

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
DE SANTA CATARINA
CAMPUS SÃO JOSÉ

JÉSSICA GOMES CARRICO

**DETECÇÃO DE NÓDULOS EM IMAGENS MAMOGRÁFICAS
UTILIZANDO APRENDIZADO PROFUNDO E TRANSFORMADA
*WAVELET***

SÃO JOSÉ

2025

Jéssica Gomes Carrico

Detecção de nódulos em imagens mamográficas utilizando aprendizado profundo e transformada *Wavelet*

Monografia apresentada ao Curso de Engenharia de Telecomunicações do Instituto Federal de Santa Catarina, para a obtenção do título de bacharel em Engenharia de Telecomunicações.

Orientador: Prof. Ramon Mayor Martins,
Dr.

Coorientadora: Profa. Elen Macedo Lobato,
Dra.

São José

2025

Jéssica Gomes Carrico

Detecção de nódulos em imagens mamográficas utilizando aprendizado profundo e transformada *Wavelet*

Monografia apresentada ao Curso de Engenharia de Telecomunicações do Instituto Federal de Santa Catarina, para a obtenção do título de bacharel em Engenharia de Telecomunicações.

São José, 11 de dezembro de 2025.

Prof. Ramon Mayor Martins, Dr.
Instituto Federal de Santa Catarina

Professor Diego da Silva de Medeiros, Dr.
Instituto Federal de Santa Catarina

Professora Deise Monquelate Arndt , Dra.
Instituto Federal de Santa Catarina

AGRADECIMENTOS

Agradeço à minha família, que sempre esteve ao meu lado, me apoiando e incentivando a seguir em frente. Vocês são a minha base e a razão pela qual busco sempre o melhor. Agradeço especialmente à minha mãe, Janieli Aparecida Ribeiro, que sempre valorizou os estudos e me ensinou a importância do conhecimento; ao meu esposo, Ronalthy Souza Seberino, cujo apoio constante manteve tudo em equilíbrio para que eu pudesse me dedicar a esta jornada; e à minha sogra, Vanderleia de Souza, que foi um pilar essencial ao longo da minha trajetória no IFSC. Agradeço também ao meu filho, Ravi Seberino, cuja presença me motiva a seguir em frente e a buscar sempre evoluir.

Agradeço ao meu orientador, Professor Ramon Mayor Martins, que me auxiliou em toda a trajetória acadêmica, desde o ensino médio integrado ao técnico até o curso de engenharia, oferecendo orientação, apoio e incentivo em cada etapa. Agradeço também à minha coorientadora, Professora Elen Macedo Lobato, pelas contribuições, direcionamentos e disponibilidade ao longo deste trabalho.

Agradeço ao Instituto Federal de Santa Catarina, que me proporcionou uma formação sólida e de qualidade, e a todos os professores que contribuíram para o meu aprendizado.

Agradeço também aos meus colegas de curso, que compartilharam comigo momentos de aprendizado e crescimento. Juntos, enfrentamos desafios e celebramos conquistas.

RESUMO

O câncer de mama permanece como uma das principais causas de mortalidade entre mulheres no mundo, sendo a detecção precoce essencial para aumentar as chances de tratamento eficaz. Nesse contexto, este trabalho propõe o desenvolvimento e a comparação de algoritmos para detecção automática de nódulos mamários em mamografias, integrando técnicas de pré-processamento baseadas na Transformada *Wavelet* e redes neurais convolucionais (*CNNs*), treinadas tanto do zero quanto com *Transfer Learning*. As imagens utilizadas provêm da base pública *Mini-MIAS*. O desempenho dos modelos é avaliado em diferentes cenários, com e sem a aplicação da *Wavelet* e com o uso de *data augmentation*, utilizando métricas como acurácia, precisão, recall e F1-score. Os resultados mostram que o pré-processamento baseado na família *Coiflet* exerce maior impacto no desempenho do que a *data augmentation* isolada, reduzindo o desequilíbrio entre as classes e melhorando a identificação de padrões relevantes. Entre as arquiteturas avaliadas, os modelos *ResNet* apresentaram desempenho mais consistente, com destaque para a *ResNet34* combinada com o pré-processamento *Coiflet*, que obteve os melhores resultados gerais.

Palavras-chave: Câncer de mama. Redes neurais convolucionais. Transformada *Wavelet*. Reconhecimento de padrões. Inteligência artificial

ABSTRACT

Breast cancer remains one of the leading causes of mortality among women worldwide, and early detection is essential to improving the chances of effective treatment. In this context, this work proposes the development and comparison of algorithms for the automatic detection of breast nodules in mammograms, integrating preprocessing techniques based on the Wavelet Transform and convolutional neural networks (CNNs), trained both from scratch and using Transfer Learning. The images used come from the public Mini-MIAS database. Model performance is evaluated under different scenarios, with and without the application of the Wavelet transform and with the use of data augmentation, using metrics such as accuracy, precision, recall, and F1-score. The results show that preprocessing based on the Coiflet family has a greater impact on performance than data augmentation alone, reducing class imbalance and improving the identification of relevant patterns. Among the evaluated architectures, the ResNet models presented the most consistent performance, with ResNet34 combined with Coiflet-based preprocessing achieving the best overall results.

Keywords: Breast cancer. Convolutional neural networks. *Wavelet* transform. Pattern recognition. Artificial intelligence.

LISTA DE ILUSTRAÇÕES

Figura 1 – Mamografia digital demonstrando nódulo espiculado avaliado com recursos de pós-processamento	13
Figura 2 – Representação anatômica e radiológica da mama	14
Figura 3 – Classificação das densidades mamárias (A a D) segundo o sistema BI-RADS	15
Figura 4 – Sinal no domínio do tempo com mudança de frequência	16
Figura 5 – Transformada Rápida de <i>Fourier</i> do sinal: ausência de localização temporal	17
Figura 6 – Espectrograma da Transformada <i>Wavelet</i> Contínua: análise tempo-frequência	17
Figura 7 – Funções <i>wavelets</i> comuns	18
Figura 8 – Representação em diagrama de blocos da FWT	20
Figura 9 – Arquitetura de uma Rede Neural Convolutiva	22
Figura 10 – Modelo matemático de um neurônio	23
Figura 11 – Fronteira de separação do neurônio de entrada	24
Figura 12 – Topologias de Rede Neural	24
Figura 13 – Rede neural profunda	25
Figura 14 – Exemplo de convolução 2D	26
Figura 15 – Exemplo de <i>pooling</i>	26
Figura 16 – Etapas do processo de Aprendizado de Máquina Centrado no Ser Humano (HCML)	31
Figura 17 – Exemplos das primeiras imagens do conjunto Mini-MIAS	32
Figura 18 – Exemplo de comparação visual entre a imagem original e a imagem pré-processada	35
Figura 19 – Acurácia para todas as arquiteturas	50
Figura 20 – F1-score - Classe Normal	51
Figura 21 – F1-score - Classe Anormal	51
Figura 22 – Precisão - Classe Normal	52
Figura 23 – Precisão - Classe Anormal	52
Figura 24 – Recall - Classe Normal	53
Figura 25 – Recall - Classe Anormal	53
Figura 26 – F1-score - Macro	54
Figura 27 – F1-score - Ponderado	54
Figura 28 – Recall - Macro	55
Figura 29 – Recall - Ponderado	55
Figura 30 – Precisão - Macro.	56

Figura 31 – Precisão - Ponderado 56

LISTA DE TABELAS

Tabela 1 – Resultados da CNN Simples sob diferentes configurações	41
Tabela 2 – Resultados da ResNet18 sob diferentes configurações	42
Tabela 3 – Resultados da ResNet34 sob diferentes configurações	43
Tabela 4 – Resultados da ResNet50 sob diferentes configurações	44

SUMÁRIO

1	INTRODUÇÃO	10
1.1	OBJETIVOS GERAIS	11
1.2	OBJETIVOS ESPECÍFICOS	11
2	FUNDAMENTAÇÃO TEÓRICA	12
2.1	CÂNCER DE MAMA	12
2.2	MAMOGRAFIA DIGITAL	12
2.2.1	Mama	13
2.3	LIMIARIZAÇÃO	15
2.4	TRANSFORMADA <i>WAVELET</i>	16
2.5	REDES NEURAS CONVOLUCIONAIS	22
2.5.1	Neurônios (<i>Perceptrons</i>)	23
2.5.2	Rede Neural	24
2.5.3	Convolução	25
2.5.4	<i>Pooling</i> (Subamostragem)	26
2.5.5	Camada Totalmente Conectada	27
2.5.6	<i>Transfer Learning</i> (Aprendizado por transferência)	27
3	ESTADO DA ARTE	28
4	PROPOSTA	30
5	METODOLOGIA	31
5.1	COLETA DE DADOS	31
5.2	EXPERIMENTO	33
5.2.1	Pré-processamento das Imagens	35
5.2.2	<i>Data Augmentation</i>	37
5.2.3	Procedimentos de Treinamento e Teste	37
5.3	IMPLEMENTAÇÃO	39
6	RESULTADOS E DISCUSSÃO	40
7	CONCLUSÃO	45
7.1	TRABALHOS FUTUROS	45
	Referências	46
	APÊNDICE A – VISUALIZAÇÃO DOS RESULTADOS GLOBAIS	50

1 INTRODUÇÃO

O câncer de mama permanece como a principal causa de mortalidade por câncer entre mulheres em todo o mundo, sendo responsável por um grande número de óbitos em diversos países (Cadrin-Chênevert, 2023; Sechopoulos; Teuwen; Mann, 2021). Apesar dos avanços em métodos de triagem e tratamento, a detecção precoce ainda enfrenta desafios consideráveis, como a variabilidade na interpretação dos exames por profissionais de saúde e a ocorrência de falsos positivos, que podem gerar ansiedade e procedimentos desnecessários para as pacientes.

Nesse cenário, a inteligência artificial (IA) tem se mostrado uma ferramenta promissora no apoio ao diagnóstico do câncer de mama. A utilização de redes neurais convolucionais (CNN), por exemplo, para a detecção de anomalias de forma automática, juntamente com uma avaliação médica, contribui para evitar procedimentos desnecessários, minimizar o impacto psicológico nas pacientes e otimizar os recursos do sistema de saúde (Sechopoulos; Teuwen; Mann, 2021; Cadrin-Chênevert, 2023).

Para que seja possível utilizar redes neurais na detecção de anormalidades em um exame de mama, é necessário treiná-las previamente. Esse processo envolve fornecer uma grande quantidade de imagens já classificadas, permitindo que a rede aprenda a reconhecer padrões específicos associados. O intuito é de que após esse treinamento, a rede seja capaz de analisar uma nova imagem e fazer uma previsão com base no conhecimento adquirido. Entretanto, treinar redes neurais do zero exige muitos dados rotulados e alto poder computacional. Para contornar essas limitações, utiliza-se o *Transfer Learning*, que reaproveita redes já treinadas em grandes bases, permitindo bons resultados mesmo com poucos dados e infraestrutura reduzida (Isin; Ozdalili, 2017).

Além disso, as imagens precisam passar por um pré-processamento. Essa etapa é necessária para que as imagens estejam em um formato e qualidade que facilite a extração de informações relevantes pelos algoritmos. Uma das técnicas utilizadas nesse processo é a Transformada *Wavelet*, que permite decompor a imagem em níveis de resoluções diferentes (Leite; Henriques; Cunha, 2018). Isso facilita a identificação de detalhes sutis, como microcalcificações e bordas de nódulos, ao mesmo tempo em que reduz ruídos indesejados. Dessa forma, o uso da *Wavelet* contribui para tornar os dados mais adequados ao aprendizado da rede neural, aumentando a eficiência na detecção de padrões associados ao câncer de mama.

Diante desse cenário, esta pesquisa tem como objetivo treinar tanto uma arquitetura convolucional simples, desenvolvida do zero, quanto modelos baseados em *Transfer Learning* para a detecção de anomalias em imagens de mamografia utilizando uma base de

dados pública. Este trabalho de conclusão de curso compara o desempenho dessas abordagens com e sem a aplicação da Transformada *Wavelet* como técnica de pré-processamento, além de avaliar o impacto do uso de *data augmentation*, entendida como o conjunto de transformações aplicadas às imagens (como rotações, espelhamentos e variações de escala) com o objetivo de aumentar a diversidade dos dados disponíveis para treinamento. A proposta é verificar se a aplicação da *Wavelet* e das técnicas de aumento de dados contribui para melhorar a acurácia dos modelos, facilitando a identificação de padrões relevantes e auxiliando na detecção precoce do câncer de mama.

1.1 OBJETIVOS GERAIS

Desenvolver e comparar algoritmos para detecção automática de nódulos mamários em mamografias, integrando a decomposição *Wavelet* multinível como técnica de pré-processamento e diferentes arquiteturas de redes neurais convolucionais, incluindo modelos treinados do zero e com *Transfer Learning*, implementados em Python. O objetivo é investigar se a aplicação da *Wavelet* e do *data augmentation* contribui para elevar a precisão diagnóstica, avaliando o impacto de cada estratégia no desempenho das redes.

1.2 OBJETIVOS ESPECÍFICOS

- Selecionar e pré-processar imagens de mamografia provenientes da base pública *Mini-MIAS*.
- Implementar um algoritmo de decomposição *Wavelet* para destacar padrões relevantes nas imagens e reduzir ruídos.
- Aplicar técnicas de *data augmentation*, compreendidas como transformações geométricas e fotométricas que aumentam a variabilidade das imagens disponíveis para o treinamento.
- Treinar redes neurais convolucionais (CNNs) nas arquiteturas ResNet-18, ResNet-34 e ResNet-50 com o uso de *Transfer Learning*.
- Avaliar o desempenho do modelo em diferentes cenários de entrada (com e sem *Wavelet*, com e sem *data augmentation*), utilizando métricas como acurácia, precisão, recall e F1-score.
- Comparar os resultados obtidos para verificar o impacto das estratégias de pré-processamento e aumento de dados na detecção automática de nódulos mamários.

2 FUNDAMENTAÇÃO TEÓRICA

A proposta desta pesquisa, é realizar a detecção de nódulos mamários em mamografias por meio de uma rede neural convolucional. Nesse processo, as imagens passam inicialmente por uma etapa de pré-processamento com a Transformada *Wavelet*, aplicada para remoção de ruído (*denoising*), antes de serem processadas pela CNN. Assim, nessa seção é explanado os principais tópicos para alcançar os objetivos pretendidos.

2.1 CÂNCER DE MAMA

Esse tipo de câncer pode se desenvolver a partir de alterações nas células dos lóbulos e ductos mamários. Inicialmente, essas alterações podem incluir um crescimento exagerado de células (hiperplasia) ou um padrão anormal de crescimento (hiperplasia atípica). Com o tempo, podem evoluir para uma forma localizada da doença, que ainda não se espalhou para os tecidos vizinhos (carcinoma in situ), e depois para uma forma mais agressiva, que invade os tecidos ao redor (carcinoma invasivo). Um dos sinais visíveis da doença é a pele da mama ficar parecida com uma casca de laranja, por causa do inchaço. Além disso, também é comum sentir caroços (linfonodos aumentados) na axila (Bravo et al., 2021).

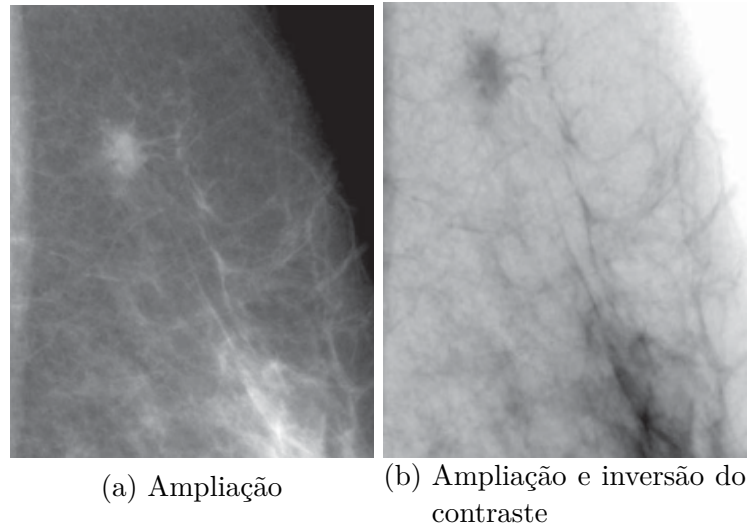
O câncer de mama é a principal causa de mortalidade por câncer entre mulheres em todo o mundo, com sua incidência influenciada por fatores genéticos, hormonais e de estilo de vida. A detecção precoce da doença é fundamental para aumentar as chances de cura e reduzir a mortalidade, sendo geralmente realizada por meio de exames de imagem, como a mamografia, que permite identificar alterações suspeitas ainda em estágios iniciais (Cadrin-Chênevert, 2023; Sechopoulos; Teuwen; Mann, 2021). Entre os elementos observados nas mamografias, destacam-se o tamanho do tumor e o envolvimento dos gânglios axilares, fatores que auxiliam na definição da conduta clínica, como a necessidade de quimioterapia ou cirurgia de retirada (Bravo et al., 2021). Além disso, a forma e o contorno dos nódulos também são considerados: nódulos com formato oval, bordas bem definidas e aspecto regular tendem a ser benignos, enquanto alterações irregulares podem indicar malignidade (Passig e Silva, 2022)

2.2 MAMOGRAFIA DIGITAL

A mamografia digital é uma técnica avançada de diagnóstico por imagem utilizada para detectar alterações no tecido mamário, com ênfase na identificação precoce do câncer de mama. Diferente da mamografia convencional, que utiliza filmes radiográficos, a versão digital emprega detectores eletrônicos para capturar a radiação transmitida pela mama,

convertendo-a em sinais digitais. Esses sinais são processados por algoritmos computacionais, permitindo ajustes precisos de brilho, contraste e ampliação da imagem (Figura 1), sem a necessidade de nova exposição à radiação (Freitas et al., 2006).

