamente entre 52 e 1024). O CSI, parâmetro presente no preâmbulo do frame do sinal Wi-Fi, é registrado em cada pacote de dados recebido em uma subportadora, indicando a qualidade do canal para sua frequência

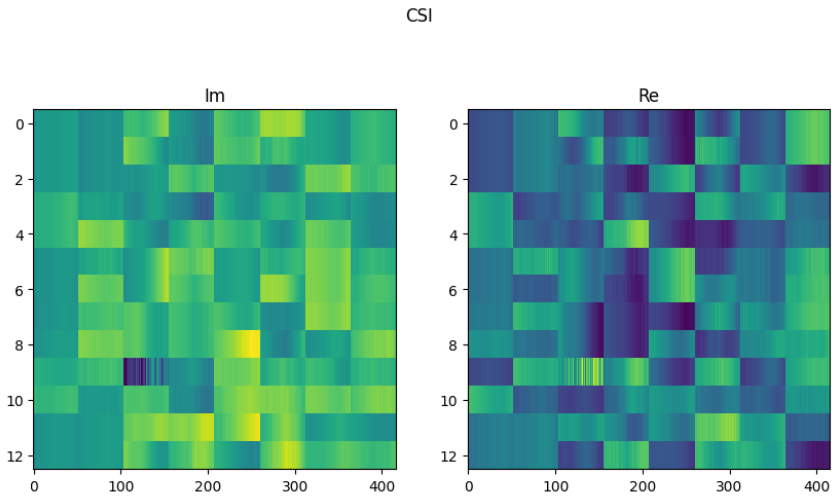
Figura 2.4 – Sinal OFDM no domínio de frequências. Representação esquemática das subportadoras.



Fonte: https://helpfiles.keysight.com/csg/89600B/Webhelp/Subsystems/wlan-ofdm/content/ofdm_basicprinciplesoverview.htm.

específica. Devido às condições de propagação particulares entre transmissor e receptor, a sequência de sinais CSI ao longo do tempo forma uma imagem/assinatura da atividade desenvolvida (Xiao *et al.*, 2016).

Figura 2.5 – Sinal CSI representado em forma de imagem. Parte real e parte imaginárias divididas em duas subfiguras.



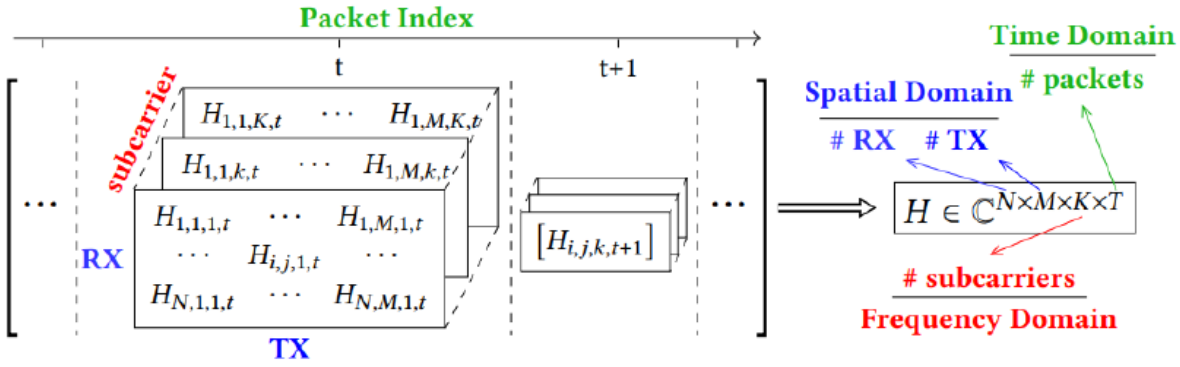
Fonte: (Bauchspiess, 2023)

Nesse trabalho, têm-se 2 transmissores (T_x) e 4 receptores (R_x), o número de links é de $n = T_x \times R_x = 8$. Cada portadora apresenta m subportadoras, para o caso em estudo, $m = 52$. A partir disso, a matriz de CSI de um pacote é descrita por

$$H = \begin{bmatrix} h_{1,1} & \cdots & h_{1,m} \\ \vdots & \vdots & \vdots \\ h_{n,1} & \cdots & h_{8,52} \end{bmatrix}, \tag{2.1}$$

em que cada elemento $h_{i,j} = |h_{i,j}|e^{j\theta}$ representa amplitude ($|h_{i,j}|$) e fase (θ) de cada subportadora do pacote (Zhang *et al.*, 2023).

Figura 2.6 – Tensor CSI 4D, expande a matriz de CSI para uma quarta dimensão temporal.



Fonte: (Zhuravchak, 2020)

A figura 2.6 ilustra o sistema CSI em 4 dimensões, sendo o tempo a quarta dimensão, ampliando a estrutura apresentada na equação (2.1). Essa representação oferece uma visão mais detalhada do processo de transmissão, demonstrando a evolução dos pacotes ao longo do tempo.

2.5 Redes Neurais Artificiais

Redes Neurais Artificiais (RNA) são modelos computacionais inspirados no cérebro humano. Sua estrutura baseia-se no comportamento dos neurônios biológicos, que tomam decisões processando sinais elétricos. Hodiernamente, as aplicações envolvendo RNAs são inúmeras, principalmente nos casos de processamento de sistemas não-lineares (Silva; Spatti; Flauzino, 2016).

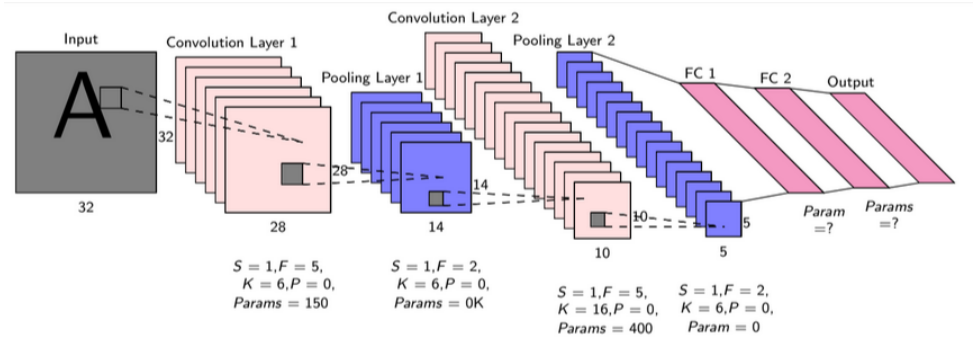
2.5.1 Redes Neurais Convolucionais

Redes Neurais Convolucionais (*Convolutional Neural Network, CNN*) são um tipo de redes neurais de Aprendizado Profundo, subconjunto de Aprendizado de Máquina. Ao contrário dos modelos tradicionais de aprendizado de máquina, as CNNs aprendem e compartilham os pesos do classificador e os filtros/kernels que devem ser utilizados.

Esse tipo específico de redes neurais é bastante utilizado no processamento de imagens. Como a coleta da frequência das subportadoras em determinado tempo forma uma imagem, esse tipo de algoritmo se mostrou adequado ao projeto.

Uma de suas características mais significativa está no design multicamadas interconectadas capazes de melhorar, a cada etapa, a categorização/previsão por meio de diversos tipos de operações e compartilhamento de informações. O procedimento mais significativo dessa categoria de algoritmos é a convolução, a qual será simplificadamente descrita a seguir. A figura 2.7 mostra uma representação geral desde a camada de entrada até a saída de um CNN genérica.

Figura 2.7 – Representação de uma Rede Neural Convolucional arbitrária. Sua estrutura multicamadas é capaz de compartilhar informações.



Fonte: (Khapra, 2017).

2.5.1.1 Convolução

Convolução é um tipo de operação linear sobre duas funções que resulta em uma terceira. Ela está presente em diversas aplicações de Matemática e Engenharia. No caso unidimensional, a integral de convolução é definida pela equação (2.2) (Lathi, 2006):

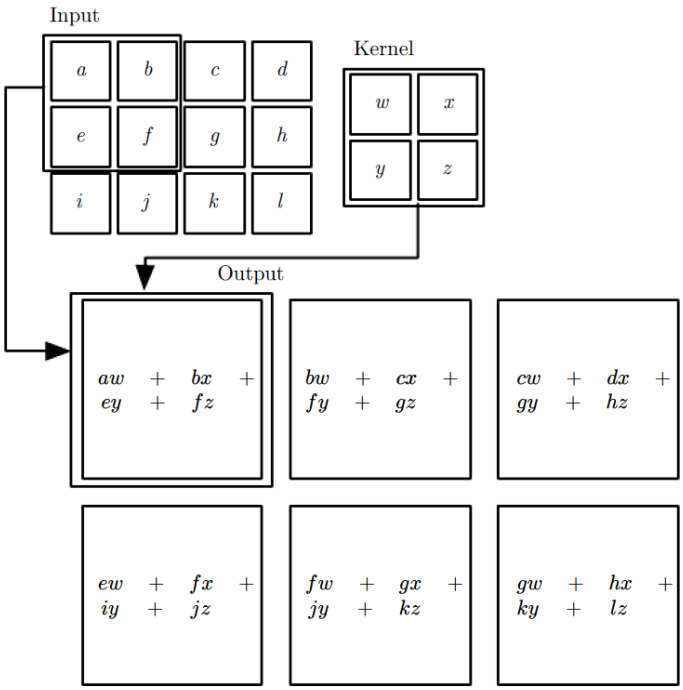
$$s(t) = (x_1 * x_2) \equiv \int_{-\infty}^{+\infty} x_1(\tau)x_2(t-\tau)d\tau. \tag{2.2}$$

No aprendizado de máquina, geralmente, o primeiro argumento (entrada/input) é uma matriz com dados e o segundo (kernel) é uma matriz de parâmetros. Segundo (Goodfellow; Bengio; Courville, 2016), é possível adequar e modelar a operação pelo somatório discreto:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m,j-n). \tag{2.3}$$

A figura 2.8 representa de forma esquemática o funcionamento da convolução no caso 2D , ou seja, a aplicação da equação (2.3) num contexto bidimensional. Demonstrando o que ocorre no processamento de imagens nas camadas convolucionais.

Figura 2.8 – Representação de convolução 2D. Versão gráfica do somatório de convolução.



Fonte: (Goodfellow; Bengio; Courville, 2016).

2.5.2 Few-Shot Learning

É um caso particular aprendizado de máquina, em que a quantidade de dados rotulados é escassa, principalmente em casos onde não é viável nem prático obter-se um banco de dados volumoso. Um modelo pré-treinado generaliza sobre novas categorias de dados a partir de alguns exemplos rotulados por classe.

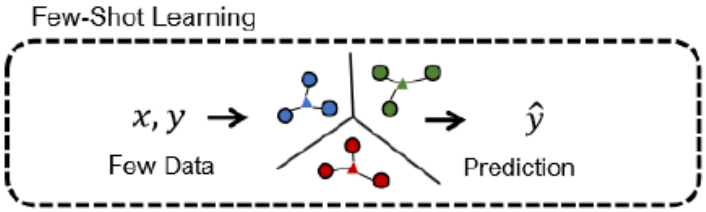
Suas implementações, geralmente, envolvem aprendizagem contrastiva e aprendizagem prototípica (em inglês *contrastive learning* e *prototypical learning*) (Yang *et al.*, 2022).

A aprendizagem contrastiva avalia dados não rotulados de um banco de dados, compara as amostras e aprende atributos em comum entre elas. As imagens são *contrastadas* umas com as outras para o treinamento entender quais são semelhantes e quais são diferentes.

Redes prototípicas (Snell; Swersky; Zemel, 2017) treinam para separar características de amostras em um espaço vetorial, criando protótipos representativos de cada classe. Durante o treinamento, as amostras são classificadas com base na proximidade desses protótipos a cada iteração. A figura 2.9 mostra uma representação simplificada desse paradigma.

A etapa de pré-treinamento das redes Few-Shot é feita por uma CNN tradicional utilizando um extenso banco de dados, com intuito de aprender representações discriminativas das amostras. Posteriormente, na fase de classificação, a entrada da parte prototípica (linear) é dividida em dois conjuntos: o *support set*, que contém exemplos rotulados usados

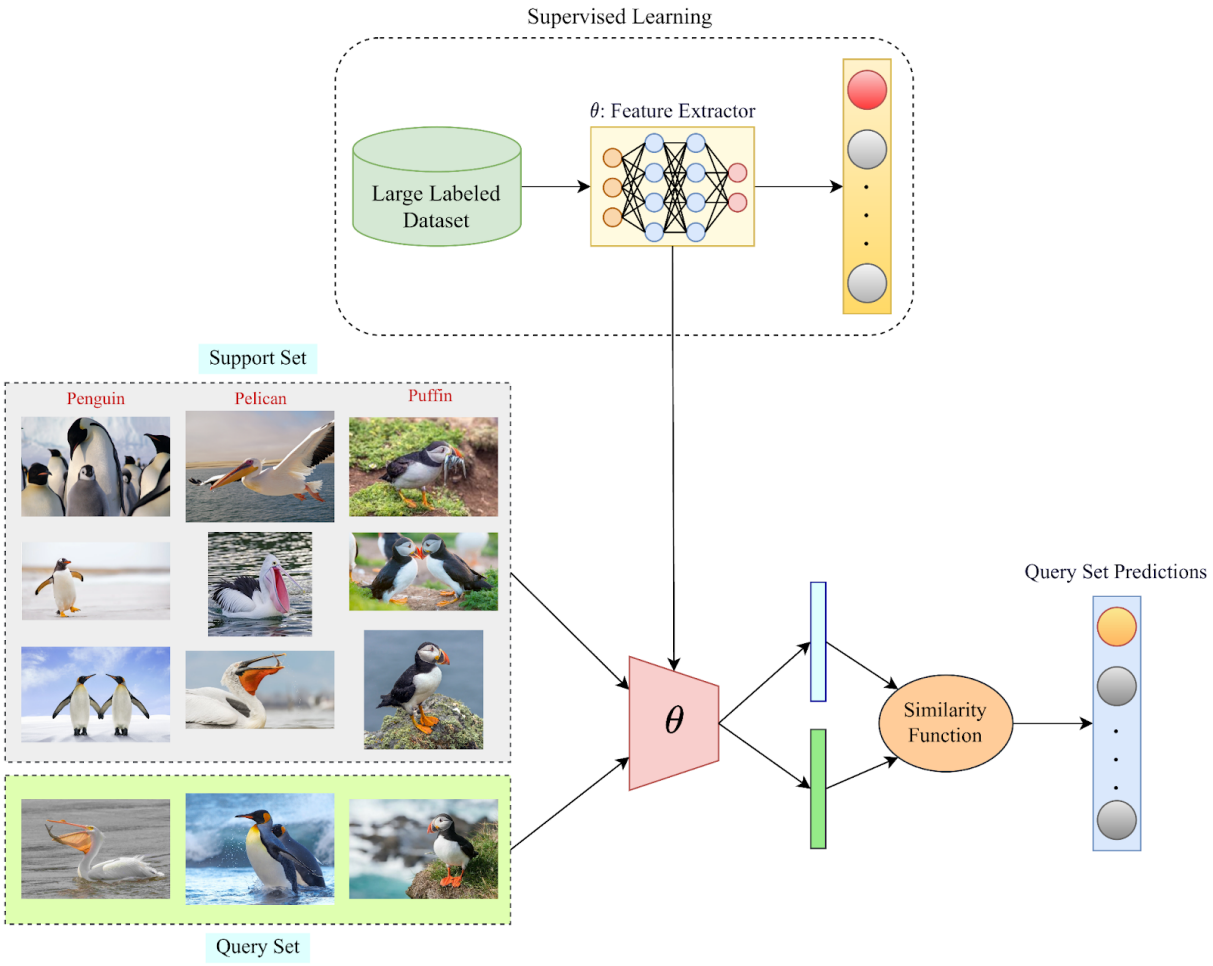
Figura 2.9 – Representação simplificada de uma rede prototípicas. As características das amostras são separadas em um espaço vetorial.



Fonte: Adaptado de (Yang *et al.*, 2022).

para calcular os protótipos das classes, e o *query set*, que avalia a capacidade do sistema em classificar novas amostras. A classificação ocorre comparando a similaridade das amostras do query set com os protótipos. A figura 2.10 é uma visão geral do método citado.

Figura 2.10 – Representação de uma rede Few-Shot. Mostrando a etapa de treinamento com CNN, a entrada dividida em conjunto de suporte e de consulta e a fase linear com a função de similaridade.



Fonte: https://www.digitalocean.com/community/tutorials/few-shot-learning.

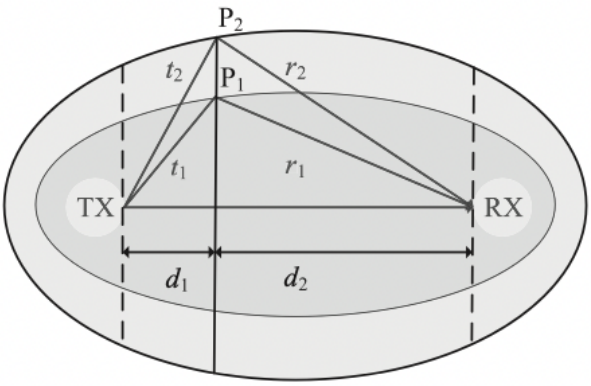
2.6 Zonas de Fresnel

Descreve regiões elípticas no caminho entre transmissores e receptores (figura 2.11). O raio da n-ésima zona de fresnel é dado por (Fei *et al.*, 2020):

$$R_n = \sqrt{\frac{n \cdot d_1 \cdot d_2}{d_1 + d_2} \cdot \frac{\lambda}{2}}, \tag{2.4}$$

em que $d1$ é a distância do transmissor até o ponto de interesse, $d2$ é a distância do receptor até o ponto de interesse, λ é o comprimento de onda da frequência utilizada, e n é o número da zona de Fresnel.

Figura 2.11 – Modelo de Zonas de Fresnel.



Fonte: Adaptado de (Fei *et al.*, 2020).

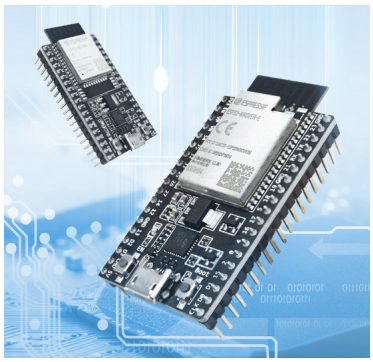
3 Materiais e Ferramentas

O atual capítulo discorre sobre equipamentos utilizados no planejamento e na execução do projeto.

3.1 ESP32

ESP32 é a plataforma de desenvolvimento da empresa chinesa Espressif System que apresenta um microcontrolador robusto de dois núcleos projetado para aplicações IoT. A placa possui suporte integrado a uma variedade de protocolos de comunicação sem fio, incluindo Wi-Fi e Bluetooth.

Figura 3.1 – Foto do ESP32-DevKitC.



Fonte: https://www.espressif.com/en/products/devkits/esp32-devkitc.

Segundo a proprietária, sua forte integração *"busca atender às demandas contínuas de uso eficiente de energia, design compacto, segurança, alto desempenho e confiabilidade"* (Systems, 2023b). A tabela 3.1 apresenta algumas características técnicas relevantes do produto.

Tabela 3.1 – Características principais do ESP32 destacadas pela Espressif.

Características do ESP32
Wi-Fi (banda de 2.4GHz)
Bluetooth
Coprocessador de ultra baixa potência
Múltiplos periféricos
Dois núcleos de CPU Xtensa® LX6 de 32 bits de alto desempenho

Fonte: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/get-started/.

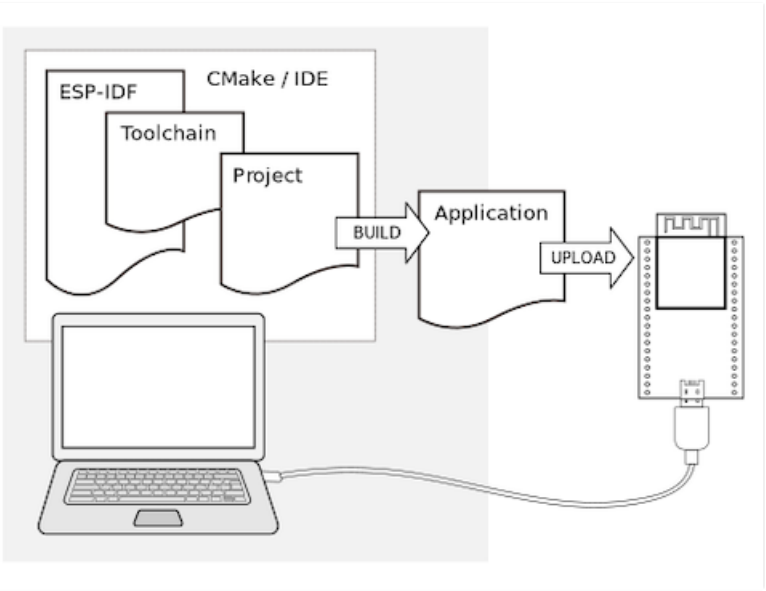
Em adição a isso, o site da corporação menciona que, em questão de hardware extra, para aplicações básicas, é necessário apenas: placa ESP32, cabo USB e computador (Linux, macOS ou Windows).

A principal linguagem de programação usada na plataforma é C/C++. Há várias ferramentas suportadas pelo dispositivo, escolheram-se as opções fornecidas pela própria empresa na preparação do ambiente de desenvolvimento.

Os módulos utilizados no projeto são do tipo **ESP32-WROOM-32UE**, cuja documentação (folha de especificações) pode ser encontrada em https://www.espressif.com/sites/default/files/documentation/esp32-wroom-32e_esp32-wroom-32ue_datasheet_en.pdf. A seguir, faz-se um resumo das ferramentas suportadas pelo componente, as quais foram usadas no trabalho.

3.1.1 ESP-IDF

Figura 3.2 – Desenvolvimento de aplicações para Esp32.



Fonte: https://docs.espressif.com/projects/esp-idf/en/v4.3-beta3/esp32/get-started.

Para a aquisição de dados, foi utilizado o Espressif IoT Development Framework (versão 5.2), via linha de comando, no sistema operacional Windows.

3.1.2 ESP32 ESP-NOW

ESP-NOW é um protocolo de comunicação sem fio da empresa Espressif e foi projetado para aplicações de baixa potência que exijam respostas rápidas e diretas. Por não ser necessário que os módulos sejam registrados em pontos de acesso, a latência na troca de informações é reduzida. Esse protocolo sem fio opera exclusivamente na camada de enlace,

eliminando a necessidade de transmissão de dados por múltiplas camadas, como ocorre no modelo teórico OSI (Systems, 2023a).

O protocolo oferece suporte a uma variedade de recursos avançados, incluindo MIMO e CSI, que podem ser utilizados para aumentar a confiabilidade e eficiência da comunicação sem fio em aplicações de IoT.

3.1.3 ESP32 ESP-NOW OFDM-MIMO

OFDM é uma técnica de transmissão de dados que utiliza múltiplas portadoras ortogonais para enviar o mesmo pacote de dados em diferentes frequências, aumentando a robustez contra interferências e distorções.

MIMO é uma abordagem empregada em sistemas de comunicação sem fio para aprimorar a qualidade do sinal e aumentar as taxas de transmissão de dados, por meio do uso de múltiplas antenas tanto no transmissor quanto no receptor.

3.1.4 ESP32 ESP-NOW CSI

CSI é uma medição das informações do estado do canal, que descreve as condições do canal entre o transmissor e o receptor. O ESP32 suporta o feedback CSI para comunicação ESP-NOW, o que significa que ele pode medir as condições do canal e usar essas informações para ajustar os parâmetros de transmissão para melhor desempenho.

O feedback CSI pode ser usado para estimar a relação sinal-ruído (SNR), a capacidade do canal e a qualidade do canal, entre outros parâmetros. Essas informações podem ser usadas para ajustar os parâmetros de transmissão para melhor desempenho e confiabilidade.

3.1.5 Antenas

Optou-se pela utilização de uma antena externa articulada (figura 3.3) de propagação radial, cujo design e flexibilidade visam minimizar interferências e maximizar o alcance do sinal.

Figura 3.3 – Antena Externa Articulada.

Fonte: Autor desconhecido, imagem encontrada na internet, 2024.

3.2 Python

Todas as aplicações de Machine Learning/Aprendizado de Máquina foram implementadas utilizando-se a linguagem Python, em função de sua simplicidade e vasta gama de bibliotecas especializadas como NumPy, Pandas, Scikit-learn, TensorFlow e Keras. Atualmente, Python é o instrumento mais popular para Análise de Dados, sendo amplamente adotado pela comunidade científica. Além dos motivos já citados, é uma linguagem de código aberto com excelentes documentação e comunidade, facilitando seu uso para as mais diversas aplicações.

4 Métodos

O banco de dados de treinamento foi criado a partir da obtenção de dados das atividades realizadas na sala de estudos do LARA. Em seguida, o treinamento foi executado por meio de uma rede neural convolucional (*CNN Few-Shot*). Por fim, realizou-se uma etapa de validação com dados da sala de reuniões (*Cross-Validation*), aplicando-se *Transfer Learning* de um ambiente para o outro.

4.1 Aquisição de Dados - Banco da Sala de Estudos

O LARA localiza-se no segundo andar do prédio SG-11 Laboratório de Engenharia Elétrica da UnB. As salas que foram utilizadas ficam à esquerda da entrada principal do laboratório. A figura 4.1 mostra uma abstração dos espaços, adjacentes, que foram utilizadas no experimento. Os ambientes são separados por uma parede física.

Figura 4.1 – Visão abstrata das salas do LARA que foram utilizadas no experimento.

A primeira aquisição de dados foi feita na sala de estudos. O espaço do laboratório foi dividido em *zonas* onde os sensores foram distribuídos. Utilizou-se seis conjuntos, definidos como: mestres 1 (M1) e 2 (M2) e quatro sensores (S).

4.1.1 Topologia da Rede de Sensores

A organização e a quantidade de sensores - diversidade espacial(Bahadori; Ashdown; Restuccia, 2022) - foram escolhidas para preencher boa parte dos ambientes e aumentar a confiabilidade dos dados.