Figura 1 – Mamografia digital demonstrando nódulo espiculado avaliado com recursos de pós-processamento

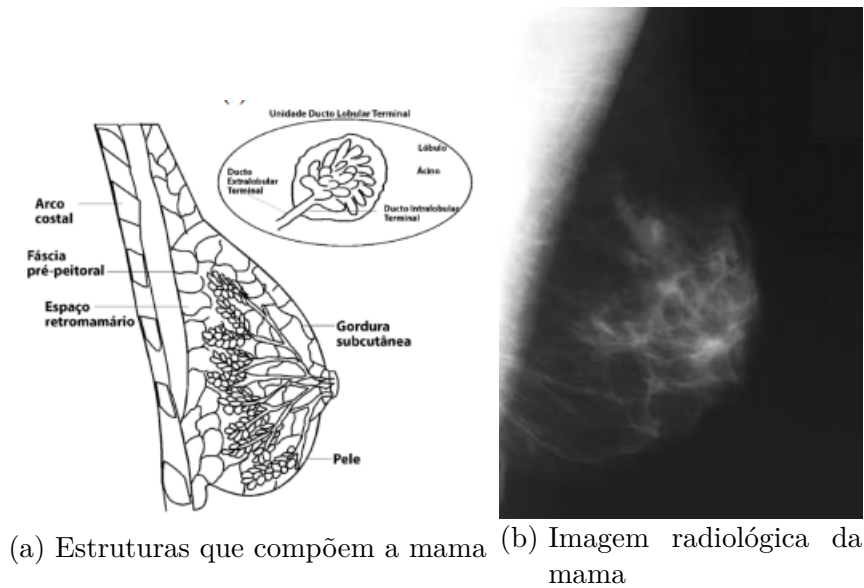


Fonte: (Freitas et al., 2006).

2.2.1 Mama

A mama é formada por diferentes tipos de tecido, como o tecido gorduroso (adiposo), o tecido fibroglandular e os ductos que transportam o leite. Nas imagens de mamografia, o tecido gorduroso aparece mais claro, enquanto o tecido fibroglandular aparece mais escuro (observa-se em Figura 2), o que ajuda a identificar diferentes áreas da mama. A quantidade desses tecidos varia de mulher para mulher, o que pode influenciar na visualização das imagens e na detecção de possíveis alterações (Sousa, 2017).

Figura 2 – Representação anatômica e radiológica da mama



Fonte: (Sousa, 2017).

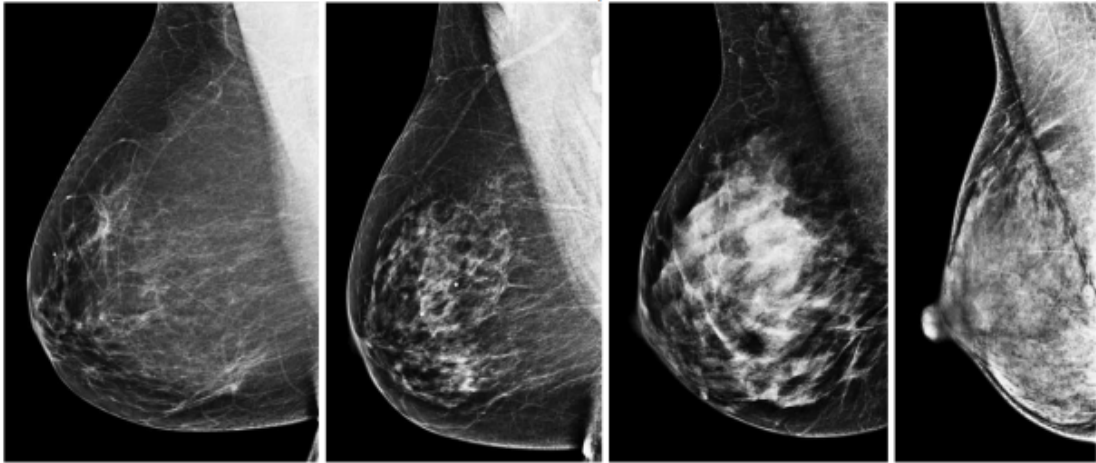
Devido à variação de densidade mamária entre as pacientes, o sistema BI-RADS (*Breast Imaging Reporting and Data System*) propôs uma classificação padronizada que descreve o grau de densidade observado nas imagens mamográficas. Essa categorização auxilia radiologistas na avaliação da qualidade da imagem e na estimativa da possibilidade de lesões não serem visualizadas, contribuindo diretamente para a conduta clínica e o acompanhamento da paciente.

O BI-RADS classifica a densidade mamária em quatro categorias:

- **A:** mama composta quase inteiramente por gordura;
- **B:** presença de áreas dispersas de tecido fibroglandular;
- **C:** mama heterogeneamente densa, podendo ocultar pequenas lesões;
- **D:** mama extremamente densa, com maior risco de mascaramento de alterações.

Na Figura 3, observa-se que mamas com maior proporção de tecido gorduroso tendem a facilitar a identificação de nódulos, uma vez que o contraste entre estruturas é mais evidente. Por outro lado, em mamas densas, a sobreposição de tecidos pode dificultar a detecção de alterações sutis, o que representa um desafio diagnóstico em exames mamográficos (Sousa, 2017).

Figura 3 – Classificação das densidades mamárias (A a D) segundo o sistema BI-RADS



Fonte: (Sousa, 2017).

2.3 LIMIAZIZAÇÃO

Uma das técnicas de processamento de imagem utilizada para o tratamento de imagens mamográficas é a limiarização, que consiste em aplicar um valor limite (limiar) para separar os pixels importantes do fundo em uma imagem. Pixels com intensidade igual ou superior ao limiar recebem valor 1, e os demais, 0, gerando uma imagem binária. Multiplicando essa imagem binária pelos valores originais, mantém-se apenas os pixels relevantes, eliminando o fundo (Passig e Silva, 2022). A seguir, apresenta-se um exemplo prático de limiarização:

$$I = \begin{bmatrix} 10 & 50 & 200 \\ 30 & 100 & 180 \\ 20 & 40 & 220 \end{bmatrix}, \quad T = 100 \quad (2.1)$$

$$B(x, y) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B(x, y) = \begin{cases} 1, & I(x, y) \geq T \\ 0, & \text{caso contrário} \end{cases} \quad (2.2)$$

$$I'(x, y) = I(x, y) \times B(x, y) = \begin{bmatrix} 0 & 0 & 200 \\ 0 & 100 & 180 \\ 0 & 0 & 220 \end{bmatrix} \quad (2.3)$$

onde I é a matriz da imagem original contendo os valores de intensidade dos pixels, T é o limiar definido para separar o fundo dos objetos de interesse, B é a imagem binária resultante da aplicação do limiar, e I' é a imagem final obtida após a multiplicação elemento a elemento, que mantém apenas os pixels com intensidade igual ou superior a T .

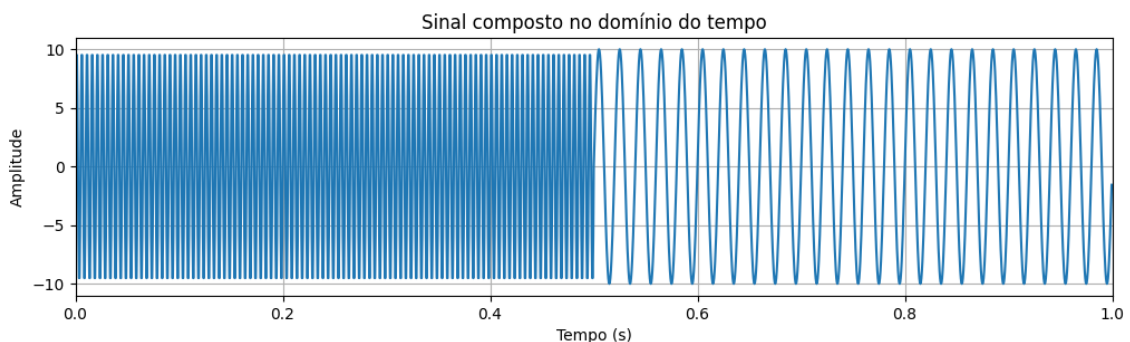
Embora uma imagem não tenha variação temporal (como um sinal de áudio, por exemplo), ela possui variações espaciais na intensidade dos pixels que podem ser interpretadas em termos de frequência. Áreas com transições suaves correspondem a baixas frequências, enquanto detalhes, bordas ou texturas correspondem a altas frequências. Por isso, técnicas como a Transformada Wavelet (seção 2.4) podem ser aplicadas a imagens para decompor essas variações em diferentes escalas, permitindo destacar padrões importantes e reduzir ruídos, o que facilita a extração de informações relevantes para etapas posteriores de processamento.

2.4 TRANSFORMADA WAVELET

A Transformada *Wavelet* surge como uma solução para uma limitação da Transformada de Fourier. Ambas têm como objetivo representar um sinal no domínio do tempo e no domínio da frequência. No entanto, a Transformada de *Fourier* não permite identificar em que momento uma determinada frequência ocorre, fazendo com que a sua aplicação torne-se limitada na análise de sinais com variações ao longo do tempo (não estacionários) (Domingues et al., 2016).

A Figura 4 ilustra um sinal no domínio do tempo, composto por uma frequência de 200 Hz no intervalo entre 0 e 0.5 segundos, e por uma frequência de 50 Hz a partir de 0.5 segundos.

Figura 4 – Sinal no domínio do tempo com mudança de frequência

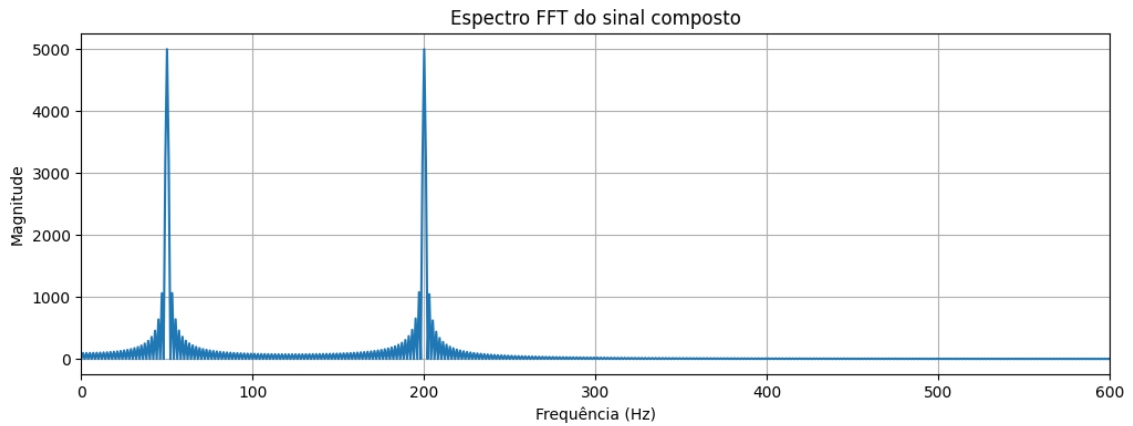


Fonte: Elaborada pela autora.

Ao aplicar a Transformada Rápida de *Fourier* ao sinal, obtém-se a representação no domínio da frequência (Figura 5). Embora ela revele as frequências presentes, não indica em que momento essas frequências ocorrem no sinal.

Para contornar essa limitação e extrair informações simultaneamente no tempo e na frequência, é necessário adotar uma abordagem diferente, como o janelamento do sinal (Domingues et al., 2016), o que leva ao uso da Transformada *Wavelet*. Embora sua fundamentação seja semelhante à da Transformada Rápida de *Fourier* (FFT), a Transformada *Wavelet* difere ao decompor o sinal não em senóides, mas em funções *wavelets* (ondas de

Figura 5 – Transformada Rápida de *Fourier* do sinal: ausência de localização temporal

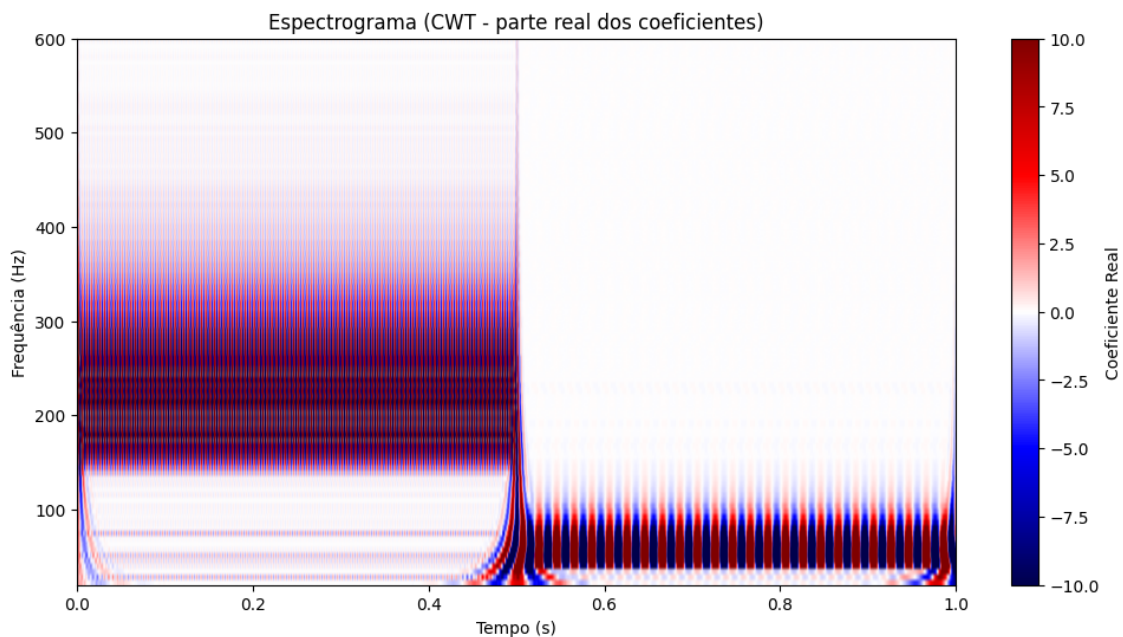


Fonte: Elaborada pela autora.

curta duração que se atenuam rapidamente). Isso permite capturar eventos transitórios e mudanças abruptas com maior precisão. As *wavelets* são funções localizadas tanto no tempo, por meio de translações, quanto na frequência ou escala, por meio de dilatações, possibilitando uma análise mais refinada de sinais não estacionários.

A Figura 6 mostra o espectrograma gerado pela aplicação da Transformada *Wavelet* Contínua (CWT, do inglês *Continuous Wavelet Transform*) ao sinal. É possível observar não apenas as frequências presentes, mas também os momentos em que elas ocorrem, evidenciando a vantagem da análise conjunta no tempo e na frequência.

Figura 6 – Espectrograma da Transformada *Wavelet* Contínua: análise tempo-frequência



Fonte: Elaborada pela autora.