Testaram-se duas topologias. A primeira, ilustrada pela figura 4.2, foi descartada em função da região no centro da sala, a qual não é coberta pelos links (figura 4.3).

Figura 4.2 – Primeira topologia de distribuição dos módulos Esp32 nas duas salas.

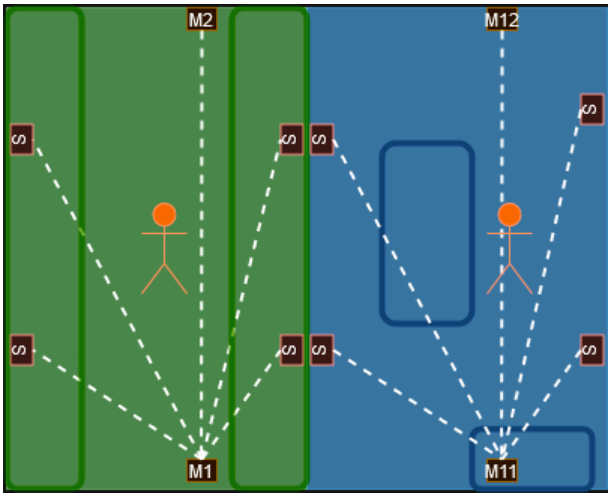
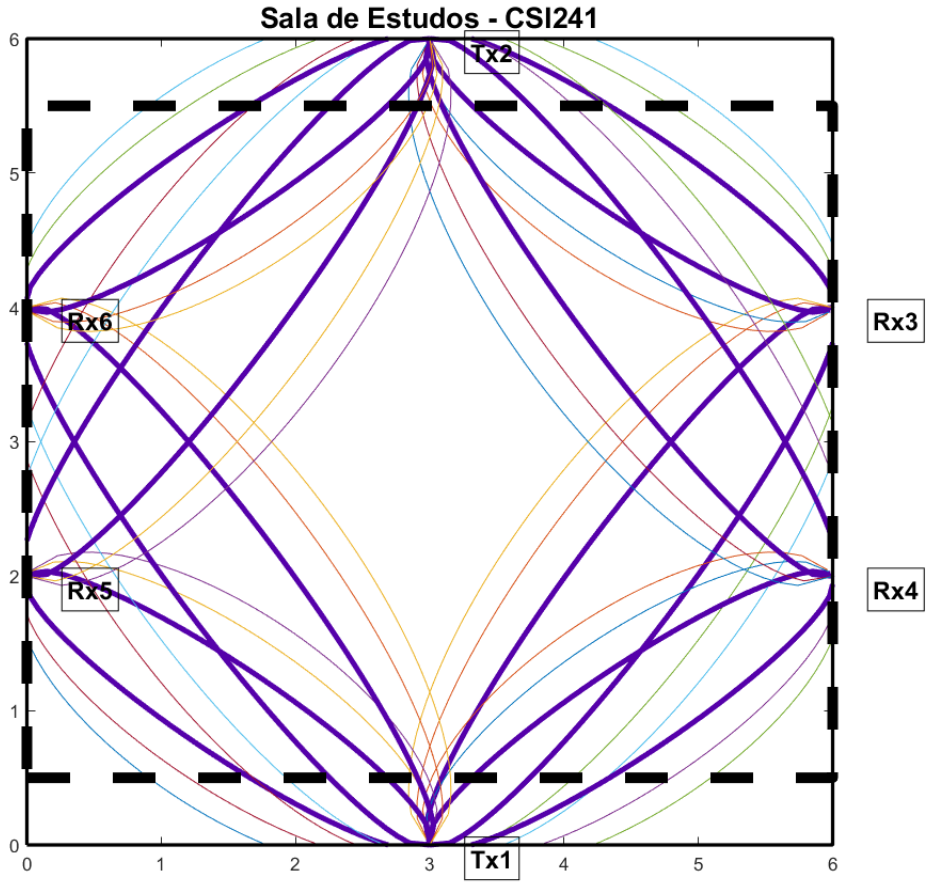


Figura 4.3 – Links entre transmissores e receptores da primeira topologia da rede de sensores - Zonas de Fresnel. 1ª zona em linha grossa.



Fonte: Professor Adolfo Bauchspiess.

A segunda e definitiva topologia desenvolvida é representada na figura 4.4. Compa-

rando as regiões de influência dos módulos, observa-se na figura 4.5 que, mesmo com menos receptores, essa topologia cobre o espaço estudado de forma mais eficiente.

Figura 4.4 – Segunda topologia de distribuição dos módulos Esp32 nas duas salas.

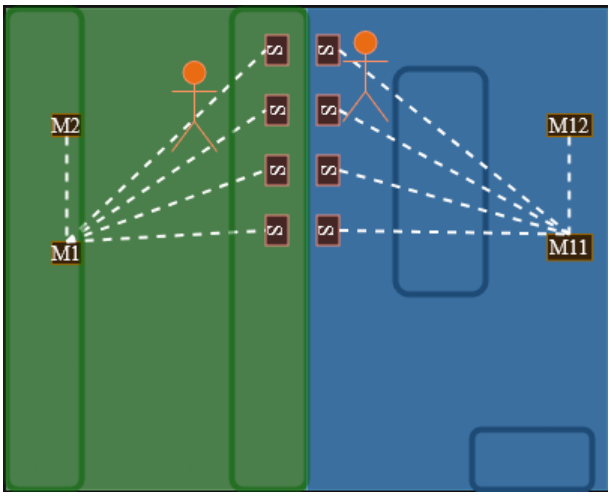
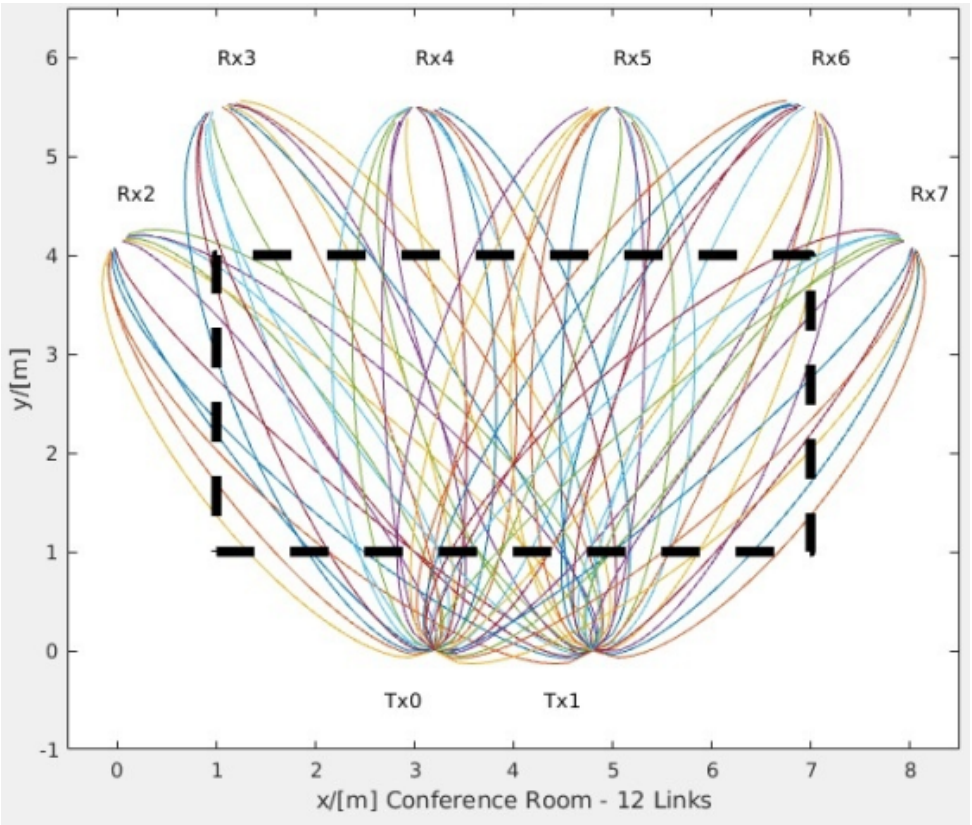


Figura 4.5 – Links entre transmissores e receptores, rede com 2 transmissores e 6 receptores - Zonas de Fresnel.



Fonte: Adaptado de (Bauchspiess, 2023).

A tabela 4.1 apresenta uma comparação entre diferentes topologias e sistemas uti-

lizados em trabalhos que serviram de inspiração para este estudo. O CSI261 (figura 4.5) é proveniente de (Bauchspiess, 2023), com o CSI241 sendo sua adaptação direta utilizada neste trabalho, enquanto o S133 tem origem em (Bahadori; Ashdown; Restuccia, 2022).

Tabela 4.1 – Comparação entre topologia de redes de sensores utilizadas em estudos.

	CSI241	**CSI261**	**S133**
Links	8	12	48
Antenas	1	1	4
Subcarriers	52	52	242
Imagem	26x26 (Amplitude/Fase)	34x34 (Amplitude/Fase)	242x242 (Amplitude)
Bandwidth	40 MHz	40 MHz	80 MHz
Hardware	6xESP32	8xESP32	1xR78006, 3xRT-AC86U

Fonte: Adaptado de (Bauchspiess *et al.*, 2024).

A figura 4.4 mostra a região de influência da topologia escolhida nos dois ambientes. Têm-se 2 transmissores T_x (M1 e M2) e 4 receptores R_x (S6, S7, S8, S9) na esquerda e T_x (M11 e M12) e R_x (S3, S4, S5, S6) na direita. O agente em laranja indica o posicionamento da cadeira e consequentemente das atividades "sentado"e "em pé".

4.1.2 Rotina de Aquisição

A rotina de aquisição de dados começa com o M1 fazendo uma transmissão *broadcast* para todos os módulos, ativando o CSI. Em seguida, o M2 faz outra transmissão *broadcast* após receber o sinal do M1, desativando o CSI. A seguir, cada sensor na sua vez, sequencialmente, devolve para o M1 as informações de CSI de cada sinal recebido. Essa sequência de passos é repetida no ciclo seguinte. O processo completo de cada ciclo está representado na figura 4.6, a qual consegue resumir de forma abstrata os programas de aquisição de dados.

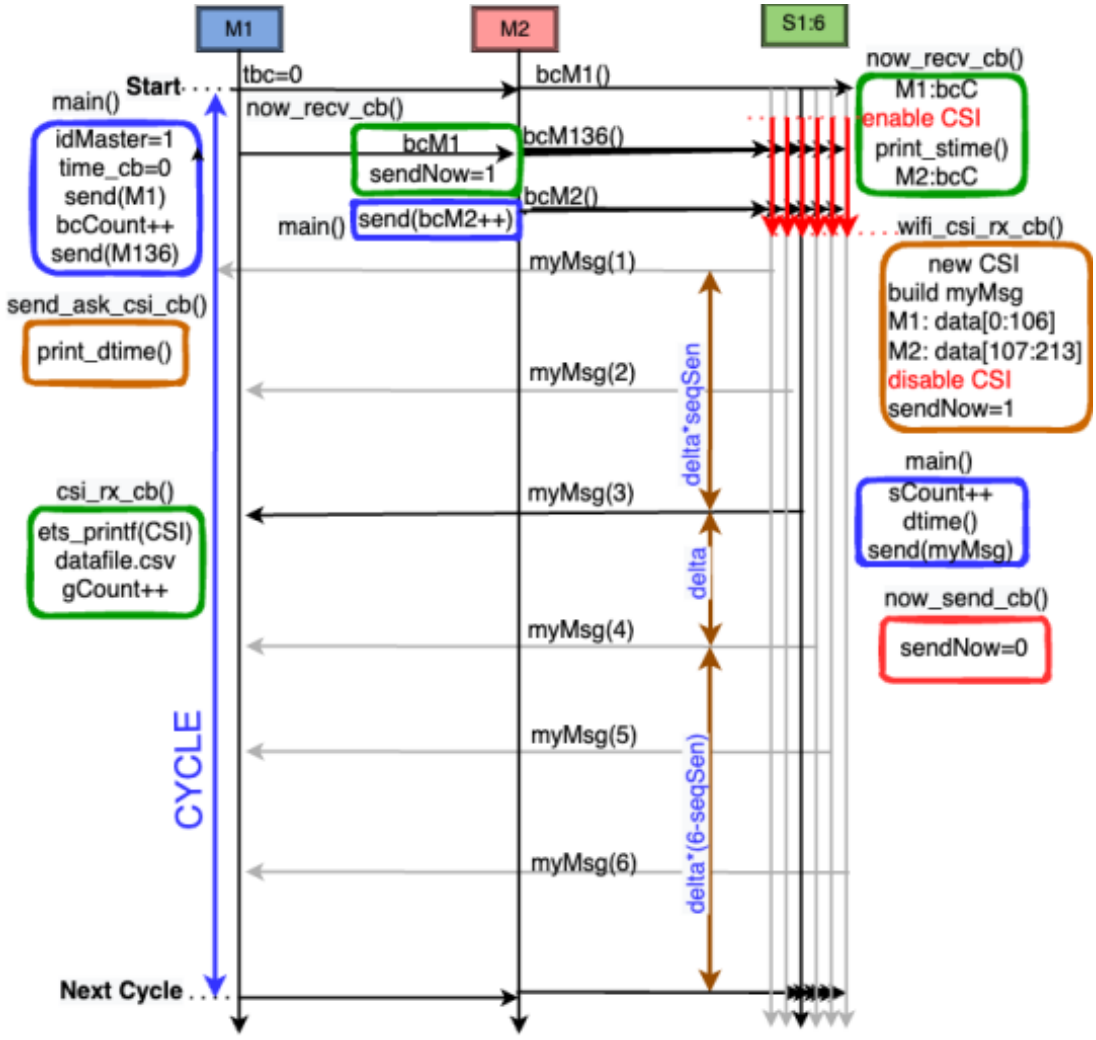
Alguns parâmetros são fundamentais para determinar os tempos de transmissão de dados no projeto: o tempo alocado para cada transmissão individual dentro do ciclo (Δ), o número de sensores (NS = 4) e a duração total do ciclo (T_{Ciclo}), obtida a partir do número de unidades de tempo no ciclo (CYCLE). Com base nesses parâmetros, é possível verificar se há folga no ciclo, o que garante a flexibilidade do sistema. Essa estratégia visa reduzir a sobrescrita e a colisão de dados.

O tempo de folga (T_{Folga}) do sistema é dado pelo tempo total do ciclo menos o tempo total de transmissão, ou seja:

$$T_{Folga} = T_{Ciclo} - T_{Transmissao}, \tag{4.1}$$

em que $T_{Transmissao} = \Delta \times (unidade\ do\ ciclo) \times NS$ e $T_{Ciclo} = CYCLE \times (unidade\ do\ ciclo)$.

Figura 4.6 – Fluxo Temporal da Aquisição de Sinal CSI241. Neste projeto $NS = 4$, $CYCLE = 400m$ e $\Delta = 90ms$.



Fonte: (Bauchspiess, 2023)

Define-se $CYCLE = 40$ e *unidade do ciclo* $= 10ms$, logo $T_{Ciclo} = 400ms$. Como $NS = 4$ e *unidade do ciclo* $= 10ms$ são fixos, a equação (4.1) torna-se $T_{Folga} = 400 - (\Delta \times 40)$.

Com o objetivo de otimizar o aproveitamento do sistema de aquisição, foram testados 4 valores de Δ distintos: 7, 8, 9 e 10. Pelos resultados apresentados na tabela 4.2, observa-se que $\Delta = 10ms$ se torna inviável, pois não há folga disponível. Já os valores $\Delta = 7ms$ e $\Delta = 8ms$ foram desconsiderados nas etapas subsequentes (pré-processamento de dados).

Portanto, a escolha de $\Delta = 9ms$ foi feita com base nos resultados obtidos nas aquisições e em seu T_{Folga}, levando em consideração a necessidade de garantir uma folga adequada no ciclo de transmissão. Sua margem de $40ms$ se mostrou a mais adequada para acomodar pequenas variações, garantindo flexibilidade no sistema.

Tabela 4.2 – Valores de Δ e T_{Folga} no ciclo de transmissão.

Δ	$T_{Folga}(ms)$
7	120
8	80
9	40
10	0

4.1.3 Atividades da Rotina

As atividades ao longo do experimento são organizadas conforme a tabela 4.3. Cada janela de tempo de 10 segundos representa uma atividade.

Tabela 4.3 – Distribuição de atividades da rotina fixa no tempo.4 minutos/experimento.

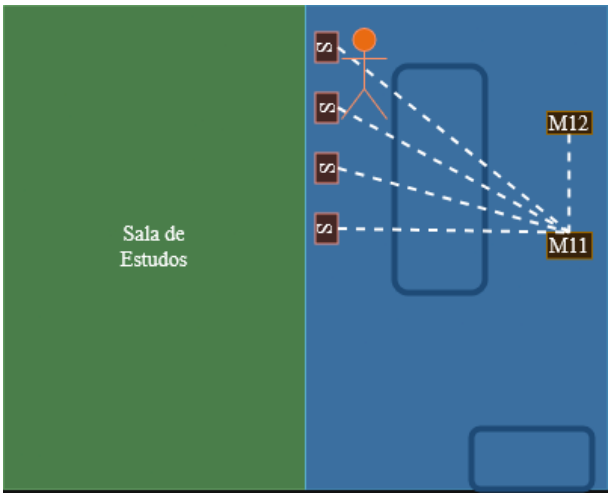
t(min)	→ 10*s*	→ 20*s*	→ 30*s*	→ 40*s*	→ 50*s*	→ 60*s*
0-1	Empty	Empty	Empty	Empty	Empty	Empty
1-2	Walk	Walk	Walk	Sit	Sit	Sit
2-3	Sit	Sit	Sit	Stand	Stand	Stand
3-4	Stand	Stand	Stand	Walk	Walk	Walk

Fonte: Adaptado de (Bauchspiess, 2023).

4.2 Aquisição de Dados - Banco da Sala de Reunião

Uma nova aquisição para testes em novos ambientes foi feita na Sala de Reunião, cuja topologia está representada na figura 4.7.

Figura 4.7 – Distribuição módulos Esp32 - Sala de Reunião.



Em função dos endereços MACs utilizados na primeira sala, os módulos da sala de reuniões eram diferentes e receberam outros rótulos: M11 e M12 tornaram-se os novos mestres M1 e M2 e a sequência de sensores virou S3-S4-S5-S6.

4.3 Pré-processamento de Dados

Em função de diversas interferências, os dados coletados nos arquivos *.csv* apresentaram irregularidades como ruídos, redundâncias, ausências etc. Portanto, há a necessidade de alguns processos prévios aos treinamentos.

4.3.1 Correção da Sequência de Sensores

Antes de se iniciarem os treinamentos, os dados coletados passaram por uma etapa de pré-processamento. Nesta fase, são removidas as linhas redundantes e preenchidas as informações ausentes. O algoritmo de preenchimento verifica sequencialmente as linhas dos arquivos em busca de valores ausentes de algum sensor. Quando encontrados, os valores são preenchidos com os últimos dados registrados para aquele sensor, assegurando a consistência da estrutura dos dados e a correta sequência dos sensores. Ao final dessa etapa, o programa indica qual a porcentagem de inserções de linhas incluídas.

4.3.2 Filtro Passa-baixas

Em seguida, os dados passam por um filtro digital analógico passa-baixas (Butterworth), bloqueando a passagem de valores superiores à frequência de corte escolhida. A implementação é feita por meio da função *butter()* (documentação disponível em https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.butter.html), sendo possível determinar: a ordem do filtro, a frequência de corte (f_0), o tipo de filtro etc. Escolheu-se um filtro de segunda ordem com $f_0 = 2Hz$.

4.3.3 Sanitização

Os valores de fase passaram por uma sanitização com intuito de manter os dados dentro do intervalo $[-\pi, \pi]$, adicionando-se ou subtraindo-se 2π caso a diferença entre elementos consecutivos ultrapasse o limite definido.

4.4 Few-Shot Learning

4.4.1 ProtoNet

A implementação de Few-Shot Learning utilizada é uma adaptação da rede prototípica **ProtoNet** presente em Bahadori, Ashdown e Restuccia (2022).

A principal ideia por trás do ProtoNet é a construção de representações prototípicas para cada classe, calculados como a média das características das amostras de suporte no espaço de representação.

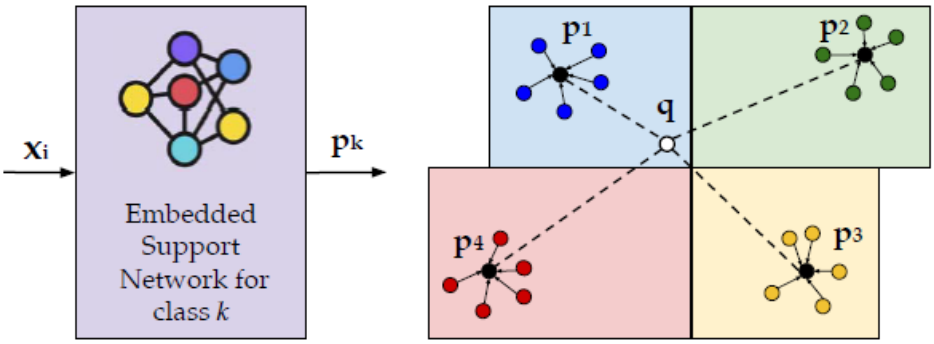
Durante o treinamento, cada tarefa envolve um conjunto de suporte (support set), que contém um pequeno número de exemplos rotulados, e um conjunto de consulta (query set), composto por amostras não vistas que precisam ser classificadas. O modelo aprende a projetar os dados em um espaço métrico onde as amostras de uma mesma classe ficam agrupadas próximas ao seu protótipo correspondente.

Matematicamente, o protótipo de uma classe k é definido como a média das representações das amostras de suporte:

$$p_k = \frac{1}{|D_k|} \sum_{(s_i, y_k) \in D_k} f_\varphi(s_i), \tag{4.2}$$

em que $f_\varphi(s_i)$ denota a função de representação que transforma os exemplos brutos em um espaço de características.

Figura 4.8 – Representação da rede prototípica de (Bahadori; Ashdown; Restuccia, 2022).



Fonte: Adaptado de (Bahadori; Ashdown; Restuccia, 2022).

4.4.2 Rede Few-Shot Adaptada

O diagrama de blocos da figura 4.9 descreve o algoritmo de Few-Shot Learning aplicado aos dados de CSI nos experimentos na UnB. Treina-se o modelo com os dados provenientes da sala de estudos e, posteriormente, aplica-se *Transfer Learning* para adaptar o modelo ao ambiente da sala de reuniões.

A leitura dos dados é realizada pela função `read_CSIKIT(trainD)`, e a separação em 80% para treinamento e 20% para teste ocorre em `train, test = split()`. A tarefa `extract_sample()` tem como saída um vetor contendo o número de classes e a quantidade de amostras no conjunto de suporte (*support set*) e no de consulta (*query set*).

O modelo recebe como entrada imagens em tons de cinza de 512x512 pixels e utiliza uma CNN com quatro camadas convolucionais sequenciais. Cada camada é seguida por normalização em lote (*BatchNorm*), ativação *ReLU* e MaxPooling 2D, reduzindo progressivamente a dimensionalidade dos mapas de características conforme apresentado na tabela 4.4.

A extração de características resulta em um vetor de dimensão final 64, que representa cada amostra no espaço latente. A classificação ocorre por meio da comparação entre as representações das amostras e os protótipos das classes, utilizando a distância euclidiana como métrica de similaridade.

Tabela 4.4 – Redução de dimensionalidade no Encoder.

Camada	**Dimensão**
Input	$512 \times 512 \times 1$
1ª camada	$256 \times 256 \times 248$
2ª camada	$128 \times 128 \times 248$
3ª camada	$64 \times 64 \times 248$
4ª camada	$32 \times 32 \times 64$
Flatten	Vetor de 32×32×64=65536 elementos

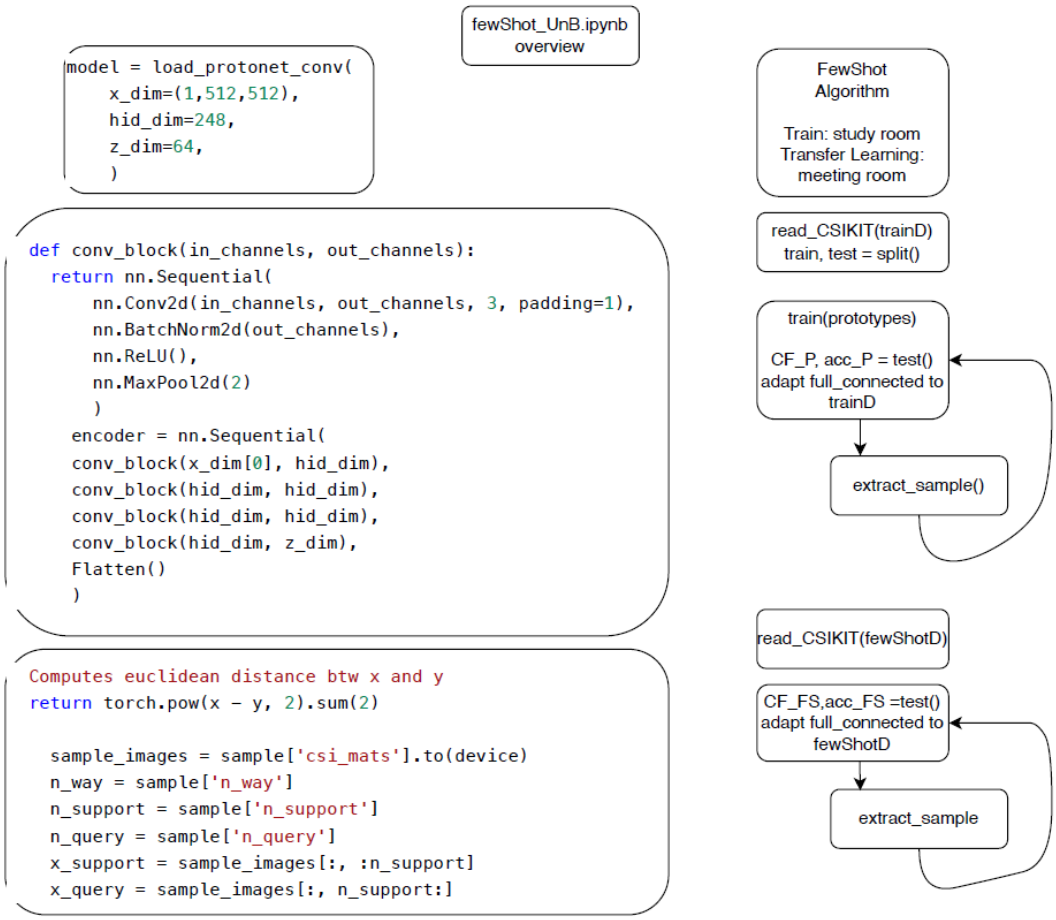
4.5 Desempenho das Redes Neurais

Para avaliar-se o desempenho das redes neurais, utilizaram-se as seguintes definições:

$$\text{Acurácia} = \frac{\text{Número de acertos}}{\text{Número total de previsões}}$$

$$\text{Acurácia (\%)} = \frac{\text{Número de acertos}}{\text{Número total de previsões}} \times 100\%$$

Figura 4.9 – Visão geral da rede Few-Shot adaptada ao trabalho na UnB.