A Transformada *Wavelet* Contínua de uma função $x(t) \in L^2(\mathbb{R})$ pode ser definida

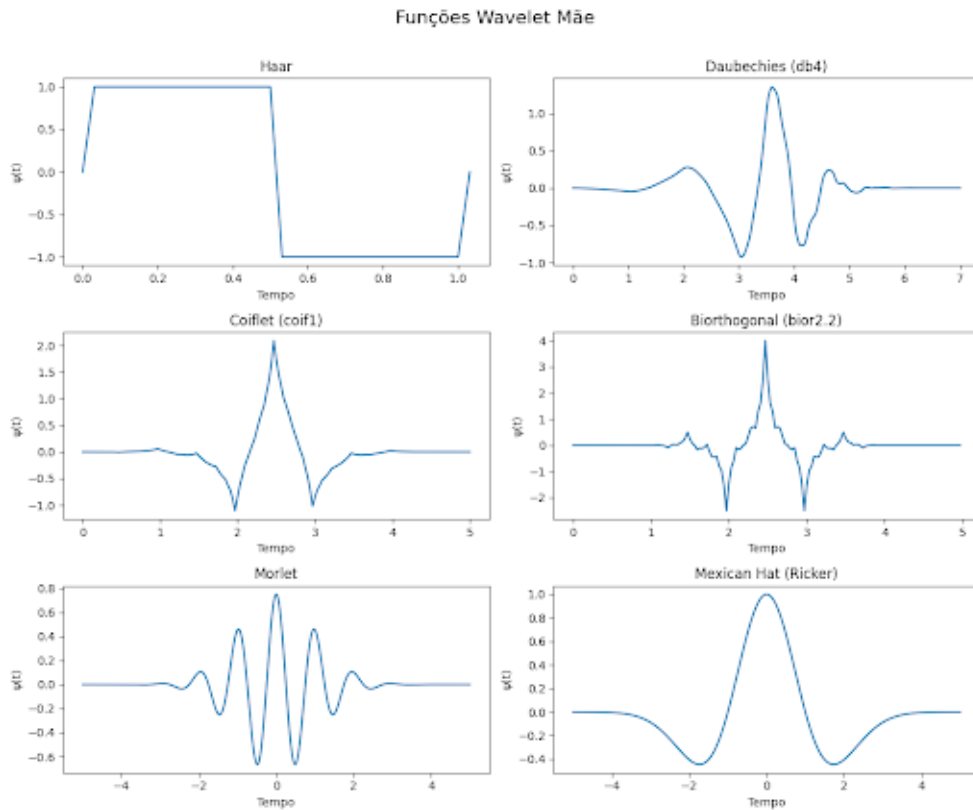
conforme a Equação 2.4. Nessa expressão, o parâmetro a representa o *fator de dilatação* (ou escala), b é o *deslocamento temporal* (ou translação), e ψ é a *wavelet mãe* (Goswami; Chan, 1999).

$$W_{\psi,x}(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.4)$$

Wavelets mães

Na Transformada *Wavelet*, as funções *wavelets* são obtidas a partir de uma função geradora, sendo denominada como *wavelet* mãe. As *Wavelets* mães são sinais que ao serem deslocados, multiplicados, ou outras transformações, formam uma família (Passig e Silva, 2022). Dentre as famílias de *wavelets* mais conhecidas estão a *Coiffet*, *Daubechies*, *Haar*, *Biorthogonal* e *Morlet* e *Mexican Hat* (Chapéu Mexicano), que são ilustradas na Figura 7

Figura 7 – Funções *wavelets* comuns



Fonte: Elaborada pela autora.

Tem-se que a *wavelet* mãe apresentada na Equação 2.4 deve satisfazer duas condições essenciais (Magrini, 2020):

- Ter valor médio nulo (Equação 2.5);
- Possuir energia finita (Equação 2.6).

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2.5)$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (2.6)$$

Transformada *Wavelet* Discreta

A Transformada *Wavelet* Discreta (DWT, do inglês *Discrete Wavelet Transform*) é construída a partir da discretização dos parâmetros de escala e translação utilizados na Transformada *Wavelet* Contínua (CWT). Essa discretização permite reescrever a integral da CWT como uma soma discreta (Goswami; Chan, 1999). A Equação 2.7 descreve esta discretização do sinal:

$$f[n] = \sum_{k=-\infty}^{\infty} c_{j,k} \phi_{j,k}[n] + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k}[n] \quad (2.7)$$

O sinal $f[n]$ pode ser decomposto a partir de duas funções base fundamentais: as funções escalares $\phi_{j,k}[n]$, que representam as componentes de aproximação associadas às baixas frequências, e as funções *wavelet* mãe $\psi_{j,k}[n]$, que representam as componentes de detalhe associadas às altas frequências (Passig e Silva, 2022). Essa decomposição depende diretamente dos parâmetros j e k . O parâmetro $j \in \mathbb{Z}$ representa o *nível de escala*: valores mais altos de j produzem *wavelets* mais dilatadas, capazes de capturar tendências globais (baixas frequências), enquanto valores menores de j geram *wavelets* mais comprimidas, adequadas para identificar variações rápidas (altas frequências). Já o parâmetro $k \in \mathbb{Z}$ determina a *posição temporal* da análise, controlando onde cada função base é aplicada ao sinal; ao variar k , varre-se o sinal no tempo para localizar eventos locais.

Por fim, essa representação é quantificada pelos coeficientes associados a cada função base: $c_{j,k}$ corresponde aos coeficientes de aproximação no nível mais baixo (mais grosseiro), enquanto $d_{j,k}$ representa os coeficientes de detalhe associados a cada nível j .

Cálculo de $\phi_{j,k}[n]$ e $\psi_{j,k}[n]$

As funções $\phi_{j,k}[n]$ e $\psi_{j,k}[n]$ são versões escaladas e transladadas das funções $\phi[n]$ (função escalar) e $\psi[n]$ (função *wavelet* mãe) (Passig e Silva, 2022), definidas da seguinte forma:

$$\phi_{j,k}[n] = a_0^{-j/2} \cdot \phi\left(\frac{n - kb_0 a_0^j}{a_0^j}\right) \quad (2.8)$$

$$\psi_{j,k}[n] = a_0^{-j/2} \cdot \psi\left(\frac{n - kb_0 a_0^j}{a_0^j}\right) \quad (2.9)$$

onde fator $a_0^{-j/2}$ garante a normalização da energia das funções em diferentes escalas (Goswami; Chan, 1999). O termo $a_0^j n$ realiza a dilatação (controlando a frequência), sendo o equivalente ao a da CWT, enquanto $kb_0 a_0^j$ define a translação no tempo (controlando a posição da análise), equivalente ao b da CWT. Deste modo, a DWT é formalmente definida como Equação 2.10, o que permite a decomposição do sinal em níveis de resolução, separando componentes de baixa frequência (aproximação) e de alta frequência (detalhe).

$$\begin{aligned} W_f(j, k) &= \sum_{n=-\infty}^{\infty} f[n] \cdot \psi_{j,k}[n] \\ &= \sum_{n=-\infty}^{\infty} f[n] \cdot a_0^{-j/2} \psi\left(\frac{n - kb_0 a_0^j}{a_0^j}\right) \quad ; \quad a_0 > 1; b_0 \neq 0. \end{aligned} \quad (2.10)$$

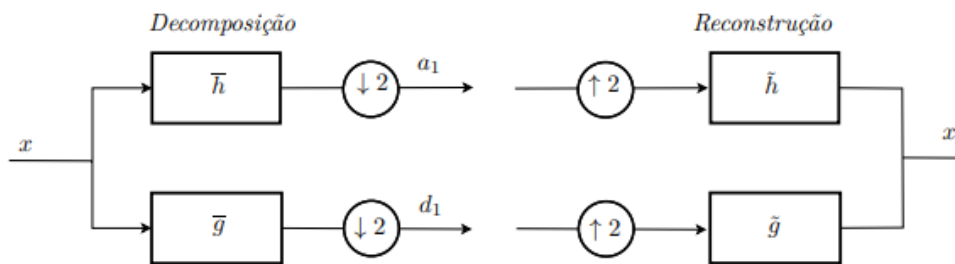
Banco de filtros

A Transformada *Wavelet* Discreta pode ser implementada de maneira eficiente por meio de um banco de filtros digital, estrutura conhecida como *Fast Wavelet Transform* (FWT). Esse método evita a construção explícita das funções $\phi_{j,k}[n]$ e $\psi_{j,k}[n]$, ao calcular diretamente os coeficientes de aproximação e de detalhe a partir de operações de filtragem e subamostragem.

O procedimento consiste em aplicar dois filtros ao sinal: um filtro passa-baixa (\bar{h}), que preserva as componentes de baixa frequência, e um filtro passa-alta (\bar{g}), que isola as componentes de alta frequência. Após a filtragem, realiza-se a subamostragem (redução pela metade da taxa de amostragem), o que resulta nos coeficientes de aproximação e de detalhe para aquele nível de análise (Magrini, 2020).

Esse processo pode ser repetido recursivamente sobre os coeficientes de aproximação, permitindo a decomposição do sinal em múltiplos níveis de resolução. A estrutura do banco de filtros, composta pelas etapas de decomposição e reconstrução, é ilustrada na Figura 8.

Figura 8 – Representação em diagrama de blocos da FWT



Fonte: (Magrini, 2020).

Essa abordagem é amplamente utilizada em aplicações reais, como compressão de sinais, análise de transientes e remoção de ruído, por ser computacionalmente eficiente.

Wavelet denoising

O processo de remoção de ruídos utilizando a Transformada *Wavelet* consiste basicamente em três etapas principais (Passig e Silva, 2022):

1. **Decomposição:** O sinal original x é decomposto em coeficientes *wavelet* C através da transformada wavelet, conforme a Equação (2.11):

$$C = W(x) \quad (2.11)$$

2. **Limiarização:** Aplica-se um operador $D(C, \alpha)$, que realiza a remoção do ruído por meio da limiarização dos coeficientes, onde α é o limiar definido para distinguir ruído de sinal significativo, conforme a Equação (2.12):

$$C_d = D(C, \alpha) \quad (2.12)$$

Segundo Mupparaju e Jahnavi (2013), podemos definir o limiar universal (ou VisuShrink) α como:

$$\alpha = \sigma \sqrt{2 \log P} \quad (2.13)$$

onde σ é a variância do ruído e P é o número total de pixels da imagem. Apesar de simples, esse método tende a produzir uma estimativa excessivamente suavizada quando P é grande.

Além disso, a limiarização pode ser aplicada de duas formas:

- **Hard thresholding (Limiar rígido):**

$$T_{hard}(d, \lambda) = d \cdot \mathbf{1}(|d| > \lambda) \quad (2.14)$$

- **Soft thresholding (Limiar suave):**

$$T_{soft}(d, \lambda) = \text{sign}(d) \cdot (|d| - \lambda)_+ \quad (2.15)$$

onde d representa o coeficiente de detalhe do sinal e λ o limiar aplicado.

3. **Reconstrução:** Por fim, a transformada *wavelet* inversa é aplicada aos coeficientes limiarizados para obter o sinal *denoised* $s(t)$, conforme a Equação (2.16):

$$s(t) = W^{-1}(C_d) \quad (2.16)$$

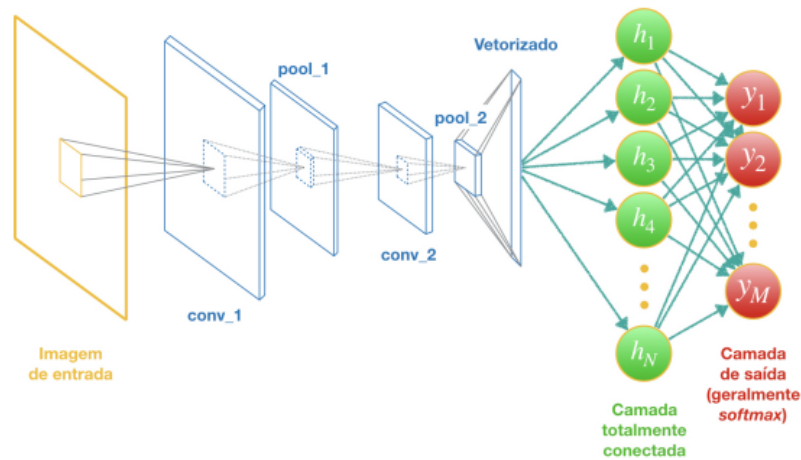
O nível de decomposição, ou seja, a quantidade de vezes que o sinal é subdividido em componentes de alta e baixa frequência, também influencia o resultado do *denoising*, podendo afetar a preservação do sinal original.

2.5 REDES NEURAS CONVOLUCIONAIS

As Redes Neurais Convolucionais (CNNs, do inglês *Convolutional Neural Networks*) constituem uma arquitetura específica de redes neurais profunda que incorpora a operação matemática de convolução em pelo menos uma de suas camadas (Rosa, 2025). Essa abordagem favorece a extração eficiente de padrões espaciais e estruturais dos dados de entrada (Miyazaki, 2017). Ao abordar imagens como objetos de entrada, (que podem ser compreendidas como uma grade bidimensional de pixels) as CNNs se demonstram adequadas (Cunha, 2020).

De forma geral, a arquitetura de uma *CNN* segue um fluxo bem definido: os dados de entrada (como uma imagem) passam por camadas convolucionais, que extraem características locais; em seguida, por camadas de *pooling*, que reduzem a dimensionalidade; depois, por camadas totalmente conectadas (*fully connected*), responsáveis pela tomada de decisão com base nas características extraídas; até, finalmente, gerar a saída (Miyazaki, 2017). A Figura 9 ilustra esse processo de forma simplificada.

Figura 9 – Arquitetura de uma Rede Neural Convolucional

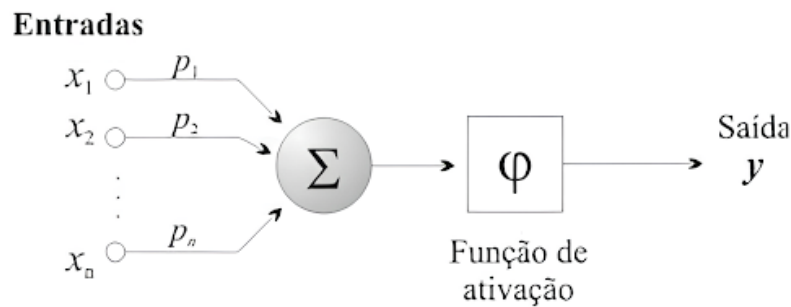


Fonte: (Silva Junior, 2022).

2.5.1 Neurônios (*Perceptrons*)

Para compreendermos o funcionamento das redes neurais profunda, é essencial entender primeiro o que faz um único neurônio (perceptrons). Sua função é receber uma ou mais entradas, calcular uma combinação linear ponderada dessas entradas com base em pesos associados e, em seguida, aplicar uma função de ativação para gerar uma saída, conforme mostra a Figura 10. Geometricamente, esse processo define um hiperplano no espaço das entradas, que serve para separar os dados conforme suas classes ou características.

Figura 10 – Modelo matemático de um neurônio



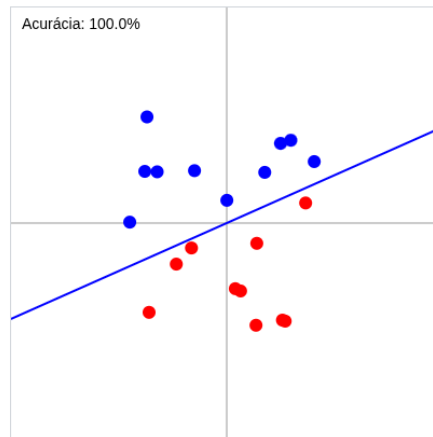
Fonte: (Ferneda, 2006).

Na Figura 11, os pesos atribuídos são $w_1 = -0,40$, $w_2 = 0,90$, e o *bias* é $b = 0,00$. A equação resultante do hiperplano é:

$$-0,40x + 0,90y + 0,00 = 0$$

Os pontos azuis representam a classe positiva, enquanto os vermelhos indicam a classe negativa. Essa visualização ilustra como os pesos e o bias afetam diretamente a fronteira de decisão.

Figura 11 – Fronteira de separação do neurônio de entrada















Fonte: (Martins, 2025).

Em redes neurais com múltiplos neurônios, cada um define seu próprio hiperplano. A combinação desses hiperplanos permite que a rede realize decisões mais complexas, separando dados de maneira mais precisa mesmo em casos em que as classes não são linearmente separáveis.

2.5.2 Rede Neural

Chamamos de rede neural quando há mais de um neurônio trabalhando em conjunto para realizar decisões. Existem diversas topologias possíveis, como por exemplo: *single-layer* (camada única), *two-layer* (duas camadas), *three-layer* (três camadas), entre outras configurações, dependendo da complexidade da tarefa a ser resolvida, conforme observa-se na Figura 12.

Figura 12 – Topologias de Rede Neural

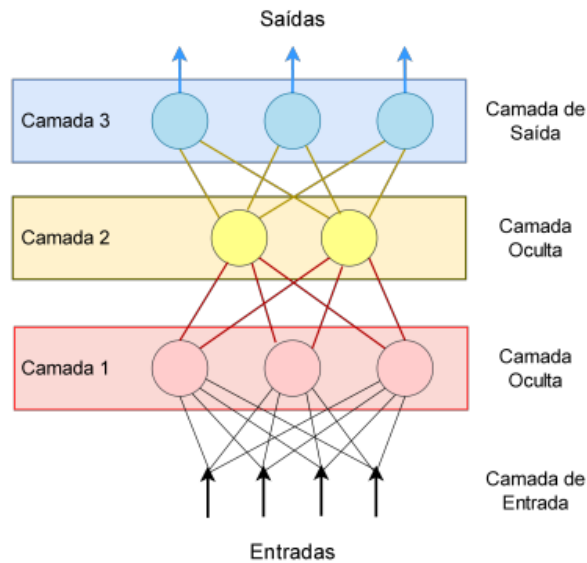
	Tipo de Regiões de Decisão	Problema OU Exclusivo	Classes com Regiões Entrelaçadas	Formas de Regiões Mais Gerais
Camada Única 	Semi-plano Delimitado por Hiperplano			
Duas Camadas 	Regiões Convexas Abertas ou Fechadas			
Multicamadas 	Arbitrárias (Complexidade Limitada pelo Número de Nós)			

Fonte: Adaptada de Very Technology.