Fonte: Professor Adolfo Bauchspiess.

5 Desenvolvimento

5.1 Instalação e Configuração de Ambiente

Antes do experimento em si, é preciso ter o ambiente de desenvolvimento da Espressif instalado e configurado no computador. No caso do sistema operacional Windows, fazendo a instalação pelo executável, cria-se um atalho (figura 5.1) de prompt de comando na Área de Trabalho que já carrega as configurações de ambiente ao ser clicado (figura 5.2).

Figura 5.1 – Atalho do terminal CMD.

É importante também verificar se alguns outros requisitos estão sendo respeitados. Uma lista pode ser encontrada em https://github.com/espressif/esp-idf/tree/master/tools/requirements.

Figura 5.2 – Terminal CMD com as configurações de ambiente sendo carregadas.

Uma vez com o terminal aberto, pode-se dar continuidade ao experimento.

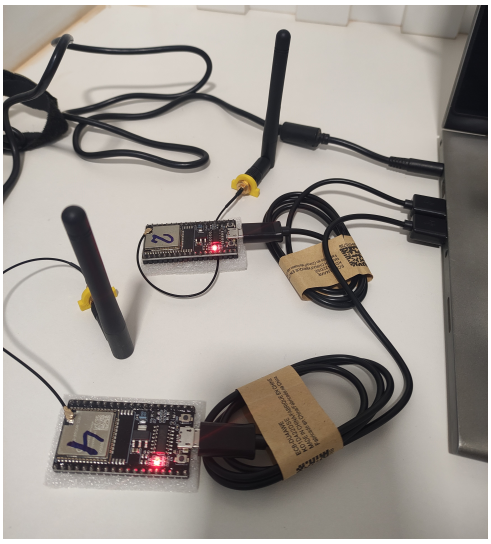
5.2 Gravação Módulos

A gravação de todos os módulos segue, praticamente, a mesma sequência. A diferença está nos diretórios dos códigos e na última linha de comando. A seguir, detalha-se o procedimento para o M1, bastando repetir os passos para todos os conjuntos. Recomenda-se,porém, que o M2 e os S sejam gravados primeiro.

5.2.1 Gravação M1/M11

O M1/M11 é ligado diretamente ao computador que fará a aquisição via cabo USB (figura 5.3). Abre-se o prompt de comando conforme figura 5.1 no diretório onde o código principal do mestre 1 se encontra.

Figura 5.3 – Gravação dos Módulos Esp32



A partir disso, a sequência de comandos abaixo começa a ser executada.

```
idf.py set-target esp32 fullclean
idf.py menuconfig
```

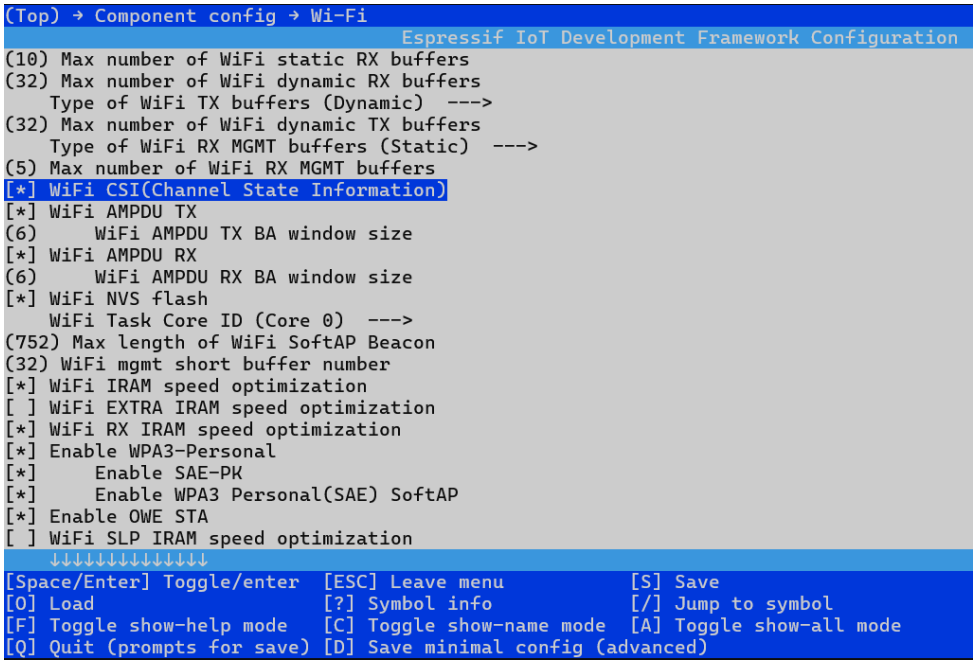
Quando o menu de configuração for aberto, é necessário sempre habilitar a opção de CSI (figura 5.4) pelo caminho *Component config* - - -> *Wi-Fi* - - ->. E continuar com os próximos comandos.

Por fim, executam-se os últimos comandos (listados abaixo).

```
idf.py build
idf.py -p COMx flash
idf.py -p COMx monitor | findstr "CSI_DATA" > nome_arquivo.csv
```

O número da porta utilizada (*COMx*) pode ser encontrado no **Gerenciador de Dispositivos** do Windows.

Figura 5.4 – Opção Wi-Fi CSI(Channel State Information) habilitada no menu de configuração.



5.2.2 Gravação M2/M12 e S

A gravação dos outros módulos (M2, M12, S6, S7, S8, S9, S3, S4, S5) segue a mesma sequência de comandos até o flash.

5.3 Procedimento Experimental

Os módulos gravados são dispostos pelas salas do LARA como mostram as figura 5.5 e figura 5.6.

O experimento consiste na reprodução de uma sequência de atividades fixa por 4 minutos nas salas do laboratório 5.5. A posição dos módulos e os elementos no ambiente são fixos, mantendo o padrão em todas as aquisições.

Os dados provenientes do procedimento prático são salvos em arquivos de extensão *csv*. Nas figura 5.7 e figura 5.8, em verde, observa-se uma sequência correta dos sensores, e em vermelho, uma sequência que será corrigida no pré-processamento. Os tempos do ciclo TDMA são destacados em azul. Entre os elementos do arquivo, incluem-se: os números de identificação dos sensores (idS), o tempo total da aquisição (t_M), o tempo relativo decorrido no ciclo atual (tin_cycle) e os dados de CSI ($[data]$).

Em seguida, iniciam-se as fases de pré-processamento e limpeza, seguidas pelo treinamento e validação das redes neurais.

Figura 5.5 – Sala de Estudos do LARA com indicação do posicionamento dos módulos.

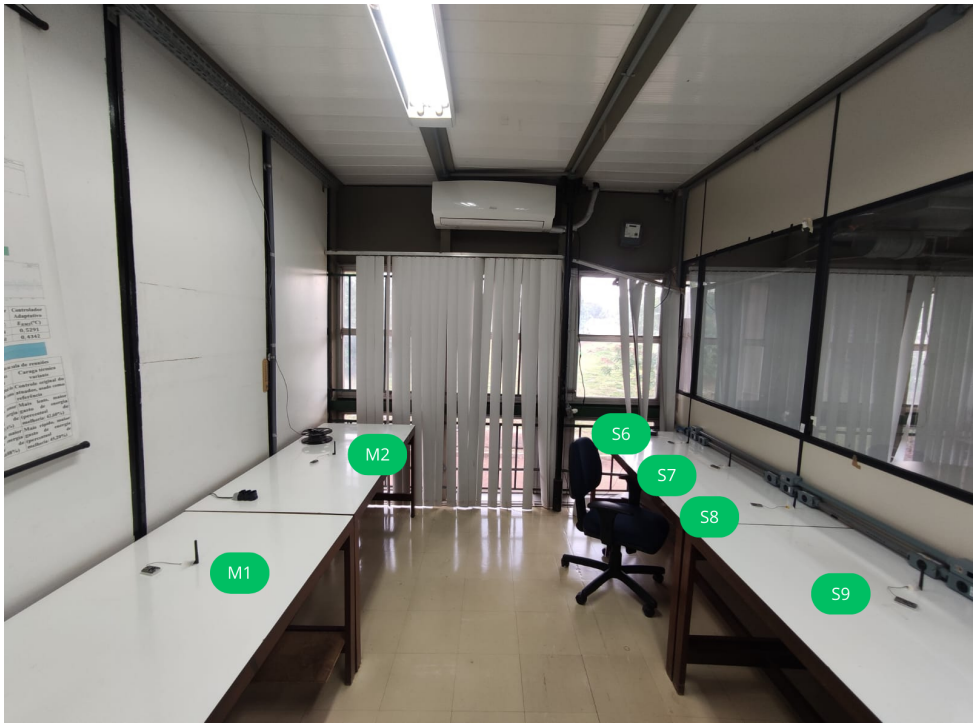


Figura 5.6 – Sala de Reunião do LARA com indicação do posicionamento dos módulos.

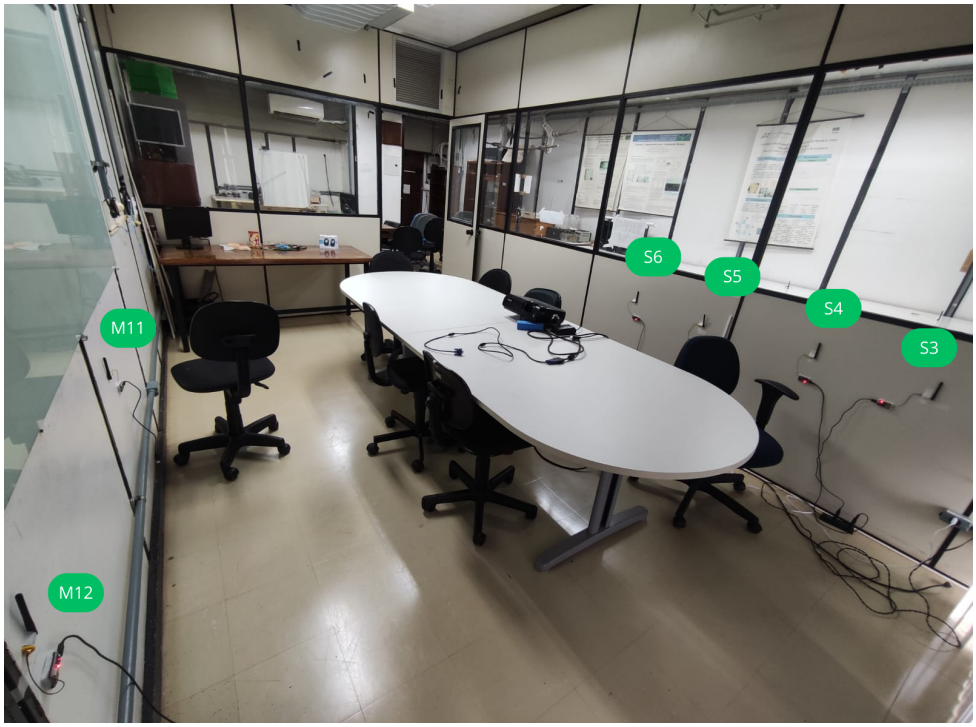


Figura 5.7 – Parte de um arquivo gerado para uma rotina na sala de estudos antes de passar pelo algoritmo de inserções.

CSI_DATA: gC	idS	sC	sC_M	bc1	bc2	bcM1	retry	RSSI1	RSSI2	tCSI1_S	tCSI2_S	tdma_S	t_M(ms)	dt_rxM	tin_cycle	[data]
CSI_DATA: 1	6	1	1	0	0	0	0	-30	-26	4294830.5	4294837.963	0.91	27.644	27.644	27.644	[-5,11,-4,12,-3,11,-4,13,-3,10,-3,12,-2,11,
CSI_DATA: 2	7	1	1	0	0	0	0	-35	-36	4294825.09	4294832.939	89.506	108.494	108.494	108.494	[-11,1,-11,2,-10,3,-10,3,-9,3,-9,2,-8,3,-7,3
CSI_DATA: 3	8	1	1	0	0	0	0	-42	-29	4294822.287	4294830.216	171.88	192.11	192.11	192.11	[-18,14,-17,13,-16,14,-18,13,-17,13,-17,1
CSI_DATA: 4	9	1	1	0	0	0	0	-31	-36	4294824.9	4294832.2	263.234	284.476	284.476	284.476	[8,-9,9,-10,7,-9,8,-11,7,-10,8,-11,7,-11,8,-
CSI_DATA: 5	6	2	2	1	1	1	0	-31	-26	252.663	261.209	18.532	418.247	133.771	19.871	[-12,-10,-12,-10,-12,-8,-13,-9,-13,-8,-12,-
CSI_DATA: 6	7	2	2	1	1	1	0	-35	-37	247.641	256.188	108.391	508.59	223.583	109.683	[-12,4,-11,4,-10,4,-10,5,-10,5,-9,4,-9,4,-8
CSI_DATA: 7	8	2	2	1	1	1	0	-41	-28	244.915	253.462	190.761	592.492	308.16	194.116	[-22,4,-22,4,-21,4,-22,5,-21,3,-21,3,-20,2
CSI_DATA: 8	9	2	2	1	1	1	0	-32	-36	246.72	255.267	282.122	681.88	397.404	283.504	[3,-11,3,-12,2,-11,2,-12,2,-12,2,-12,2,-12,
CSI_DATA: 9	6	3	3	2	2	2	0	-29	-26	642.664	651.24	18.596	808.224	126.344	19.848	[1,-13,0,-12,0,-13,0,-13,-1,-13,0,-12,-1,-1
CSI_DATA: 10	7	3	3	1	2	2	0	-35	-35	247.641	646.218	108.429	898.888	217.08	110.512	[-12,4,-11,4,-10,4,-10,5,-10,5,-9,4,-9,4,-8
CSI_DATA: 11	8	3	3	2	2	2	0	-41	-28	634.916	643.493	190.801	980.519	298.639	192.143	[23,0,22,0,22,-1,22,-1,22,-1,21,0,20,-1,2(
CSI_DATA: 12	9	3	3	2	2	2	0	-33	-36	636.721	645.298	282.135	1072.108	390.228	283.732	[3,8,3,8,4,8,4,8,4,8,4,8,4,8,4,8,4,7,5,7,5,8
CSI_DATA: 13	6	4	4	5	3	5	0	-31	-26	1812.659	1821.249	18.449	1978.257	906.149	19.86	[-5,13,-5,14,-3,12,-4,14,-3,12,-4,13,-3,12
CSI_DATA: 14	8	4	4	5	3	5	0	-39	-30	1804.914	1813.503	190.621	2150.43	1078.322	192.33	[-18,-5,-19,-5,-17,-4,-20,-4,-18,-4,-19,-5,-
CSI_DATA: 15	9	4	4	5	3	5	0	-35	-38	1806.717	1815.306	281.98	2241.865	1169.757	283.468	[-3,-6,-4,-7,-4,-6,-4,-7,-4,-6,-5,-7,-5,-6,-5,-

Figura 5.8 – Parte de um arquivo gerado para uma rotina na sala de reuniões antes de passar pelo algoritmo de inserções.

CSI_DATA: gC	idS	sC	sC_M	bc1	bc2	bcM1	retry	RSSI1	RSSI2	tCSI1_S	tCSI2_S	tdma_S	t_M(ms)	dt_rxM	tin_cycle	[data]
CSI_DATA: 1	3	1	1	0	0	0	0	-39	-40	4294671.28	4294668.489	0.564	24.212	24.212	24.212	[15,18,15,18,16,17,16,18,15,18,15,18,14,18,14,
CSI_DATA: 2	4	1	1	0	0	0	0	-67	-60	4294673.252	4294670.465	83.479	100.28	100.28	100.28	[4,-6,4,-6,3,-8,5,-8,5,-8,6,-10,6,-10,7,-9,8,-10,8,-
CSI_DATA: 3	5	1	1	0	0	0	0	-39	-52	4294668.09	4294665.214	175.807	195.426	195.426	195.426	[6,-9,6,-9,5,-10,6,-10,5,-10,5,-11,5,-10,5,-11,5,-1
CSI_DATA: 4	6	1	1	0	0	0	0	-53	-50	4294674.81	4294671.294	269.97	286.775	286.775	286.775	[10,-4,9,-4,9,-6,9,-5,9,-5,8,-6,7,-6,7,-6,7,-6,6,-7,7
CSI_DATA: 5	3	2	2	1	1	1	0	-39	-40	89.272	95.55	16.245	417.204	130.429	17.706	[21,-7,21,-7,21,-8,22,-7,21,-7,21,-7,20,-7,20,-6,2
CSI_DATA: 6	4	2	2	1	1	1	0	-67	-60	91.249	97.526	99.329	500.265	213.49	100.767	[-6,-6,-6,-6,-7,-6,-7,-7,-8,-7,-7,-7,-9,-8,-9,-9,-9,
CSI_DATA: 7	5	2	2	1	1	1	0	-38	-52	85.997	92.274	191.595	593.137	306.362	193.639	[-10,-1,-11,-1,-10,0,-12,0,-11,0,-12,0,-11,0,-12,1,
CSI_DATA: 8	6	2	2	1	1	1	0	-52	-50	92.64	98.341	285.754	686.68	399.905	287.182	[8,10,9,10,9,9,9,9,9,8,9,8,9,7,10,6,10,6,10,5,11,5
CSI_DATA: 9	3	3	3	2	2	2	0	-39	-40	479.314	485.191	15.997	806.823	120.143	17.325	[11,-11,11,-12,10,-11,12,-12,11,-11,11,-11,10,-11,
CSI_DATA: 10	5	3	3	2	2	2	0	-38	-51	476.38	481.914	191.78	982.57	295.89	193.72	[-1,9,-1,10,0,9,0,11,0,10,1,11,1,10,1,10,1,10,2,1(
CSI_DATA: 11	6	3	3	2	2	2	0	-52	-50	482.106	487.982	285.871	1080.439	393.759	290.941	[-10,-9,-11,-8,-11,-7,-12,-8,-11,-6,-11,-6,-11,-5,-1
CSI_DATA: 12	3	4	4	3	3	3	0	-40	-41	869.292	875.137	15.959	1196.783	116.344	17.286	[-4,12,-5,13,-3,11,-5,14,-4,12,-5,13,-5,12,-4,12,-5
CSI_DATA: 13	4	4	4	3	3	3	0	-67	-60	871.271	877.116	99.465	1280.22	199.781	100.723	[-2,5,-2,6,-2,6,-3,8,-3,6,-3,9,-4,8,-4,9,-5,8,-4,8,-4
CSI_DATA: 14	5	4	4	3	3	3	0	-38	-52	866.15	871.86	191.799	1375.583	295.144	196.86	[4,12,4,14,5,11,6,15,5,12,6,15,5,13,6,13,5,13,7,
CSI_DATA: 15	3	5	5	4	4	4	0	-39	-41	1259.475	1267.652	16.823	1590.65	214.482	20.567	[14,-11,14,-11,13,-11,14,-11,14,-11,14,-10,13,-10

5.4 Resultados

5.4.1 Aquisição de Dados e Pré-processamento

Na etapa da sala de estudos, foram adquiridos 20 arquivos com rotina completa. O algoritmo de pré-processamento indicou que a quantidade de inserções de linhas de correção variou entre 19% e 35% (tabela 5.1), aproximadamente.

Tabela 5.1 – Quantidade de inserções de linhas no pré-processamento (Dados Sala de Estudos).

Arquivos da Aquisição de Dados	Inserções de Linhas
1	28.32%
2	31.31%
3	25.91%
4	31.58%
5	25.34%
6	25.91%
7	21.33%
8	20.64%
9	23.43%
10	19.70%
11	23.74%
12	20.35%
13	21.86%
14	24.01%
15	20.86%
16	27.20%
17	27.66%
18	30.33%
19	22.55%
20	32.46%

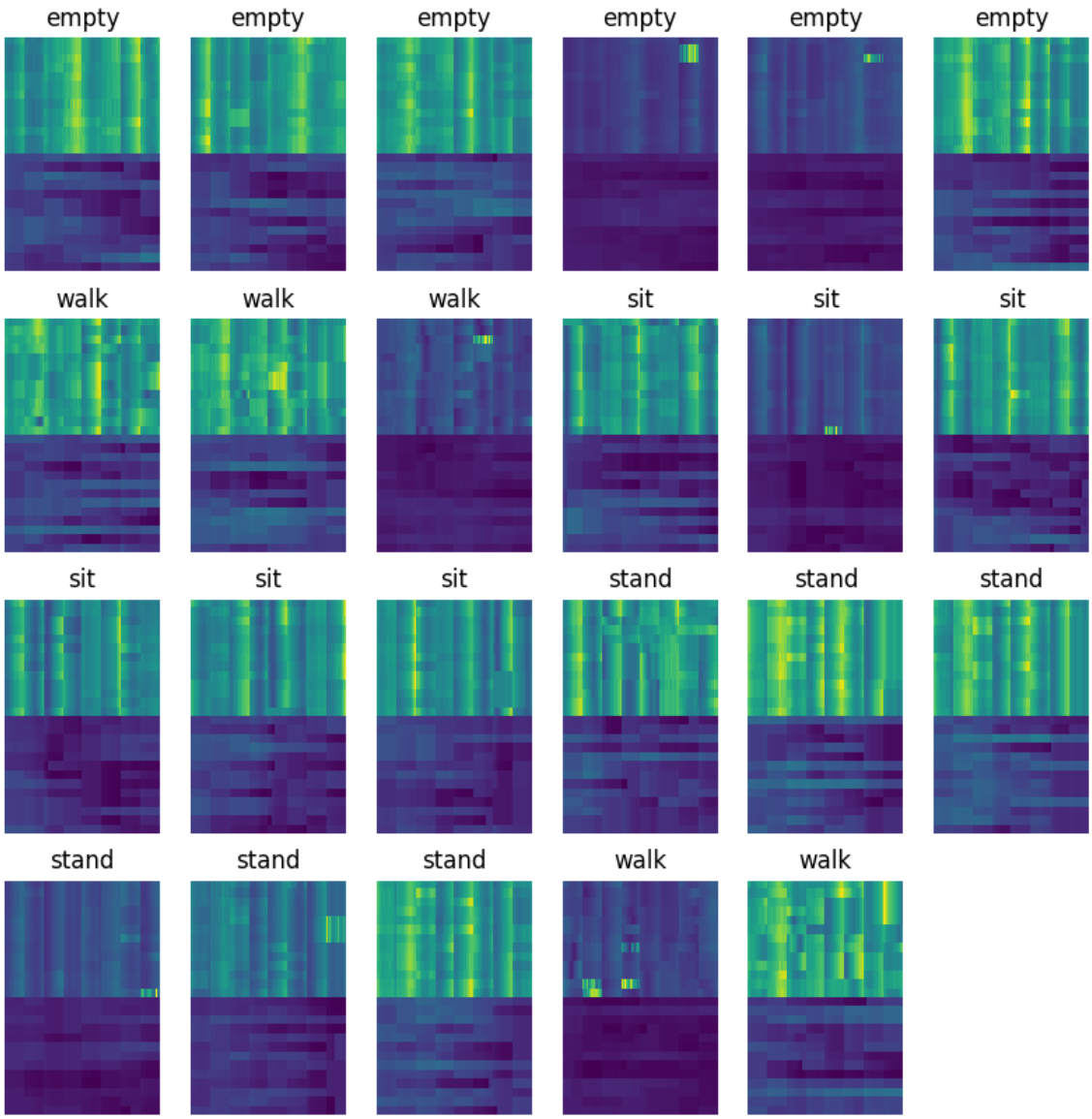
Na sala de reunião, foram adquiridos 6 arquivos com a rotina completa. A quantidade de inserções de linhas pode ser vista na tabela 5.2.

Tabela 5.2 – Quantidade de inserções de linhas no pré-processamento (Dados Sala de Reunião).

Arquivos da Aquisição de Dados	Inserções de Linhas
1	29.39%
2	25.57%
3	28.02%
4	26.79%
5	34.08%
6	20.81%

5.4.2 CSI

Figura 5.9 – Sinal CSI de um rotina completa de aquisição de dados em forma de imagem.



A figura 5.9 representa o sinal CSI de uma rotina completa de aquisição de dados de 4 minutos. Cada imagem isolada representa 10s, o equivalente a 100 linhas. Da aquisição inteira, espera-se que sejam obtidas, aproximadamente, 2400 linhas em cada arquivo. As partes reais das imagens encontram-se na parte de cima e as imaginárias na parte de baixo.

A figura 5.10 apresenta um total de $8 \times 52 = 416$ amostras, correspondentes às partes real e imaginária da coleta de 4 sensores em resposta a 2 mestres. Inicialmente, essas partes são representadas separadamente em duas figuras e, posteriormente, combinadas em uma única imagem com 26 linhas, onde as linhas de 0 a 13 correspondem à parte real e as de 14 a 25 à parte imaginária. Na última subfigura, os dados passaram pelo procedimento de

Análise de Componentes Principais (ACP, em inglês Principal Component Analysis, PCA), eliminando redundâncias e comprimindo a informação em uma imagem 26x26 pixels.