Rede Neural Profunda

Quando se forma uma estrutura em que os neurônios são organizados em múltiplas camadas, tem-se uma rede neural profunda (DNN, do inglês *Deep Neural Network*), como mostrado na Figura 13. Esse tipo de rede permite o processamento hierárquico dos dados, em que cada camada utiliza as saídas da anterior para extrair informações mais complexas.

Figura 13 – Rede neural profunda

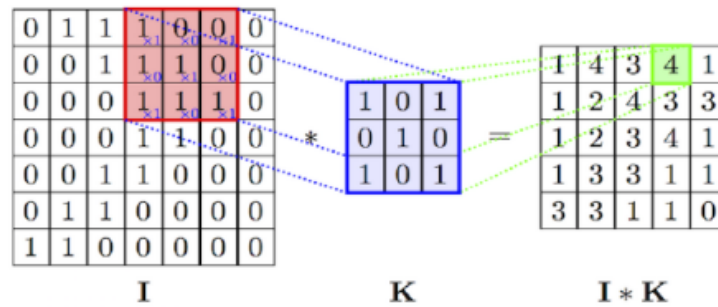


Fonte: (Rosa, 2025).

2.5.3 Convolução

A principal camada de uma Rede Neural Convolutacional é a camada de Convolução. No contexto de uma convolução 2D (aplicada a imagens não coloridas), a imagem de entrada é representada como uma matriz de pixels e processada por um pequeno filtro, conhecido como *kernel*. Esse *kernel* é deslocado sobre a imagem, sendo posicionado em regiões sucessivas. Em cada posição, realiza-se a multiplicação elemento a elemento entre os valores do *kernel* e os da região correspondente da imagem, seguida pela soma desses produtos para gerar um único valor. Cada valor obtido compõe um ponto de uma nova matriz chamada mapa de características (*feature map*) (Rosa, 2025). Como ilustrado na Figura 14.

Figura 14 – Exemplo de convolução 2D



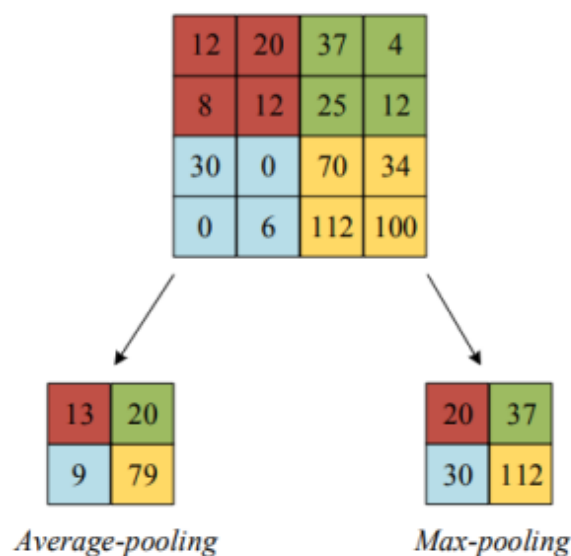
Fonte: (Rosa, 2025).

2.5.4 Pooling (Subamostragem)

Outra camada importante em uma *CNN* é a camada Pooling. Dois tipos comuns de operações de agrupamento (*pooling*) em redes neurais convolucionais são o *Max Pooling* e o *Average Pooling* (Figura 15). Ambas as técnicas atuam sobre regiões locais da imagem de entrada, definidas por um *kernel* deslizante.

O *Max Pooling* retorna, para cada região coberta pelo *kernel*, o valor máximo entre os elementos daquela vizinhança. Essa abordagem é a mais frequentemente utilizada, pois preserva os recursos mais significativos detectados durante a convolução. Além disso, contribui para a supressão de ruídos, descartando ativações de baixa intensidade.

Já o *Average Pooling*, por outro lado, calcula a média dos valores contidos na região coberta pelo *kernel*. Embora também reduza a dimensionalidade, tende a suavizar a representação da imagem, o que pode levar à perda de informações mais marcantes.

Figura 15 – Exemplo de *pooling*

Fonte: (Rosa, 2025).

2.5.5 Camada Totalmente Conectada

Após passarem por diversas camadas de convolução e pooling, os *feature maps* são progressivamente reduzidos até serem transformados em uma representação unidimensional. Esse vetor resultante é então utilizado como entrada para uma camada totalmente conectada, na qual cada neurônio se conecta a todos os neurônios da camada anterior. Essa etapa permite à rede combinar as características extraídas e identificar padrões de maior complexidade (Rosa, 2025).

2.5.6 *Transfer Learning* (Aprendizado por transferência)

O *Transfer Learning* tem se destacado como uma abordagem eficaz para reduzir o tempo e os recursos computacionais necessários no treinamento de grandes modelos de aprendizado profundo. Essa técnica consiste em reutilizar modelos previamente treinados em grandes bases de dados, para extrair características relevantes mesmo em cenários com poucos dados rotulados (Isin; Ozdalili, 2017).

A transferência de conhecimento aproveita os vetores de características gerados por redes neurais profundas já treinadas, utilizando seus pesos sinápticos configurados para reconhecer e extrair informações de conjuntos de dados anteriores. Esses vetores servem como descritores das imagens do novo conjunto e podem ser usados para treinar um novo classificador ou inseridos diretamente nas camadas totalmente conectadas da arquitetura original, após ajustes finos (*fine-tuning*) nas camadas convolucionais (Aguiar Neto, 2017). Assim, essa estratégia tem se mostrado especialmente vantajosa em áreas como a Medicina, onde a disponibilidade de grandes volumes de dados rotulados é limitada.

3 ESTADO DA ARTE

A integração entre redes neurais convolucionais (CNNs) e transformadas *Wavelet* tem demonstrado eficácia significativa na análise de mamografias, especialmente para identificação de lesões sutis como microcalcificações e distorções arquiteturais (Blahová; Kostolný; Cimrák, 2025; Oyelade; Ezugwu, 2022). Essa abordagem extrai características multiescala que amplificam padrões relevantes para o diagnóstico precoce do câncer de mama (Oyelade; Ezugwu, 2022).

Técnicas tradicionais de aumento de dados, como rotação e espelhamento, equilibram classes raras (ex.: distorções arquiteturais), reduzindo falsos positivos em até 12% e melhorando a sensibilidade para microcalcificações em 9,3%. Essa estratégia mitiga desequilíbrios amostrais e aprimora a generalização dos modelos (Blahová; Kostolný; Cimrák, 2025)

Com a base de dados *Mini-MIAS*, a aplicação de *Wavelets Haar* como pré-processamento em arquiteturas *ResNet/VGG* alcançou 98,5% de acurácia na detecção de nódulos malignos, isolando padrões de alta frequência em bordas irregulares (Rasheed et al., 2021). Complementarmente, redes híbridas que combinam *Wavelet* e *CNN* melhoraram a discriminação de microcalcificações em tecidos densos, atingindo acurácia de 85% com *EfficientNet*, 80,9% com *ResNet101* e 83,4% com *AmoebaNet-C* (Banerjee; Kabir, 2024).

Com a base de dados *CBIS-DDSM*, modelos *Wavelet-CNN* atingiram 87,2% de acurácia na identificação de distorções arquiteturais e microcalcificações, com sensibilidade de 85,4% (Oyelade; Ezugwu, 2022). Em comparação com abordagens anteriores, a superioridade da integração *Wavelet-CNN* é evidenciada em comparações diretas:

Oyelade e Ezugwu (2022) superaram métodos como CNN-DW (Jadoon et al. (2017): 81,83% de acurácia) e *CNN* tradicional (Teare et al. (2017): 85–88%), alcançando 87,2% de acurácia e AUC 0,96 no *CBIS-DDSM*.

Embora Bakkour e Afdel (2017) tenham reportado 97,28% de acurácia com aumento de dados, sua abordagem não incorporou transformadas *Wavelet*, limitando a extração de características multiescala (Oyelade; Ezugwu, 2022).

No contexto nacional e institucional, destaca-se a produção desenvolvida no Instituto Federal de Santa Catarina (IFSC) – Campus São José, no curso de Tecnologia em Sistemas de Telecomunicações, em especial o estudo “*Deteção de Tumores em Mamografia*”, de Silva (2013), que investigou técnicas de processamento digital de imagens aplicadas à deteção de tumores mamográficos, obtendo 82% de eficiência na identificação e localização dos achados nas imagens avaliadas. Os resultados evidenciaram a influência direta do pré-processamento na qualidade da deteção, com a necessidade de equaliza-

ção de brilho e contraste e limitações impostas pelo uso de filtros com dimensões fixas, especialmente diante da variabilidade de tamanho e posicionamento das mamas. Em comparação com abordagens baseadas em análise multiescala, o trabalho indica que a ausência da transformada *Wavelet* pode ter contribuído para uma filtragem menos eficaz, reforçando o potencial dessa técnica para a remoção de ruídos e o realce de estruturas relevantes em imagens mamográficas.

Posteriormente, o Trabalho de Conclusão de Curso “*Realce de Achados em Mamografia*”, de Oliveira (2019), desenvolvido no IFSC – Campus São José, no curso de Tecnologia em Sistemas de Telecomunicações, investigou técnicas de processamento digital de imagens aplicadas ao realce de achados em mamografias. Os resultados indicaram que a utilização da transformada *Wavelet* apresentou desempenho satisfatório para o realce das imagens, sendo recomendada para detecção de bordas e remoção de ruídos. Como direcionamento para trabalhos futuros, o autor também sugeriu a integração do pré-processamento diretamente no *script* principal.

Mais recentemente, no âmbito do curso de Engenharia de Telecomunicações do IFSC – Campus São José, destaca-se o Trabalho de Conclusão de Curso “*Deteção de nódulos em imagens de mamografia usando wavelet e inteligência artificial*”, de Passig e Silva (2023). O estudo avaliou o impacto do pré-processamento de mamografias por meio da remoção de ruído via transformadas *Wavelet* na classificação com redes neurais convolucionais, identificando melhora significativa de acurácia com a família Biorthogonal, além de apontar as famílias *Biorthogonal* e *Coiflet* como promissoras para aplicações em mamografias. Como sugestão de trabalhos futuros, a autora recomendou a investigação mais aprofundada dessas famílias, motivação que fundamenta a adoção da família *Coiflet* no presente trabalho, estabelecendo continuidade e alinhamento metodológico com pesquisas previamente desenvolvidas na mesma instituição.

Assim, essa combinação entre CNNs e transformadas *Wavelet* representa uma abordagem consolidada na análise de mamografias, evidenciando ganhos quantificáveis na detecção de lesões sutis e na redução de falsos positivos. Esses avanços estabelecem as bases para a exploração de arquiteturas otimizadas e estratégias de pré-processamento desta pesquisa.

4 PROPOSTA

Esta pesquisa propõe o desenvolvimento e a avaliação de modelos de aprendizado profundo para detecção de nódulos mamários, utilizando tanto uma arquitetura convolucional simples, desenvolvida do zero, quanto redes neurais convolucionais baseadas em *Transfer Learning*. Além disso, são incorporadas técnicas de pré-processamento baseadas na Transformada *Wavelet* e estratégias de *data augmentation*. A proposta é implementada a partir de uma base pública de mamografias, contendo imagens previamente anotadas, utilizando ferramentas de código aberto.

Os modelos são treinados e avaliados em diferentes cenários, considerando ou não o uso da *Wavelet* e das técnicas de aumento de dados, e seus desempenhos são analisados por meio de métricas como acurácia, precisão, recall e F1-score. O objetivo é investigar o impacto dessas estratégias de tratamento de dados na performance dos sistemas de detecção, permitindo compreender como cada combinação de técnicas contribui para a identificação de anomalias.

5 METODOLOGIA

O desenvolvimento deste trabalho segue os princípios do processo de Aprendizado de Máquina Centrado no Ser Humano (HCML - *Human-Centered Machine Learning*) proposto por Martins et al. (2023), conforme representado na Figura 16.

Figura 16 – Etapas do processo de Aprendizado de Máquina Centrado no Ser Humano (HCML)



Fonte: (Martins et al., 2024).

Esse processo compreende as seguintes etapas: 1-análise dos requisitos, 2-preparação dos dados, 3-treinamento do modelo, 4-avaliação do modelo, 5-predição e 6-exportação. Tais fases não são necessariamente lineares, sendo comum o retorno iterativo entre o treinamento e a avaliação até a obtenção de um modelo satisfatório.

A pesquisa é de natureza quantitativa, aplicada e classificada como um quase-experimento tecnológico (Gil, 2002), por envolver base em análise estatística, utilizando dados secundários (prontos).

5.1 COLETA DE DADOS

São utilizadas imagens provenientes da base pública *Mini-MIAS*¹, compostas por dados secundários quantitativos, como imagens de mamografia e anotações feitas por radiologistas.

Mini-MIAS

A base *Mini-MIAS* (*Mammographic Image Analysis Society*) é amplamente utilizada em pesquisas sobre detecção de câncer de mama e consiste em um subconjunto da base original MIAS, contendo 322 imagens mamográficas digitalizadas a partir de filmes do programa nacional de triagem do Reino Unido. As imagens possuem resolução de 1024×1024 pixels, em formato PGM (*Portable Gray Map*). Cada imagem é acompanhada de anotações fornecidas por radiologistas, que indicam a presença, o tipo (como massa, calcificação, distorção arquitetural) e a localização aproximada das anomalias por meio de coordenadas do centro e raio de uma elipse. Vale destacar que esses rótulos representam

¹ <https://www.kaggle.com/datasets/kmader/mias-mammography>.

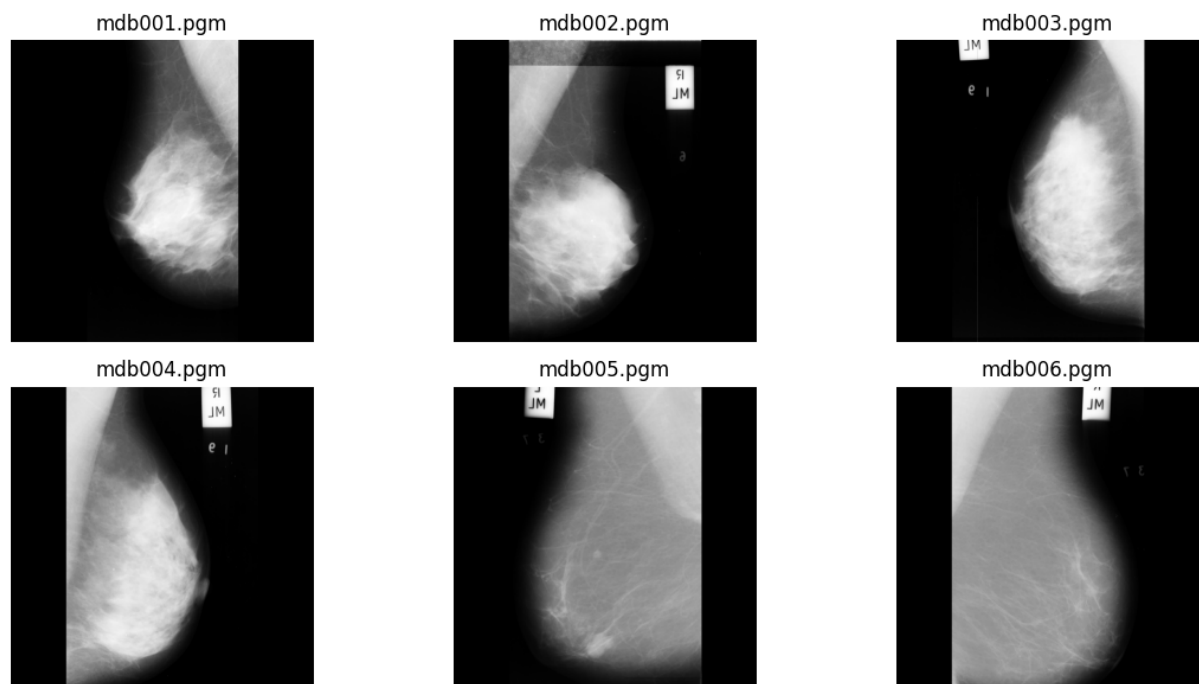
avaliações clínicas, e não confirmação histopatológica, o que deve ser considerado na análise dos resultados. Ainda assim, essa base permite o treinamento e a avaliação de modelos voltados à triagem e apoio ao diagnóstico, especialmente em contextos de investigação de densidade mamária e detecção de nódulos (Mammographic Image Analysis Homepage, 2009).

Cada imagem da base possui uma etiqueta descritiva composta por colunas que indicam: (i) o identificador do exame, (ii) a densidade do tecido mamário, podendo ser F (gorduroso), G (gorduroso-glandular) ou D (denso-glandular), (iii) o tipo de anormalidade presente, como calcificações (CALC), massas circunscritas (CIRC), massas espiculadas (SPIC), distorções arquitetônicas (ARCH), assimetrias (ASYM), outras massas mal definidas (MISC) ou ainda casos normais (NORM), (iv) a gravidade da anormalidade, com rótulos B (benigno) ou M (maligno), e (v-vi) as coordenadas centrais da lesão (x, y), seguidas de (vii) um raio estimado em pixels de um círculo que engloba a região de interesse (Mader, 2018).

As imagens estão organizadas em pares, correspondendo às mamas esquerda e direita de cada paciente. Em casos de calcificações amplamente distribuídas, as coordenadas centrais e raios podem estar ausentes. A origem das coordenadas está localizada no canto inferior esquerdo da imagem.

A Figura 17 apresenta exemplos das primeiras imagens do conjunto Mini-MIAS, ilustrando a aparência típica das mamografias utilizadas no estudo.

Figura 17 – Exemplos das primeiras imagens do conjunto Mini-MIAS



Fonte: Elaborada pela autora.

5.2 EXPERIMENTO

Para avaliar o impacto de diferentes estratégias de pré-processamento, *data augmentation* e arquiteturas de redes neurais convolucionais na tarefa de classificação binária (Normal vs. Anormal), foram conduzidos experimentos sistemáticos divididos em três etapas principais:

1. Definição da arquitetura

- **CNN simples desenvolvida manualmente:** composta por duas camadas convolucionais (8 e 16 filtros de 3×3 , com ativação ReLU), seguidas de operações de *max-pooling* 2×2 que reduzem progressivamente a resolução espacial ($224 \rightarrow 112 \rightarrow 56$). Após o achatamento do mapa de características, uma camada totalmente conectada realiza a classificação binária.
- **Modelos pré-treinados ResNet18, ResNet34 e ResNet50:** todas as variantes foram inicializadas com pesos treinados no ImageNet e tiveram sua camada final substituída por um classificador de duas saídas. As ResNets empregam blocos residuais que facilitam o fluxo de gradiente, permitindo analisar o comportamento de arquiteturas com diferentes profundidades e capacidades de representação. As arquiteturas foram selecionadas com base em sua ampla adoção e validação na literatura especializada em classificação de imagens médicas.

2. Aplicação de combinações de tratamento de dados

Para cada arquitetura foram testadas até quatro configurações:

1. Sem pré-processamento e sem *data augmentation*
2. Apenas *Data Augmentation*
3. Apenas Pré-processamento Wavelet
 - Coiflet 4: Filtragem utilizando *wavelets* da família Coiflet de ordem 4
 - Daubechies 4: Filtragem utilizando *wavelets* da família Daubechies de ordem 4
 - Symlet 4: Filtragem utilizando *wavelets* da família Symlet de ordem 4
4. Pré-processamento Wavelet + *Data Augmentation*
 - Coiflet 4 + Augmentation: Combinação da filtragem Coiflet 4 com técnicas de aumento de dados

- Daubechies 4 + Augmentation: Combinação da filtragem Daubechies 4 com técnicas de aumento de dados
- Symlet 4 + Augmentation: Combinação da filtragem Symlet com técnicas de aumento de dados

A Quadro 1 resume o planejamento experimental adotado.

Quadro 1 – Cenários de Treinamento Considerados

Arquitetura	Configuração
CNN Simples	Baseline + Augmentation + Preproc Coiflet + Preproc db4 + Preproc sym4 + Preproc Coiflet + Augmentation + Preproc db4 + Augmentation + Preproc sym4 + Augmentation
ResNet18	Baseline + Augmentation + Preproc Coiflet + Preproc db4 + Preproc sym4 + Preproc Coiflet + Augmentation + Preproc db4 + Augmentation + Preproc sym4 + Augmentation
ResNet34	Baseline + Augmentation + Preproc Coiflet + Preproc db4 + Preproc sym4 + Preproc Coiflet + Augmentation + Preproc db4 + Augmentation + Preproc sym4 + Augmentation
ResNet50	Baseline + Augmentation + Preproc Coiflet + Preproc db4 + Preproc sym4 + Preproc Coiflet + Augmentation + Preproc db4 + Augmentation + Preproc sym4 + Augmentation

3. Avaliação padronizada

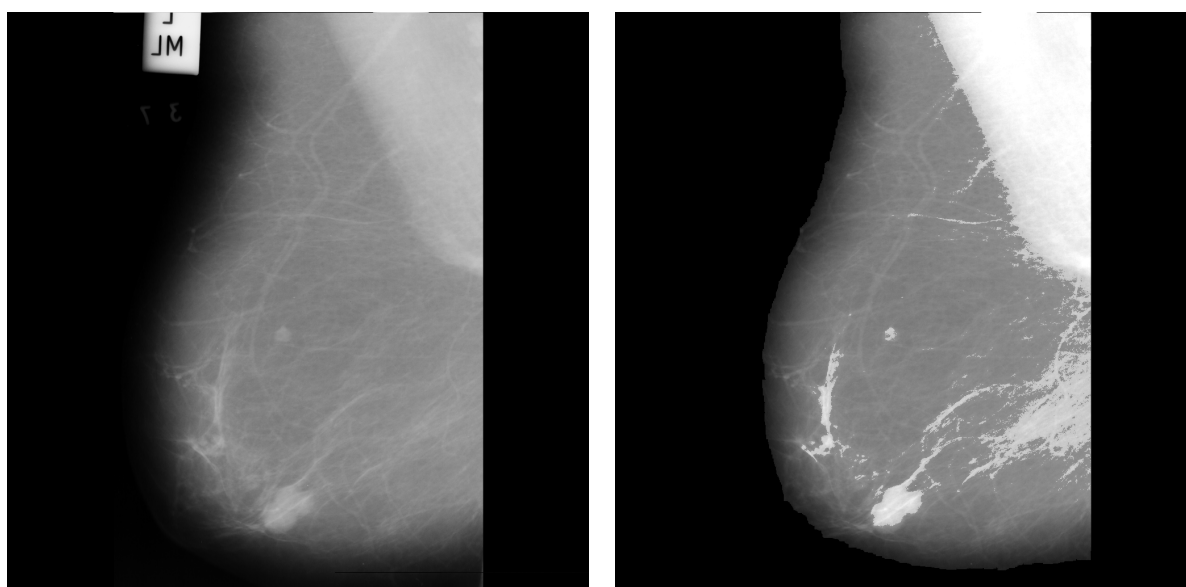
Cada modelo foi treinado e avaliado com a mesma divisão de dados, utilizando como métricas:

- **Acurácia:** proporção de predições corretas no total.
- **Precision (Precisão):** proporção de predições corretas para cada classe (Normal ou Anormal).
- **Recall (Sensibilidade):** proporção de exemplos reais de cada classe corretamente identificados.
- **F1-score:** média harmônica entre precisão e recall para cada classe.
- **Observações qualitativas:** considerações sobre desbalanceamento, viés e estabilidade do modelo.

5.2.1 Pré-processamento das Imagens

O pré-processamento foi composto por quatro etapas principais: segmentação da mama, realce de regiões densas, remoção de ruído por *wavelets* e ajuste de brilho local. Após essas etapas, as imagens foram convertidas para tensor e normalizadas com média 0,5 e desvio-padrão 0,5. A figura a seguir demonstra a comparação entre uma imagem original e uma processada.

Figura 18 – Exemplo de comparação visual entre a imagem original e a imagem pré-processada



(a) Imagem Original

(b) Imagem Pré-processada

Fonte: Elaborada pela autora.

Segmentação da Região da Mama

A primeira etapa consistiu na segmentação da área mamária, removendo fundo preto, rótulos e demais artefatos que não contribuem para a análise. O processo ocorreu em cinco passos:

1. Binarização inicial da imagem;
2. Remoção de ruídos por operações morfológicas;
3. Identificação das componentes conectadas;
4. Seleção da maior região conectada, assumida como o tecido mamário;
5. Aplicação de máscara para descartar todas as demais regiões.

A segmentação garantiu que apenas o tecido mamário fosse mantido para as etapas posteriores, produzindo imagens mais consistentes para o treinamento. A função resultante isola a região mamária de forma consistente e padronizada.

Realce de Pixels Claros (Regiões Densas)

Para destacar áreas de potencial interesse radiológico, que frequentemente apresentam maior intensidade de pixel, foi aplicada uma operação de limiarização com limiar fixo de 145, sendo um limiar de intensidade na escala 0–255 da imagem em tons de cinza. Esta abordagem, determinada empiricamente através de análise visual iterativa, mostrou-se eficaz para isolar seletivamente os pixels de maior intensidade luminosa na imagem segmentada, correspondentes às regiões de densidade glandular e possíveis anomalias. Os pixels identificados por esta máscara binária foram subsequentemente realçados nas etapas de *denoising* e ajuste de brilho que se seguiram.

Denoising por *Wavelets* — *Hard Threshold*

A redução de ruído foi realizada utilizando a transformada *wavelet* 2D com limiarização do tipo *hard threshold*. Para avaliar a sensibilidade do método à escolha da família de *wavelets*, foram testadas múltiplas famílias (incluindo Coiflet, Daubechies e Symlet). O cálculo do limiar seguiu a abordagem do limiar universal, utilizando a estimativa do desvio padrão do ruído (σ) calculada a partir da mediana dos coeficientes de detalhe (d) normalizada pela constante 0,6745.

$$\sigma = \frac{\text{mediana}(|d|)}{0,6745} \quad (5.1)$$

Esta técnica:

- Preserva bordas, microestruturas e detalhes morfológicos sutis relevantes para o diagnóstico;
- Reduz ruído granular, artefatos de aquisição e variações indesejadas de iluminação;
- Mantém a integridade das regiões não ruidosas e a coerência anatômica global da imagem mamográfica.

Ajuste de Brilho Local

Por fim, aplicou-se um aumento de brilho somente nas regiões claras previamente segmentadas, evitando alterações indesejadas no restante da mama. O ajuste utilizado foi de aproximadamente 35%. Esse procedimento aumenta o contraste interno da mama sem causar saturação e contribui para realçar padrões estruturais importantes.

5.2.2 Data Augmentation

Com o objetivo de aumentar a diversidade do conjunto de treinamento e reduzir o *overfitting*, foram aplicadas transformações aleatórias apenas nas imagens da etapa de treino. As operações utilizadas foram:

- Reflexão horizontal e vertical com probabilidade de 50%;
- Rotação aleatória no intervalo de -15° a $+15^\circ$.

A escolha dessas transformações deve-se ao fato de que elas preservam a anatomia mamária e introduzem variações realistas. Operações mais agressivas, como rotações grandes, foram evitadas por poderem gerar imagens fisicamente incoerentes. No conjunto de teste, nenhuma transformação aleatória foi aplicada; apenas a normalização padrão foi utilizada.

5.2.3 Procedimentos de Treinamento e Teste

O treinamento dos modelos foi realizado de forma padronizada para todas as arquiteturas avaliadas, garantindo comparabilidade entre os resultados. Inicialmente, o conjunto de dados foi dividido em 80% para treino e 20% para teste, mantendo a mesma proporção para todos os experimentos.

Balanceamento do conjunto de treino

Como a base apresenta desbalanceamento entre as classes Normal e Anormal, foi empregado um *Weighted Random Sampler*. Cada amostra recebeu um peso inversamente proporcional à frequência da sua classe, promovendo um processo de amostragem mais

equilibrado durante o treinamento. Esse procedimento evita que o modelo seja enviesado para a classe majoritária e foi aplicado de forma uniforme a todas as arquiteturas.

Configuração dos DataLoaders

No PyTorch, o *DataLoader* é responsável por carregar as imagens durante o treinamento e organizá-las em pequenos grupos chamados *batches*, que são utilizados a cada atualização dos pesos da rede. O tamanho desses grupos é definido pelo parâmetro *batch size*.

Os DataLoaders foram configurados com:

- *Batch size* de 16, definindo que cada iteração do treinamento utiliza um grupo de 16 imagens;
- *Sampler* balanceado no conjunto de treino, assegurando probabilidade proporcional entre as classes;
- Ordem fixa no conjunto de teste, sem embaralhamento, garantindo reprodutibilidade.

Essa configuração garante que os modelos recebam *minibatches* organizados de forma consistente e comparável entre todos os experimentos.

Processo de treinamento

Todos os modelos foram treinados sob as mesmas condições:

- Otimização: Adam (*learning rate* = 0,001);
- Função de perda: *CrossEntropyLoss*;
- Número de épocas: 30.

Durante cada época, o modelo percorreu todos os lotes do conjunto de treino, executando o *forward*, que consiste na passagem das imagens pela rede para geração das predições. Em seguida, é realizado o cálculo da perda, a retropropagação do erro e a atualização dos pesos. O valor médio da perda por época foi registrado para monitoramento da convergência.

5.3 IMPLEMENTAÇÃO

A implementação foi realizada em Python 3.8+, utilizando o framework PyTorch 2.0 para construção e treinamento das redes neurais, e PyWavelets 1.4.1 para aplicação da decomposição wavelet multinível nas imagens. A reprodutibilidade é garantida pela disponibilização do código-fonte em repositório público no GitHub², permitindo que os experimentos possam ser replicados por outros pesquisadores.

Ambiente e Configurações Técnicas

O desenvolvimento foi conduzido no ambiente Google Colaboratory, que disponibiliza uma GPU NVIDIA Tesla T4 com 16 GB de VRAM quando esse recurso está habilitado na sessão. O *PyTorch* realiza automaticamente a detecção do dispositivo e utiliza a GPU para acelerar o treinamento; caso a GPU não esteja disponível, a execução ocorre normalmente na CPU.

² <https://github.com/jessicac13/tcc>

6 RESULTADOS E DISCUSSÃO

Nesta seção, os resultados quantitativos são apresentados prioritariamente por meio de tabelas, permitindo comparação direta entre arquiteturas e configurações experimentais. As figuras globais são mantidas apenas como apoio visual no Apêndice A, enquanto a análise principal baseia-se nas métricas reportadas nas tabelas. Foram avaliadas quatro arquiteturas: CNN Simples, ResNet18, ResNet34 e ResNet50, em quatro cenários experimentais.

Para as comparações entre configurações, adotou-se como referência o pré-processamento baseado na família *Wavelet Coiflet*, uma vez que essa família apresentou o melhor equilíbrio entre as classes ao longo dos experimentos, proporcionando redução de ruído e realce adequado das estruturas relevantes.

Com isso, as análises que seguem são organizadas considerando quatro cenários para cada arquitetura: (i) sem pré-processamento e sem data augmentation (*baseline*); (ii) apenas com *data augmentation*; (iii) apenas com pré-processamento *Coiflet*; e (iv) com a combinação de *Coiflet* e *data augmentation*.

O conjunto de teste contém 42 imagens da classe Normal e 24 da classe Anormal, caracterizando um desbalanceamento moderado. Embora tenham sido utilizados pesos de classe durante o treinamento, observou-se que a classe Normal ainda tende a dominar o aprendizado em diversas configurações, tornando essencial a análise individual das métricas por classe.

Os resultados obtidos também evidenciam a influência do volume de dados no desempenho das arquiteturas avaliadas, especialmente para a classe Anormal. Conjuntos de dados mais extensos tendem a favorecer a generalização dos modelos e a estabilidade das métricas.

MODELO CNN SIMPLES

O primeiro modelo avaliado foi uma CNN simples, cujos resultados estão apresentados na Tabela 1. O modelo obteve acurácia de aproximadamente 52%. Embora a classe Normal apresente precisão de 61% e recall de 67%, esses valores não representam um desempenho satisfatório, pois em uma tarefa binária resultados próximos de 60% ficam perigosamente próximos do comportamento aleatório. Já a classe Anormal apresentou desempenho ainda mais limitado (*precision* de 30% e *recall* de 25%), evidenciando forte incapacidade de identificar corretamente os casos minoritários. Esse desequilíbrio se reflete nos f1-scores (0,64 para Normal e 0,27 para Anormal), indicando que, apesar de reconhecer alguns padrões básicos, o modelo falha em alcançar um nível de confiabilidade

mínimo para aplicações clínicas.

Ao aplicar *data augmentation*, observou-se uma melhora significativa na classe Anormal em termos de f1-score, embora acompanhada de queda na acurácia global. Isso indica que o aumento de variabilidade beneficia a generalização da classe minoritária ao reduzir o viés em direção à classe Normal. Ainda assim, o modelo não conseguiu atingir um equilíbrio satisfatório, reforçando a ideia de que a *CNN* simples não é suficientemente expressiva para capturar adequadamente as nuances do problema.

A introdução do pré-processamento baseado na *wavelet Coiflet*, mesmo sem *data augmentation*, gerou um salto expressivo de desempenho: acurácia de 0,65 e f1-score de 0,51 para a classe Anormal. Comparado ao modelo com *data augmentation* sem pré-processamento (acurácia de 0,50 e f1-scores inferiores), fica claro que o pré-processamento foi o fator determinante para a melhoria do aprendizado.