Figura 5.10 – Sinal CSI representado em forma de imagem.

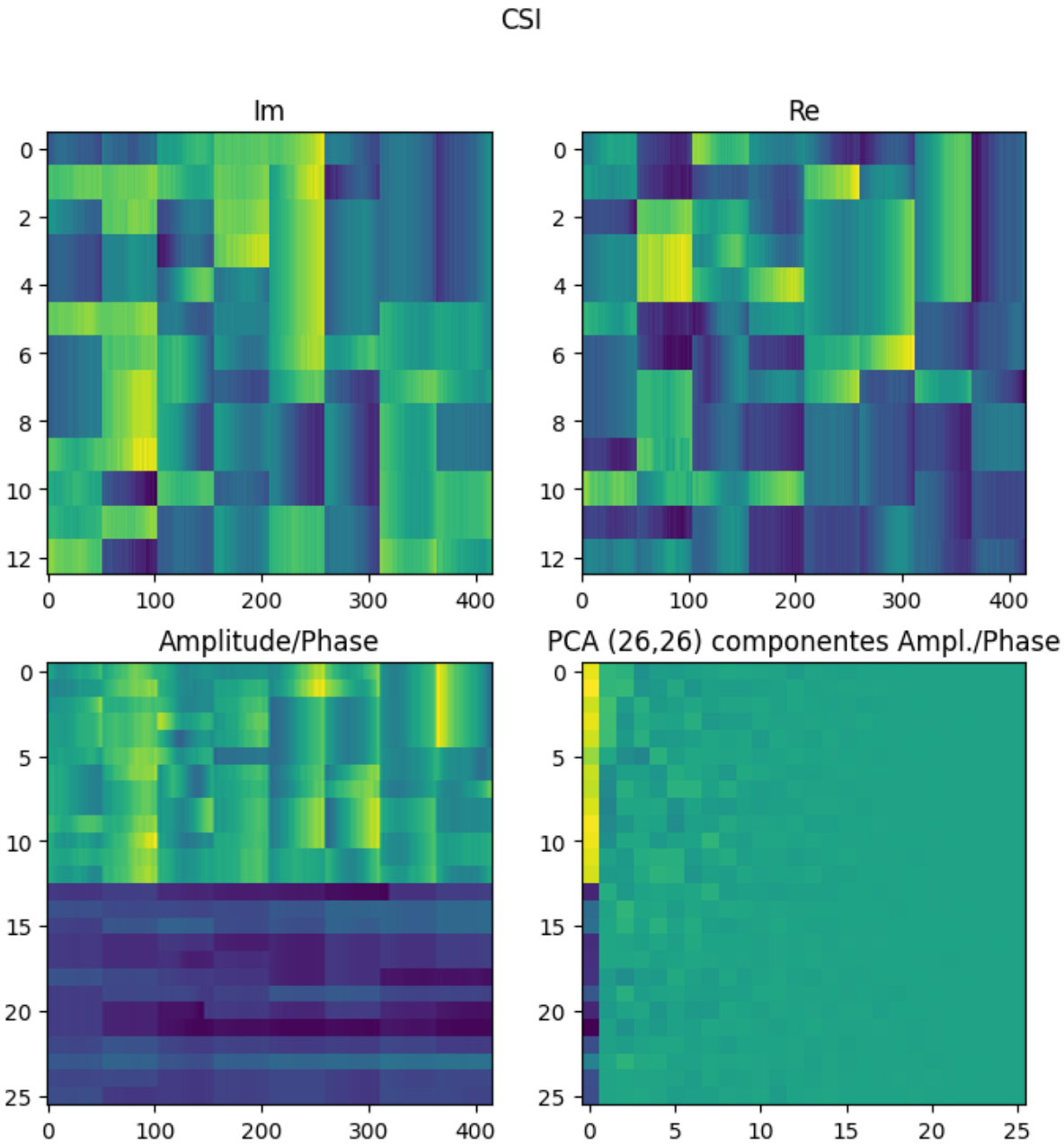
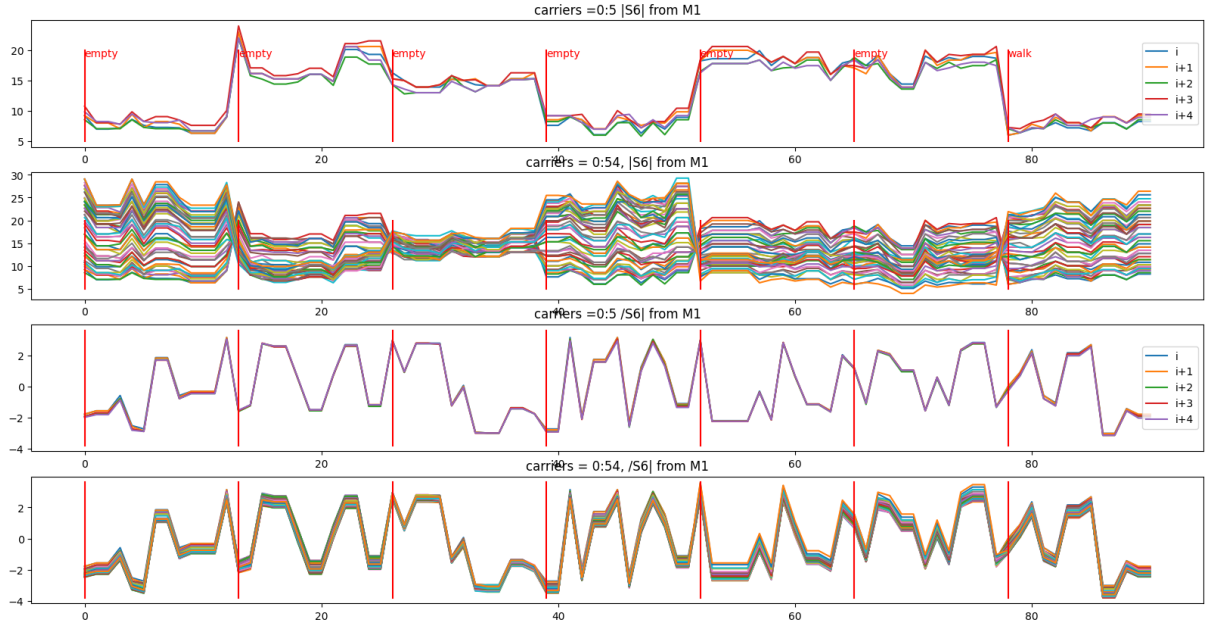


Figura 5.11 – Subportadoras CSI consecutivas, sensor 6.



Na figura 5.11, são representadas as portadoras do sensor 6 para o M1 durante um período de 1 minuto e 30 segundos de aquisição de dados. São mostradas a amplitude e a fase das subportadoras de 0 a 5 e de 0 a 52. Ao analisar o conjunto menor de subportadoras, observa-se que há pouca variação entre elas, diferentemente do cenário com as 52. Além disso, a figura indica que, em relação à fase, a dispersão entre os sinais é pequena.

5.4.3 CNN

A CNN com o melhor modelo obteve uma acurácia de treinamento de 0,9683 e uma acurácia de teste de 0,5337. O treinamento foi realizado em 2 épocas, com 1000 episódios em cada época, e a avaliação final foi feita com 100 episódios.

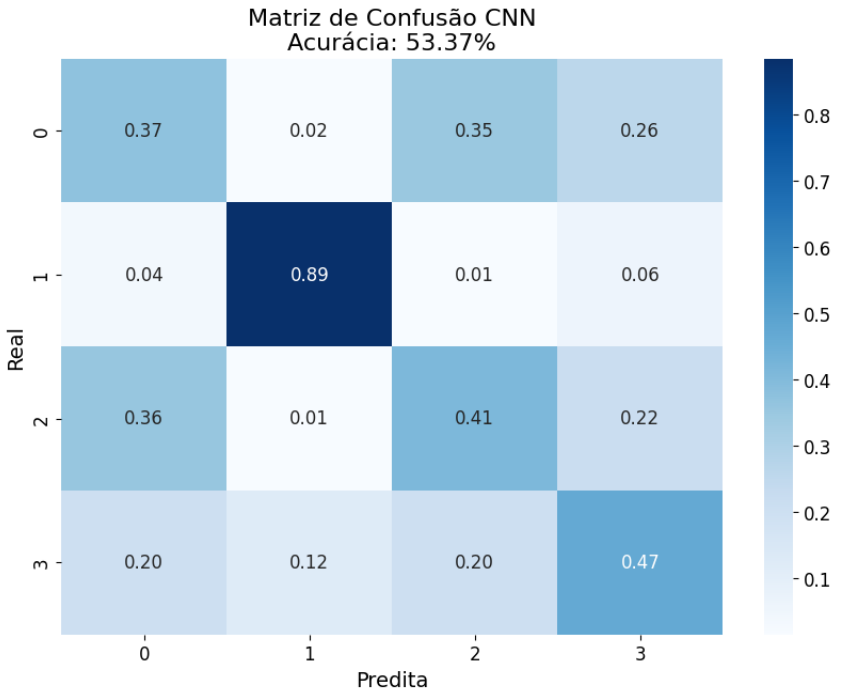
5.4.4 Few-Shot

Foram exploradas diferentes configurações da rede Few-Shot para se avaliar os impactos no desempenho do modelo. Os resultados das acurácias de validação podem ser vistos na tabela 5.3. Os parâmetros em ordem serão descritos a seguir.

O parâmetro B10s, que representa quantos blocos com 4 linhas de CSI foram usados para gerar as imagens, foi variado entre 10 e 15.

Avaliou-se, também, os efeitos da aplicação de um filtro passa-baixas, representado pelas opções nF (sem filtro) e wF (com filtro). A sanitização de fase foi representada pelas possibilidades nS (sem sanitização) e wS (com sanitização). Outro aspecto analisado foi

Figura 5.12 – Matriz de Confusão da CNN. 305 imagens de Treinamento e 132 de Testes (30%), 19 experimentos de 4 min.



a utilização da compactação PCA, indicada por nPCA (sem compactação) e wPCA (com compactação).

O número de amostras de consulta variou entre 2 e 8, enquanto o número de amostras de suporte foi testado dentro do intervalo de 3 a 4.

5.5 Discussão

5.5.1 Interferências Eletromagnéticas

Os primeiros testes demonstraram taxas de inserções maiores que 40%. Ao avaliar o ambiente, observou-se que alguns módulos estavam muito próximos a elementos ferromagnéticos. Surgiu-se a suspeita de que isso estaria influenciando nos resultados obtidos. Em função disso, os dispositivos foram rearranjados livremente sobre as mesas, e novos testes demonstraram uma redução na quantidade de inserções (tabela 5.1).

5.5.2 Esp32

Os módulos ESP32 demonstraram funcionalidade relativa no projeto. No entanto, há incertezas quanto à adequação de um dispositivo com foco em telecomunicações para aplicações de sensoriamento, especialmente considerando que a qualidade e a integridade dos dados são de importância primordial para essa aplicação. Além disso, a limitada docu-

Tabela 5.3 – Tabela de Acurácia das Configurações.

Configuração	Acurácia
acc(B10s14_wF_wS_wPCA_n4sup_n3qry)	0.613
acc(B10s14_wF_wS_wPCA_n4sup_n4qry)	0.572
acc(B10s13_wF_wS_wPCA_n4sup_n4qry)	0.566
acc(B10s15_wF_wS_wPCA_n4sup_n4qry)	0.562
acc(B10s15_wF_wS_nPCA_n4sup_n3qry)	0.558
acc(B10s11_wF_wS_wPCA_n4sup_n4qry)	0.546
acc(B10s14_nF_wS_wPCA_n4sup_n3qry)	0.545
acc(B10s15_wF_wS_nPCA_n4sup_n4qry)	0.537
acc(B10s10_wF_wS_wPCA_n4sup_n3qry)	0.533
acc(B10s10_wF_wS_wPCA_n4sup_n4qry)	0.532
acc(B10s15_wF_wS_wPCA_n4sup_n5qry)	0.531
acc(B10s12_wF_wS_wPCA_n4sup_n3qry)	0.529
acc(B10s14_wF_nS_wPCA_n4sup_n3qry)	0.527
acc(B10s13_wF_wS_wPCA_n4sup_n8qry)	0.522
acc(B10s15_wF_wS_wPCA_n4sup_n3qry)	0.516
acc(B10s12_wF_wS_wPCA_n4sup_n4qry)	0.514
acc(B10s14_nF_nS_wPCA_n4sup_n3qry)	0.504
acc(B10s12_wF_wS_wPCA_n4sup_n8qry)	0.500

mentação dos métodos e protocolos (ESP-NOW, por exemplo) da empresa Espressif levanta questionamentos sobre sua adequação ao trabalho.

Por outro lado, o dispositivo se apresenta como uma alternativa promissora em aplicações de baixo custo, como é o caso deste projeto.