Combinando pré-processamento e *data augmentation*, o modelo atingiu maior equilíbrio entre as classes, embora com acurácia moderada de 0,55. Houve perda de sensibilidade na classe Normal (recall de 0,38), mas um aumento expressivo na classe Anormal (recall de 0,83 e f1-score de 0,57, seu melhor até então). Isso demonstra que, após a limpeza e estruturação dos dados proporcionadas pela *Coiflet*, a *data augmentation* passou a atuar de forma mais eficaz, reforçando a capacidade de identificação da classe minoritária.

Tabela 1 – Resultados da CNN Simples sob diferentes configurações

Arq.	Config.	Acc	P(N)	R(N)	F1(N)	P(A)	R(A)	F1(A)
CNN Simples	Baseline	0,5152	0,61	0,67	0,64	0,30	0,25	0,27
CNN Simples	+ Augmentation	0,5000	0,65	0,48	0,55	0,37	0,54	0,44
CNN Simples	+ Preproc coif4	0,6515	0,72	0,74	0,73	0,52	0,50	0,51
CNN Simples	+ Preproc coif4 + Aug	0,5455	0,80	0,38	0,52	0,43	0,83	0,57

Fonte: Elaborada pela autora.

Destques: Os valores em **laranja** indicam o desempenho *baseline*, caracterizado por limitação na identificação da classe Anormal. Em **azul**, a *data augmentation* melhora a representação da classe Anormal, porém com impacto negativo na acurácia global. Os valores em **verde** indicam ganho consistente com o pré-processamento *Coiflet*, enquanto em **roxo** destaca-se a combinação *Coiflet + augmentation*, que favorece a sensibilidade à classe Anormal, à custa de queda no desempenho da classe Normal.

MODELO RESNET18

O uso de arquiteturas mais profundas também trouxe ganhos importantes. A ResNet18 sem pré-processamento superou a *CNN* simples, alcançando acurácia de 0,5606 e melhorias em todas as métricas (conforme apresentado na Tabela 2). A classe Normal apresentou f1-score de 0,67 e a Anormal de 0,33, indicando avanço na extração de características mesmo sem qualquer tratamento adicional.

Com *data augmentation*, a ResNet18 atingiu acurácia de 51,5%, e ainda apresentou desempenho muito distinto entre as classes: f1-score de 0,64 para Normal e apenas 0,27 para Anormal. Apesar do balanceamento por pesos, o modelo continuou tendo dificuldade para aprender padrões representativos da classe Anormal quando os dados são utilizados sem pré-processamento. Isso sugere que a variabilidade introduzida pela *data augmentation*, aplicada sobre imagens brutas, não foi suficiente para melhorar a capacidade do modelo de distinguir características relevantes da classe minoritária.

A inserção do pré-processamento *Coiflet*, sem *data augmentation*, evidenciou novamente que essa etapa é o fator mais determinante: a acurácia subiu para 0,62, e o f1-score da classe Anormal aumentou de 0,13 para 0,36. Esse salto indica que a transformação *Wavelet* reduziu ruídos e tornou estruturas relevantes mais distinguíveis, facilitando o aprendizado mesmo com pesos balanceados.

Com a combinação de pré-processamento e *data augmentation*, a ResNet18 alcançou um de seus resultados mais equilibrados: acurácia de 0,64, com f1-scores de 0,71 para Normal e 0,52 para Anormal. Nessa configuração, a *data augmentation* passou a atuar de forma benéfica, pois a entrada já havia sido estabilizada pelo pré-processamento. Assim, a variabilidade adicional não gerou inconsistências, mas contribuiu para melhorar a representação da classe Anormal sem comprometer significativamente o desempenho na classe Normal.

Tabela 2 – Resultados da ResNet18 sob diferentes configurações

Arq.	Config.	Acc	P(N)	R(N)	F1(N)	P(A)	R(A)	F1(A)
ResNet18	Baseline	0,5606	0,64	0,71	0,67	0,37	0,29	0,33
ResNet18	+ Augmentation	0,5152	0,61	0,67	0,64	0,30	0,25	0,27
ResNet18	+ Preproc coif4	0,6212	0,67	0,81	0,73	0,47	0,29	0,36
ResNet18	+ Preproc coif4 + Aug	0,6364	0,72	0,69	0,71	0,50	0,54	0,52

Fonte: Elaborada pela autora.

Destaques: Os valores em **laranja** indicam o desempenho da ResNet18 sem pré-processamento, com limitação na identificação da classe Anormal. Em **azul**, a *data augmentation* isolada não trouxe ganhos relevantes, evidenciando que a variabilidade não é eficaz sem tratamento prévio dos dados. Os valores em **verde** indicam melhora consistente com o uso do pré-processamento *Coiflet*, especialmente para a classe Anormal. Por fim, os valores em **roxo** destacam a combinação *Coiflet + augmentation*, que resulta em um dos cenários mais equilibrados entre as classes.

MODELO RESNET34

A Tabela 3 resume os resultados obtidos com a ResNet34. O modelo sem pré-processamento e sem *data augmentation* apresentou acurácia de 0,61 com forte tendência à classe Normal (f1 = 0,72) e baixo desempenho para Anormal (f1 = 0,32). Ao aplicar

apenas *data augmentation*, o desempenho caiu severamente (acurácia 0,50), mostrando que dados não tratados tornam a *data augmentation* prejudicial, introduzindo ruído e instabilidade.

Quando aplicado o pré-processamento *Coiflet*, o modelo atingiu seus melhores resultados: acurácia de 0,68 e equilíbrio muito superior ($f1 = 0,75$ para Normal e 0,57 para Anormal). Assim como nos modelos anteriores, o pré-processamento mostrou ser essencial para extrair padrões relevantes.

Já a combinação pré-processamento + *data augmentation* levou a uma queda acentuada no desempenho, com colapso total da classe Anormal ($f1 = 0$). A variabilidade artificial superou a capacidade da rede de manter separação entre as classes, evidenciando que a ResNet34 é mais sensível à *data augmentation* e depende fortemente da limpeza fornecida pela *Coiflet*.

Tabela 3 – Resultados da ResNet34 sob diferentes configurações

Arq.	Config.	Acc	P(N)	R(N)	F1(N)	P(A)	R(A)	F1(A)
ResNet34	Baseline	0,6061	0,65	0,81	0,72	0,43	0,25	0,32
ResNet34	+ Augmentation	0,5000	0,60	0,64	0,62	0,29	0,25	0,27
ResNet34	+ Preproc coif4	0,6818	0,76	0,74	0,75	0,56	0,58	0,57
ResNet34	+ Preproc coif4 + Aug	0,5455	0,60	0,86	0,71	0,00	0,00	0,00

Fonte: Elaborada pela autora.

Destques: Os valores em **laranja** indicam o desempenho da ResNet34 sem pré-processamento, caracterizado por forte viés em favor da classe Normal e limitação na identificação da classe Anormal. Em **azul**, a aplicação isolada de *data augmentation* resulta em queda acentuada de desempenho, sugerindo aumento de ruído e instabilidade no aprendizado. Os valores em **verde** destacam o melhor cenário, no qual o pré-processamento *Coiflet* promove melhora consistente e equilibrada entre as classes. Por fim, os valores em **roxo** evidenciam o colapso da classe Anormal quando pré-processamento e *data augmentation* são combinados, indicando elevada sensibilidade da arquitetura ao excesso de variabilidade.

MODELO RESNET50

Os resultados obtidos pela ResNet50 sob diferentes configurações são apresentados na Tabela 4. A arquitetura sem pré-processamento e sem *data augmentation* apresentou acurácia de 0,50 e grande desequilíbrio entre as classes. Com *data augmentation* isolada, houve melhoria aparente na acurácia (0,58), mas a classe Anormal praticamente colapsou ($f1 = 0,07$). Isso mostra que a *data augmentation* sem pré-processamento intensifica padrões da classe majoritária. O pré-processamento Novamente, com a aplicação do pré-processamento *Coiflet*, observou-se a maior evolução: acurácia de 0,65, $f1$ da classe Normal de 0,74 e $f1$ da classe Anormal de 0,49 (valores superiores aos obtidos nas configurações anteriores). O pré-processamento proporcionou estabilidade, redução de ruído e realce

das estruturas relevantes, permitindo que a ResNet50 extraísse características de forma mais consistente.

Com pré-processamento + *data augmentation*, o desempenho manteve-se estável (acurácia de 0,65), com leve queda para a classe Anormal, mas sem os colapsos observados em arquiteturas menores. Isso indica que a ResNet50 tolera melhor o aumento de variabilidade quando os dados já chegam limpos e estruturados.

Tabela 4 – Resultados da ResNet50 sob diferentes configurações

Arq.	Config.	Acc	P(N)	R(N)	F1(N)	P(A)	R(A)	F1(A)
ResNet50	Baseline	0,5000	0,59	0,69	0,64	0,24	0,17	0,20
ResNet50	+ Augmentation	0,5758	0,62	0,88	0,73	0,17	0,04	0,07
ResNet50	+ Preproc coif4	0,6515	0,71	0,76	0,74	0,52	0,46	0,49
ResNet50	+ Preproc coif4 + Aug	0,6515	0,70	0,79	0,74	0,53	0,42	0,47

Fonte: Elaborada pela autora.

Destques: Os valores em **laranja** indicam que, sem pré-processamento, a ResNet50 apresenta desempenho próximo ao aleatório e forte desequilíbrio entre as classes. Em **azul**, a *data augmentation* isolada não beneficia a classe Anormal, acentuando o desbalanceamento e comprometendo a generalização. Os valores em **verde** evidenciam ganho expressivo com o pré-processamento *Coiflet*, permitindo melhor aproveitamento da profundidade da arquitetura. Por fim, os valores em **roxo** mostram que a combinação *Coiflet + augmentation* mantém desempenho estável, indicando maior tolerância da ResNet50 à variabilidade quando os dados já estão previamente estruturados.

7 CONCLUSÃO

Esta pesquisa avaliou o impacto da Transformada *Wavelet* e de diferentes arquiteturas de redes neurais convolucionais na detecção de nódulos em mamografias da base Mini-MIAS. Os resultados mostraram que o pré-processamento com a *Wavelet Coiflet* foi a etapa que mais contribuiu para a melhoria do desempenho dos modelos, equilibrando as classes e destacando estruturas relevantes para a classificação. As arquiteturas *ResNet* apresentaram desempenho mais consistente que a *CNN* simples, com destaque para a *ResNet34* combinada ao pré-processamento, que obteve os melhores resultados gerais.

De modo geral, observou-se que a qualidade das imagens de entrada influencia mais o desempenho dos modelos do que a aplicação isolada de técnicas de *data augmentation*. Além disso, algumas combinações entre *Wavelet* e aumento de dados mostraram sensibilidade dependendo da arquitetura utilizada. Esses achados reforçam a importância de estratégias adequadas de pré-processamento e do uso de modelos mais profundos em bases pequenas e desbalanceadas como a Mini-MIAS.

7.1 TRABALHOS FUTUROS

Como continuidade deste estudo, sugerem-se:

- Investigar a integração da *Wavelet* diretamente na pipeline da *CNN*, por meio de camadas *Wavelet-CNN*, permitindo que a decomposição multiescala seja aprendida conjuntamente durante o treinamento.
- Avaliar o método em bases maiores e mais balanceadas, a fim de verificar a capacidade de generalização dos modelos.
- Explorar outras arquiteturas de *CNN*, como EfficientNet, VGG e AlexNet.

REFERÊNCIAS

- AGUIAR NETO, Décio Gonçalves de. **Transferência de conhecimento utilizando aprendizado profundo para classificação de imagens histopatológicas**. Quixadá, Brasil, 2017. Orientador: Prof. Dr. Paulo de Tarso Guerra Oliveira. Disponível em: <https://repositorio.ufc.br/bitstream/riufc/29515/1/2017_tcc_dganaguiar.pdf>.
- BANERJEE, Sheekar; KABIR, Humayun. An introductory implementation of breast cancer detection from mammograms and pixel intensity with Efficient-Net and other neural nets. **Cold Spring Harbor Laboratory**, 2024. Preprint. DOI: 10.1101/2024.05.04.592536. Disponível em: <<https://doi.org/10.1101/2024.05.04.592536>>. Acesso em: 28 jun. 2025.
- BLAHOVÁ, Lenka; KOSTOLNÝ, Jakub; CIMRÁK, Ivan. Neural network-based mammography analysis: augmentation techniques for enhanced cancer diagnosis — a review. **Bioengineering**, v. 12, p. 232, 2025. DOI: 10.3390/bioengineering12030232. Disponível em: <<https://doi.org/10.3390/bioengineering12030232>>. Acesso em: 28 jun. 2025.
- BRAVO, Barbara Silva et al. Câncer de mama: uma revisão de literatura/ Breast cancer: a literature review. **Brazilian Journal of Health Review**, v. 4, n. 3, p. 14254–14264, jun. 2021. DOI: 10.34119/bjhrv4n3-357. Disponível em: <<https://ojs.brazilianjournals.com.br/ojs/index.php/BJHR/article/view/32101>>.
- CADRIN-CHÊNEVERT, Alexandre. Unleashing the Power of Deep Learning for Breast Cancer Detection through Open Mammography Datasets. **Radiology: Artificial Intelligence**, v. 5, n. 2, e220294, 2023. DOI: 10.1148/ryai.220294. Disponível em: <<https://doi.org/10.1148/ryai.220294>>.
- CUNHA, Leonardo Cardoso da. **Redes neurais convolucionais e segmentação de imagens: uma revisão bibliográfica**. 2020. Disponível em: <https://www.monografias.ufop.br/bitstream/35400000/2872/6/MONOGRAFIA_RedNeuraisConvolucionais.pdf>.
- DOMINGUES, M. O. et al. Explorando a transformada wavelet contínua. **Revista Brasileira de Ensino de Física**, v. 38, n. 3, p. 1–12, 2016. DOI: 10.1590/1806-9126-RBEF-2016-0019. Disponível em: <<https://doi.org/10.1590/1806-9126-RBEF-2016-0019>>.
- FERNEDA, Edberto. Redes neurais e sua aplicação em sistemas de recuperação de informação. **Ciência da Informação**, v. 35, n. 1, p. 25–30, 2006. DOI: 10.1590/S0100-19652006000100003. Disponível em: <<https://www.scielo.br/j/ci/a/SQ9myjZWLxnyXfstXMgCdch/?format=pdf&lang=pt>>.
- FREITAS, Andréa Gonçalves de et al. Mamografia Digital: Perspectiva Atual e Aplicações Futuras. **Radiologia Brasileira**, v. 39, n. 4, p. 287–296, 2006. Atualização. DOI: 10.1590/S0100-39842006000400012.

- GIL, Antônio Carlos. **Como Elaborar Projetos de Pesquisa**. 2002. Disponível em: <https://files.cercomp.ufg.br/weby/up/150/o/Anexo_C1_como_elaborar_projeto_de_pesquisa_-_antonio_carlos_gil.pdf>.
- GOSWAMI, Jaideva C.; CHAN, Andrew K. **Fundamentals of Wavelets: Theory, Algorithms, and Applications**. John Wiley & Sons, 1999. ISBN 978-0-471-25578-1.
- ISIN, Ali; OZDALILI, Selen. Cardiac arrhythmia detection using deep learning. **Procedia Computer Science**, v. 120, p. 268–275, 2017. DOI: 10.1016/j.procs.2017.11.238. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S187705091732450X>>.
- JADOON, M. et al. Three-class mammogram classification based on descriptive CNN features. **BioMed Research International**, Hindawi, v. 2017, p. 1–11, 2017.
- LEITE, Francisco E. A.; HENRIQUES, Marcos V. C.; CUNHA, Maytê T. N. Análise de Imagens em Escalas Utilizando a Transformada Wavelet. **Anais do Encontro de Computação do Oeste Potiguar ECOP/UFERSA**, v. 1, n. 2, p. 59–66, 2018. Disponível em: <<https://periodicos.ufersa.edu.br/ecop/article/view/7903>>.
- MADER, Klemens. **MIAS Mammography Database**. 2018. <https://www.kaggle.com/datasets/kmader/mias-mammography>. Disponível no Kaggle. Base com imagens mamográficas digitalizadas e anotações clínicas da Mammographic Image Analysis Society (MIAS). Acesso em: 30 jun. 2025.
- MAGRINI, Luciano Aparecido. Funções wavelet e transformada wavelet contínua: representação simultânea nos domínios do tempo e da frequência. **C.Q.D. – Revista Eletrônica Paulista de Matemática**, v. 19, p. 17–32, 2020. DOI: 10.21167/cqdvol19202023169664lam1732. Disponível em: <<https://www.fc.unesp.br/Home/Departamentos/Matematica/revistacqd2228/v19a03-funcoes-wavelet-e-transformada-wavelet-continua.pdf>>.
- MAMMOGRAPHIC IMAGE ANALYSIS HOMEPAGE. **Databases – MammoImage.org**. 2009. <https://www.mammoimage.org/databases/>. Acessado em 30 de junho de 2025.
- MARTINS, R. Mayor. **Perceptron: projeto e implementação**. 2025. Website pessoal. Acesso em 26 de junho de 2025. Disponível em: <<https://rmayormartins.github.io/tools-ia-projects/perceptron.html>>.
- MARTINS, Ramon Mayor et al. Exploring the relationship between learning of machine learning concepts and socioeconomic status background among middle and high school students: a comparative analysis. **ACM Transactions on Computing Education**, v. 24, n. 3, p. 1–31, set. 2024. DOI: 10.1145/3680288. Disponível em: <<https://doi.org/10.1145/3680288>>. Acesso em: 20 jun. 2025.
- MIYAZAKI, Caio Kioshi. **Redes neurais convolucionais para aprendizagem e reconhecimento de objetos 3D**. 2017. Disponível em: <https://bdta.abcd.usp.br/directbitstream/ee692ea6-3d9c-4d48-92a6-c672e8dbef07/Miyazaki_caio_tcc.pdf>.

MUPPARAJU, Shivani; JAHNAVI, B. Naga Venkata Satya Durga. Comparison of Various Thresholding Techniques of Image Denoising. **International Journal of Engineering Research & Technology (IJERT)**, v. 2, n. 9, p. 3293–3298, 2013. Disponível em: <<https://www.ijert.org/research/comparison-of-various-thresholding-techniques-of-image-denoising-IJERTV2IS90812.pdf>>.

OYELADE, Olaide N.; EZUGWU, Absalom E. A novel wavelet decomposition and transformation convolutional neural network with data augmentation for breast cancer detection using digital mammogram. **Scientific Reports**, v. 12, p. 5913, 2022. DOI: 10.1038/s41598-022-09905-3. Disponível em: <<https://www.nature.com/articles/s41598-022-09905-3>>. Acesso em: 28 jun. 2025.

PASSIG E SILVA, Eduarda. **Detecção de nódulos em imagens de mamografia usando wavelet e inteligência artificial**. 2022. TCC (Engenharia de Telecomunicações) – Instituto Federal de Santa Catarina – Campus São José. Disponível em: <https://wiki.sj.ifsc.edu.br/images/0/09/TCC_68_EDUARDA_PASSIG_E_SILVA.pdf>.

RASHEED, Ahmed et al. **Use of transfer learning and wavelet transform for breast cancer detection**. 2021. arXiv preprint. arXiv: 2103.03602 [eess.IV]. Disponível em: <<https://arxiv.org/abs/2103.03602>>. Acesso em: 28 jun. 2025.

ROSA, Andrey Adriano da. **Reconhecimento por imagem de lances de xadrez com visão computacional e redes neurais convolucionais**. 2025. Disponível em: <https://wiki.sj.ifsc.edu.br/images/a/a0/TCC_ANDREY_ADRIANO_DA_ROSA.pdf>.

SECHOPOULOS, Ioannis; TEUWEN, Jonas; MANN, Ritse. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. **Seminars in Cancer Biology**, v. 72, p. 214–225, 2021. DOI: 10.1016/j.semcancer.2020.06.002. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1044579X20301358>>.

SILVA JUNIOR, Edvaldo Castro da. **Construção e análise de rede neural convolucional com neurônios pulsados para classificação de imagens**. 2022. Dissertação de Mestrado – Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto.

SOUSA, Maria Angélica Zucareli. **Desenvolvimento de um objeto simulador de mama: investigações da percepção visual da imagem e do desempenho de esquemas CADx**. 2017. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. Tese de Doutorado.

TEARE, Peter et al. Malignancy detection on mammography using dual deep convolutional neural networks and genetically discovered false color input enhancement. **Journal of Digital Imaging**, Springer, v. 30, n. 4, p. 499–505, 2017.

VERY TECHNOLOGY. **Machine Learning Algorithms: What is a Neural Network?**

Acesso em: 26 jun. 2025. n.d. Disponível em: <<https://www.verytechnology.com/insights/machine-learning-algorithms-what-is-a-neural-network>>.

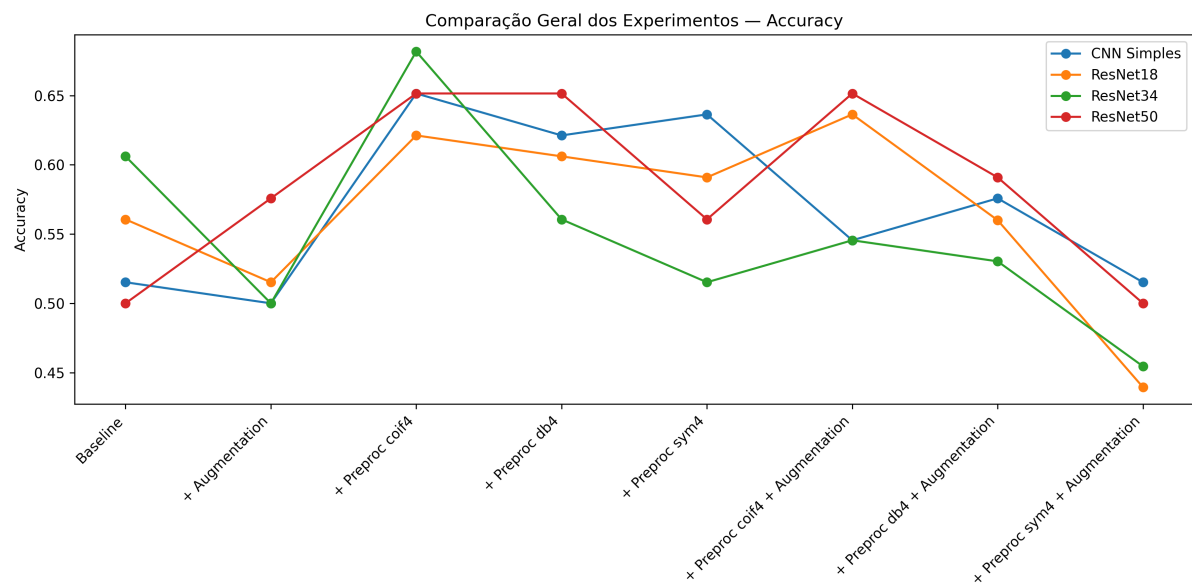
APÊNDICE A – VISUALIZAÇÃO DOS RESULTADOS GLOBAIS

A Figura 19 apresenta a acurácia das arquiteturas avaliadas. As Figuras 20 e 21 exibem, respectivamente, os valores de $F1$ -score para as classes Normal e Anormal.

As Figuras 22 e 23 mostram a precisão das duas classes, enquanto as Figuras 24 e 25 apresentam os respectivos valores de $recall$.

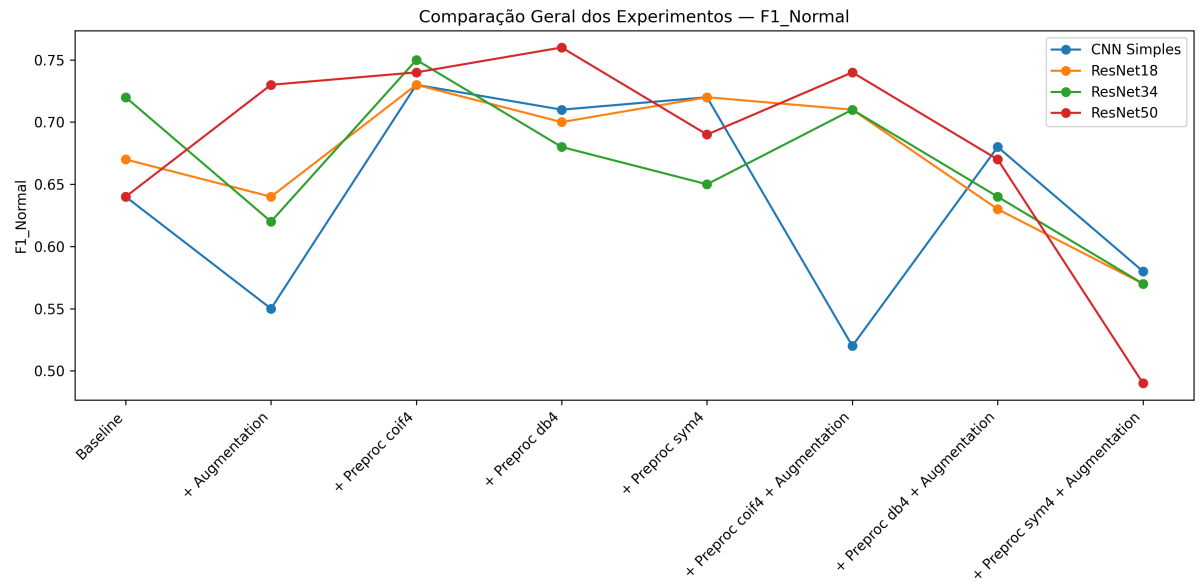
Por fim, as Figuras 26 a 31 reúnem as métricas macro e ponderadas, permitindo uma visão consolidada do desempenho global.

Figura 19 – Acurácia para todas as arquiteturas



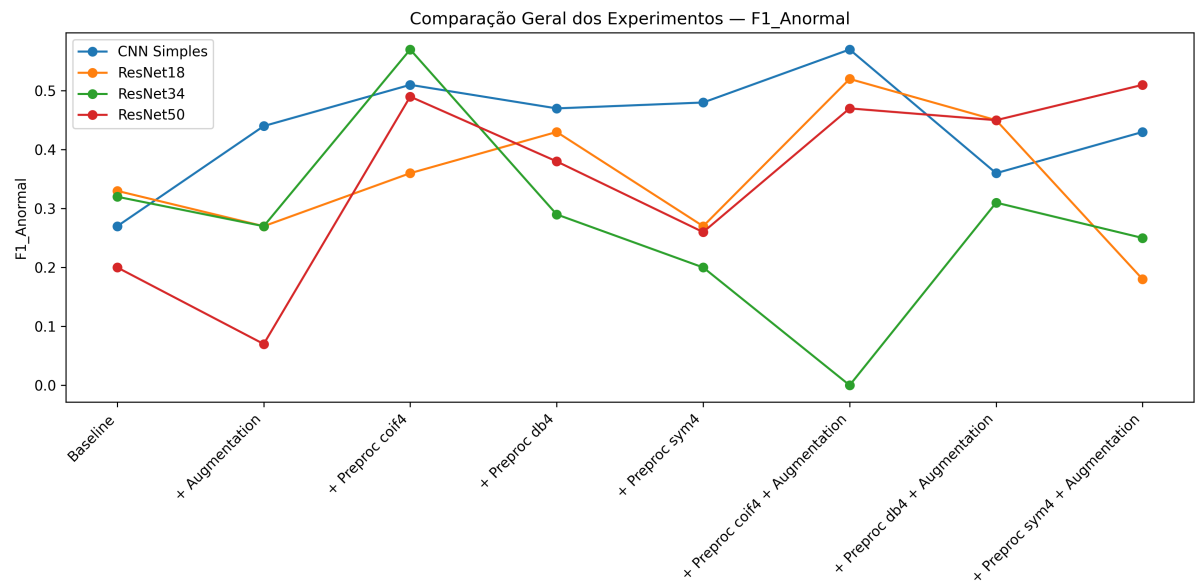
Fonte: Elaborada pela autora.

Figura 20 – F1-score - Classe Normal



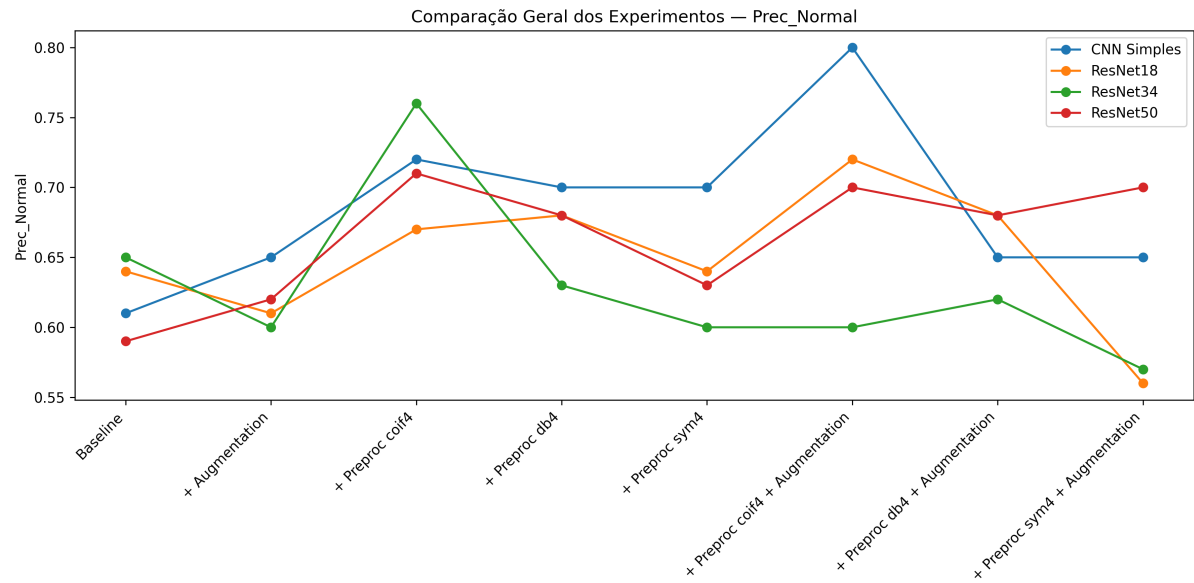
Fonte: Elaborada pela autora.

Figura 21 – F1-score - Classe Anormal



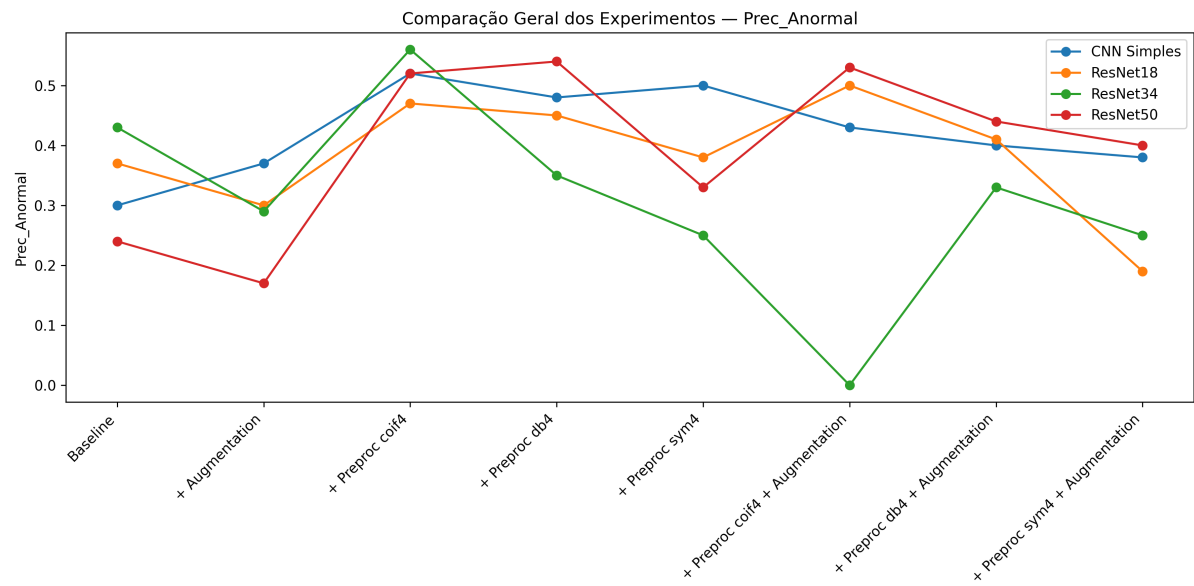
Fonte: Elaborada pela autora.

Figura 22 – Precisão - Classe Normal



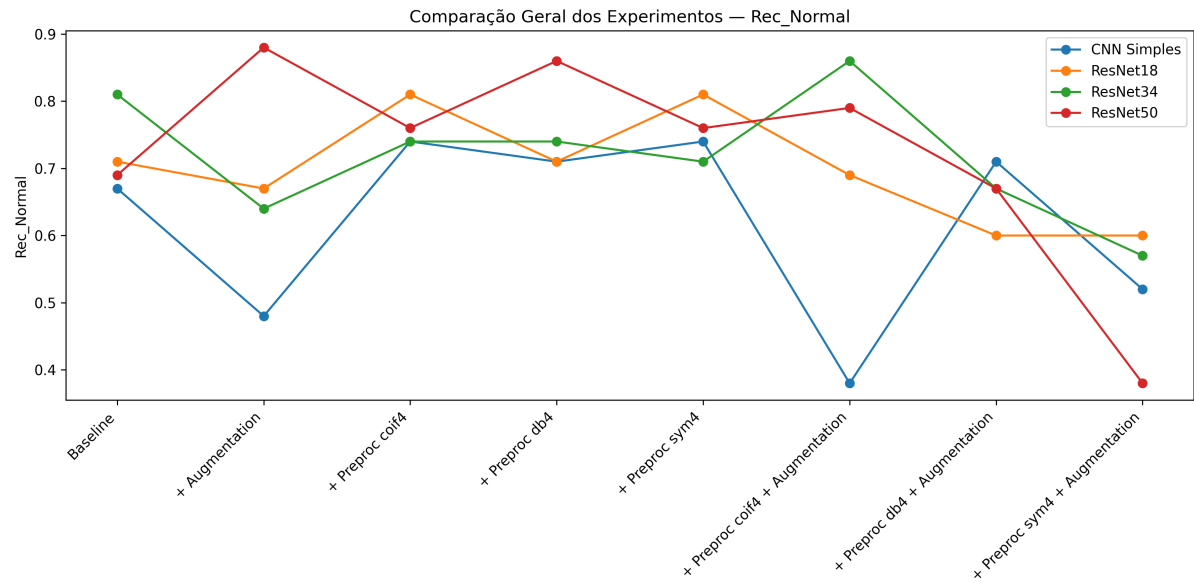
Fonte: Elaborada pela autora.