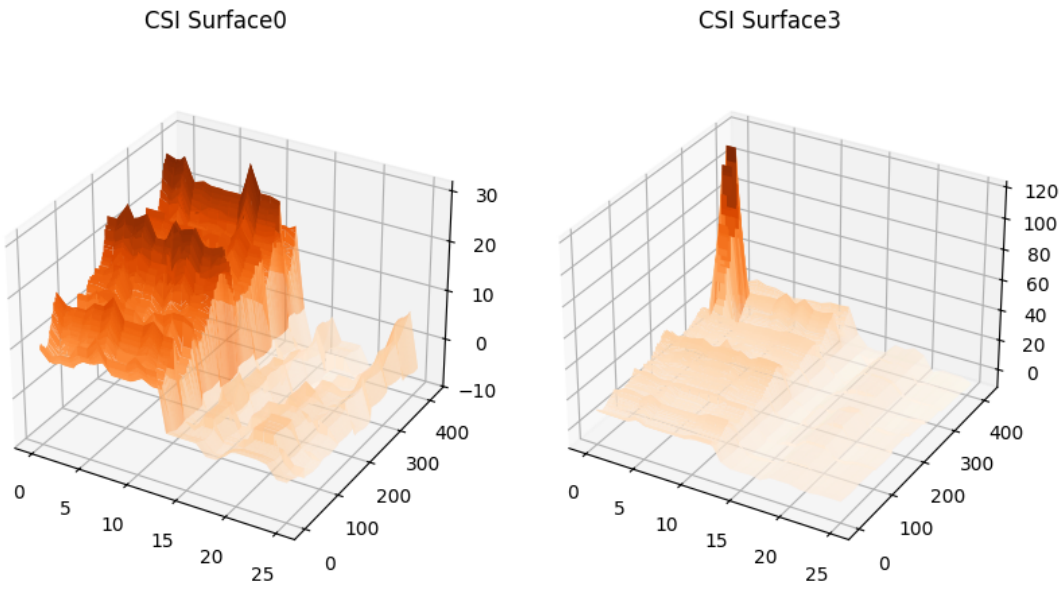
5.5.3 Qualidade dos Dados

Observa-se, nas tabela 5.1 e tabela 5.2, que, embora a maioria das inserções esteja abaixo de 30%, essa porcentagem ainda representa uma influência significativa do algoritmo de correção. Isso sugere que pode haver algum problema na coleta dos dados, o que pode ter afetado a precisão dos resultados. Para que as imagens fossem geradas de maneira adequada, esperava-se que cada arquivo resultante das rotinas experimentais contivesse, pelo menos, 2400 linhas de CSI. No entanto, ao avaliar os dados, observou-se que alguns arquivos apresentam um número consideravelmente inferior de linhas, o que justifica a quantidade de adição de informações no pré-processamento.

Nota-se pela figura 5.13 a presença de alguns valores atípicos nas imagens de CSI. Idealmente, imagens com valores muito discrepantes deveriam ser descartadas, pois não seriam adequadas para o experimento. No entanto, atualmente, não há nenhum mecanismo

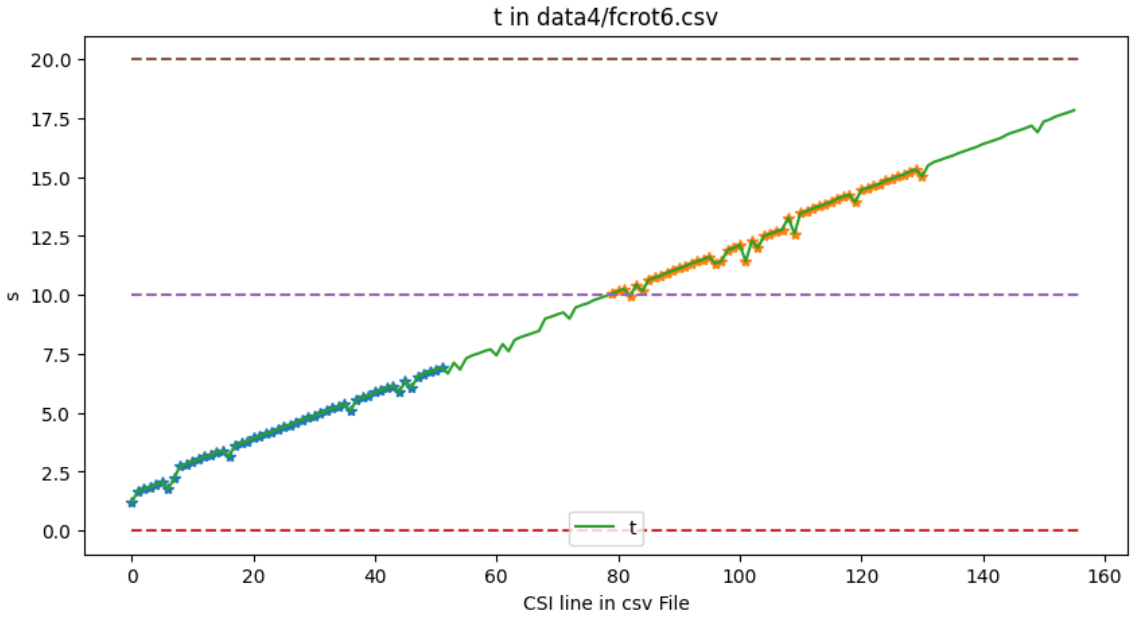
no processo que realize essa filtragem, o que permite a utilização de dados corrompidos.

Figura 5.13 – Superfícies 3D com valores de amplitude e fase de uma imagem. A subfigura da esquerda mostra valores condizentes, a da direita representa um outlier.



5.5.4 Uso Parcial dos Dados

Figura 5.14 – Demonstrativo de quantas linhas dos arquivos estão sendo utilizada ao longo de duas janelas de 10s. A linha verde representa o tempo e as marcações em azul e laranja indicam o quanto do arquivo foi utilizado na geração das imagens.



Avaliou-se pela figura 5.14 que parte dos arquivos não está sendo utilizada. Desconfia-se que a utilização dos padrão de tempo para a construção de imagens não esteja adequado,

visto que é um parâmetro impreciso em um sistema como esse.

5.5.5 Desempenho das Redes Neurais

A CNN obteve acurácia de 96,83% no treinamento e 53,37% no teste. A discrepância nos valores indica que houve *overfitting* (sobreajuste).

Com base na tabela 5.3, verifica-se que a configuração de rede Few-Shot com a maior acurácia foi: B10s=14, com filtro passa-baixas aplicado no pré-processamento, sanitização de fase realizada, compactação PCA utilizada, 4 amostras de suporte e 3 amostras de consulta. Essa combinação alcançou uma acurácia de 61,3%, sendo a melhor entre todas as configurações testadas.

6 Conclusões

Diversas abordagens para determinar o reconhecimento de atividades humanas (HAR) utilizando uma rede de sensores Wi-Fi foram exploradas. A pesquisa avaliou o desempenho de uma rede Few-Shot Learning usando um banco de dados com atividades típicas de escritórios/laboratórios adquirido no Laboratório de Automação e Robótica da Universidade de Brasília.

Apesar da acurácia de 61,3% obtida ser um bom ponto de partida, esse valor está abaixo do esperado para um modelo de aprendizado de máquina aplicado a essa tarefa. Acredita-se que diversos fatores, como interferências eletromagnéticas, o tipo de materiais presentes nos ambientes de coleta de dados, e a qualidade das aquisições, possam ter impactado os resultados.

6.1 Projetos Futuros

Sugere-se como formas gerais de continuidade do projeto apresentado aqui, as melhorias: aumento da diversidade do banco de dados de treinamento/validação e desenvolvimento de sistema voltado para tempo real.

Em relação a pontos específicos, vale ressaltar uma revisão no método de aquisição de dados com alternativas que não utilizem apenas métricas temporais na criação das imagens. Além disso, seria interessante implementar outras abordagens de filtro nas construções das imagens, evitando que valores indevidos atrapalhem o sistema.

A expansão do projeto para outros ambientes de automação tem potencial para diversas aplicações no contexto de cidades inteligentes. Espera-se que, futuramente, a aplicação contribua, não só em automações prediais, como também em cenários residenciais, dando assistência para pessoas em situação de risco.

Referências

BAHADORI, N.; ASHDOWN, J.; RESTUCCIA, F. Rewis: Reliable wi-fi sensing through few-shot multi-antenna multi-receiver csi learning. **2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)**, p. 50–59, 2022. Citado nas pp. 10, 19, 30, 33 e 37.

BAUCHSPIESS, A. **Machine Learning Algorithms for Occupancy-based Energy Smart Buildings**. [*S.l.*], 2023. Citado nas pp. 15, 20, 32, 33, 34 e 35.

BAUCHSPIESS, A.; MANE, T.; FüRST-WALTER, I.; BECKER, J. Distributed esp-now csi human activity recognition for few-shot energy saving in building automation. **2024 IEEE 14th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)**, p. 352–357, 2024. Citado na p. 33.

BRASIL. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709compilado.htm. Citado na p. 15.

ECONOMIC, U. N. D. of; AFFAIRS, S. **The Sustainable Development Goals Report 2023: Special Edition**. 2023. Disponível em: https://www.un-ilibrary.org/content/books/9789210024914/read. Citado na p. 17.

FEI, H.; XIAO, F.; HUANG, H.; SUN, L. Indoor static localization based on fresnel zones model using cots wi-fi. **Journal of Network and Computer Applications**, v. 167, p. 102709, 2020. ISSN 1084-8045. Disponível em: https://www.sciencedirect.com/science/article/pii/S1084804520301831. Citado na p. 25.

GALLO, E. A.; RIBEIRO, F. N. **Índice de Conforto Térmico ISO7730 em Automação Predial**. Dissertação (Monografia) — Universidade de Brasília, Brasília, 2007. Citado na p. 18.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [*S.l.*]: MIT Press, 2016. http://www.deeplearningbook.org. Citado nas pp. 22 e 23.

KHAPRA, M. M. **Deep Learning Tutorial**. 2017. Citado na p. 22.

LATHI, B. P. **Sinais e Sistemas Lineares**. [*S.l.*]: Bookman, 2006. ISBN 9788577803910. Citado nas pp. 19 e 22.

MA, Y.; ZHOU, G.; WANG, S. Wifi sensing with channel state information: A survey. **ACM Computing Surveys**, v. 52, n. 46, p. 1–36, 2019. Citado na p. 19.

MONTEIRO, L. de S. **CNN para Reconhecimento de Atividades em Ambientes Prediais com Rede de Sensores Wi-Fi**. Dissertação (Monografia) — Universidade de Brasília, Brasília, 2024. Citado na p. 15.

PRUDENTE, F. **Automação Predial e Residencial - Uma Introdução**. Rio de Janeiro: Grupo GEN, 2011. Citado na p. 18.

RAWAT, A.; KAUSHIK, R.; TIWARI, A. An overview of mimo ofdm system for wireless communication. **International Journal of Technical Research & Science**, p. 1–4, 2021. Citado na p. 19.

RODRIGUES, A. S.; SILVA, M. P. M. da. **Controle Antecipativo por estimativa de carga térmica em vídeo visando eficiência energética na climatização de ambiente predial**. Dissertação (Monografia) — Universidade de Brasília, Brasília, 2018. Citado na p. 15.

SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificais para Engenharia e Ciências Aplicadas**. [*S.l.*]: ARTLIBER EDITORA LTDA, 2016. ISBN 88588098879. Citado na p. 21.

SNELL, J.; SWERSKY, K.; ZEMEL, R. S. Prototypical networks for few-shot learning. **CoRR**, abs/1703.05175, 2017. Disponível em: http://arxiv.org/abs/1703.05175. Citado na p. 23.

SYSTEMS, E. **ESP-NOW Wireless Communication Protocol**. 2023. Disponível em: https://www.espressif.com/en/solutions/low-power-solutions/esp-now. Citado na p. 28.

SYSTEMS, E. **Get Started - ESP32**. 2023. Disponível em: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/get-started/. Citado na p. 26.

UN-HABITAT, T. U. N. H. S. P. **Urban Energy**. 2024. Disponível em: https://unhabitat.org/topic/urban-energy. Citado na p. 17.

XIAO, J.; ZHOU, Z.; YI, Y.; NI, L. M. A survey on wireless indoor localization from the device perspective. **ACM Computing Surveys**, v. 49, n. 25, p. 1–31, 2016. Citado na p. 20.

YANG, J.; CHEN, X.; WANG, D.; ZOU, H.; LU, C.; SUN, S.; XIE, L. Deep learning and its applications to wifi human sensing: A benchmark and a tutorial. 07 2022. Citado nas pp. 23 e 24.

ZHANG, Y.; YIN, Y.; WANG, Y.; AI, J.; WU, D. Csi-based location-independent human activity recognition with parallel convolutional networks. **Computer Communications**, v. 197, p. 87–95, 2023. ISSN 0140-3664. Disponível em: https://www.sciencedirect.com/science/article/pii/S0140366422004157. Citado na p. 21.

ZHURAVCHAK, A. **Human Activity Recognition based on WiFi CSI data**. Dissertação (Monografia) — Ukrainian Catholic University, Ucrânia, 2020. Citado na p. 21.