Figura 23 – Precisão - Classe Anormal



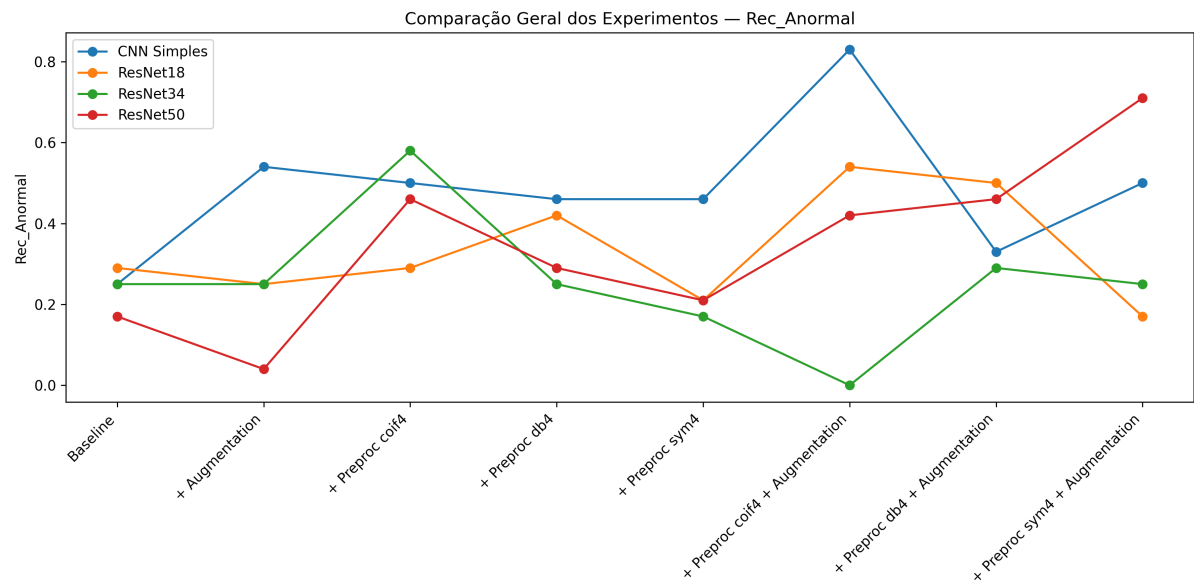
Fonte: Elaborada pela autora.

Figura 24 – Recall - Classe Normal



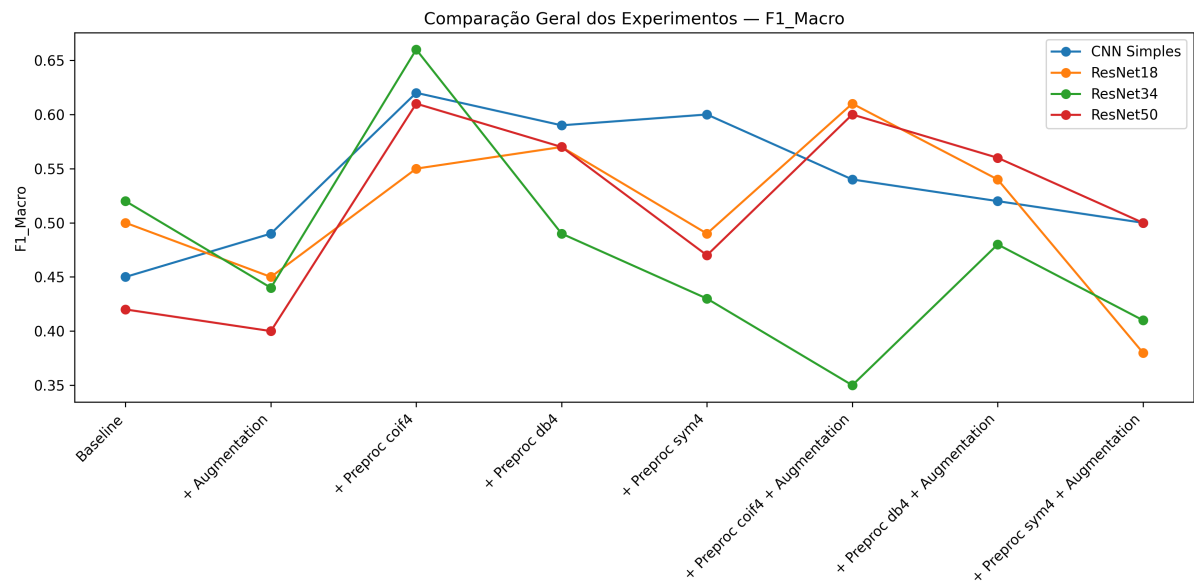
Fonte: Elaborada pela autora.

Figura 25 – Recall - Classe Anormal



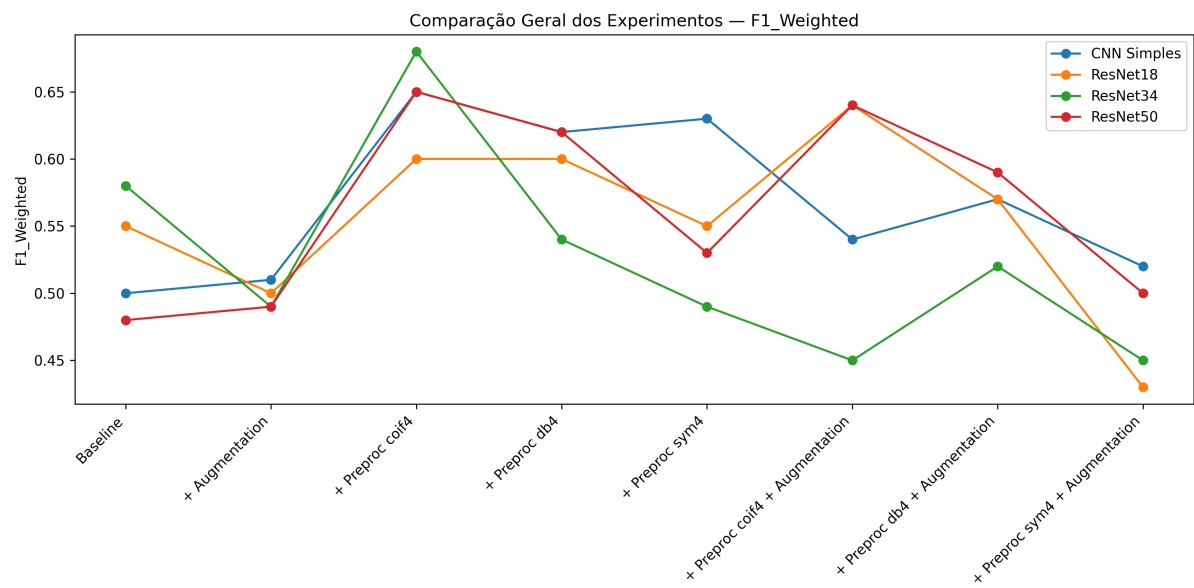
Fonte: Elaborada pela autora.

Figura 26 – F1-score - Macro



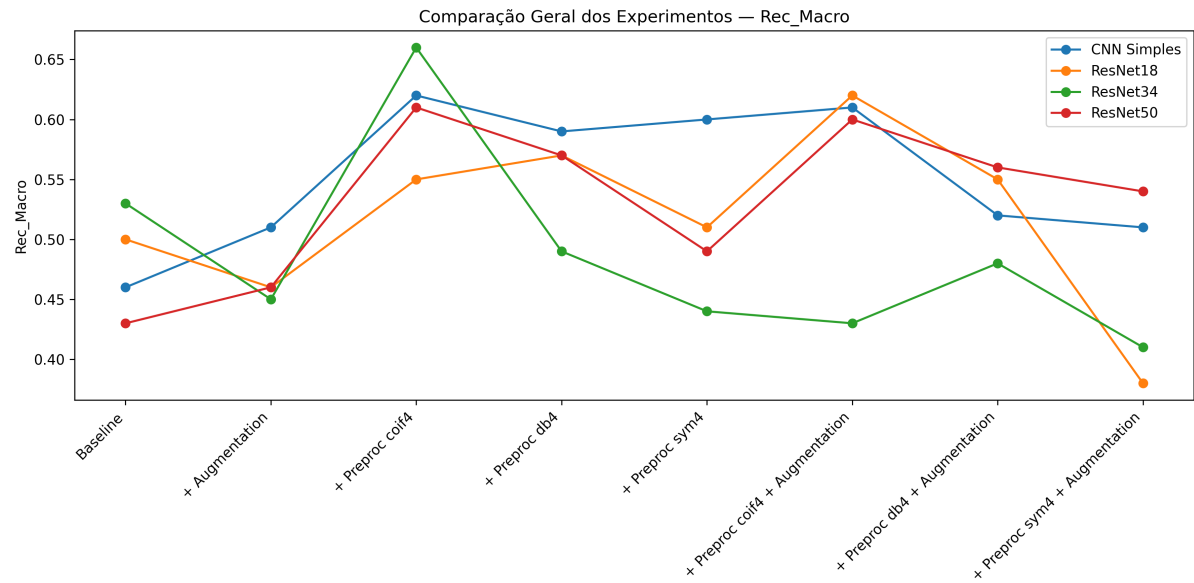
Fonte: Elaborada pela autora.

Figura 27 – F1-score - Ponderado



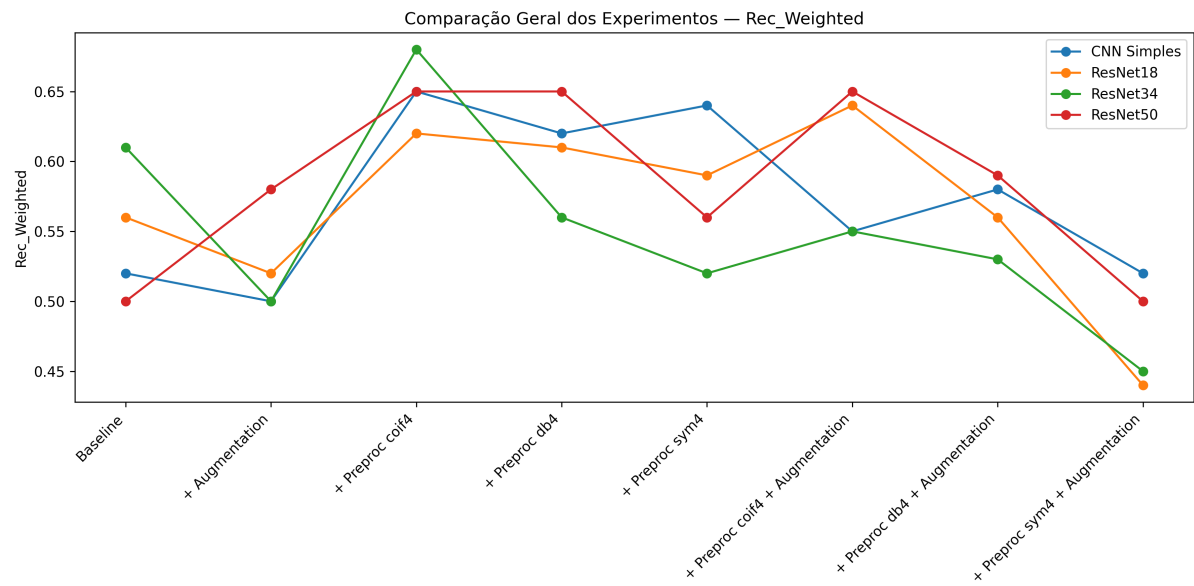
Fonte: Elaborada pela autora.

Figura 28 – Recall - Macro



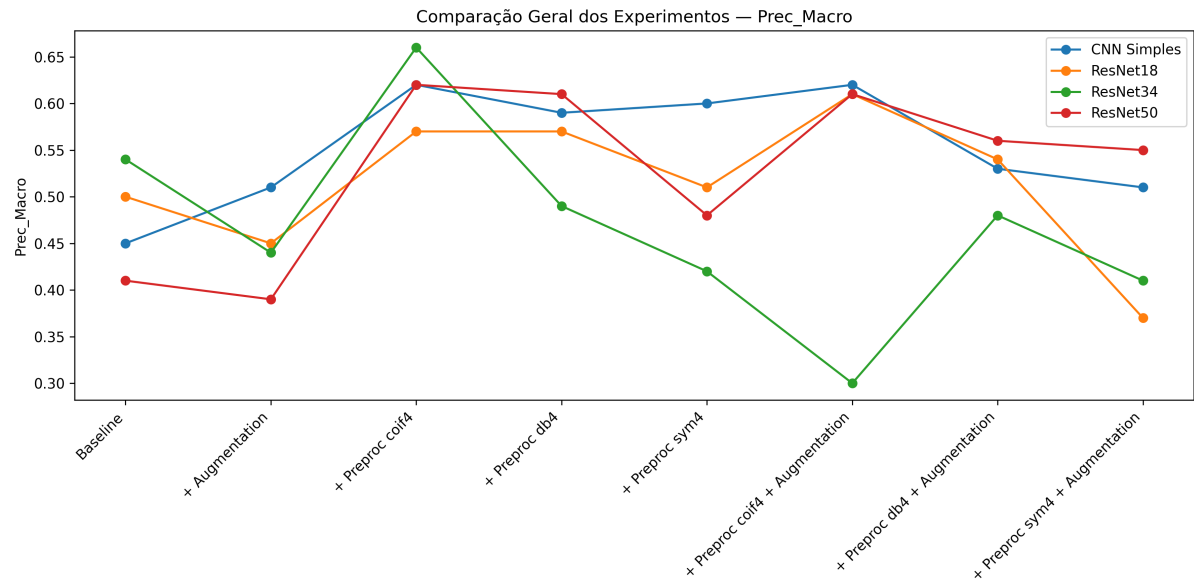
Fonte: Elaborada pela autora.

Figura 29 – Recall - Ponderado



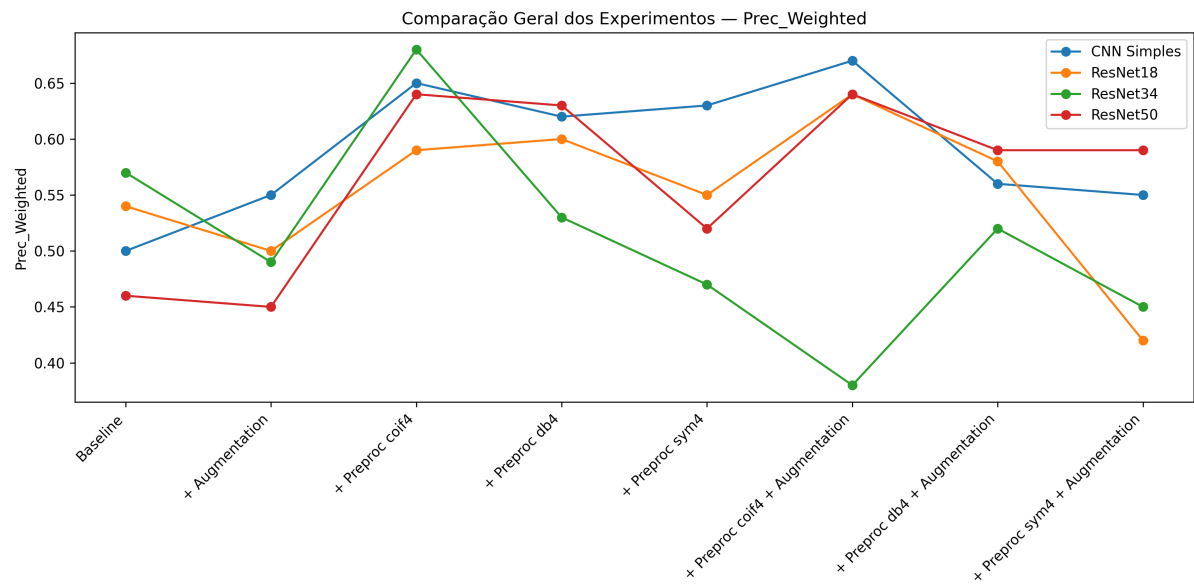
Fonte: Elaborada pela autora.

Figura 30 – Precisão - Macro.



Fonte: Elaborada pela autora.

Figura 31 – Precisão - Ponderado



Fonte: Elaborada pela autora.