

Comparação de Wavelet Pooling e Dual-Branch com CBAM e ROI para Classificação de Mamografias

Beatriz Paz Faria, Ramon Mayor Martins, Elen Macedo Lobato, Instituto Federal de Santa Catarina (IFSC), São José-SC, e-mails: estudo.bea@gmail.com, ramon.mayor@ifsc.edu.br, elen.lobato@ifsc.edu.br.

Beatriz Paz Faria, Ramon Mayor Martins e Elen Macedo Lobato

Resumo— A mamografia é o principal método de rastreamento do câncer de mama, porém a interpretação manual apresenta taxas relevantes de falsos negativos. CNNs têm se mostrado promissoras para diagnóstico auxiliado, mas o impacto isolado de ROI, atenção e wavelet pooling permanece pouco investigado. Este trabalho compara dez arquiteturas no dataset mini-MIAS via *Stratified 5-Fold Cross-Validation*, avaliando extração de ROI, CBAM, quatro famílias wavelet como pooling, estratégia híbrida e dual-branch (ResNet18 + Wavelet Coiflet). A ROI é o componente mais crítico (+35,6 p.p.); Daubechies-4 é a wavelet mais estável (Acc 92,81%, $\sigma=0,77$); e o dual-branch obtém o melhor desempenho (Acc 94,06%, F1 0,939, AUC 0,936, sensibilidade 99,5%).

Palavras-Chave— *Wavelet pooling, CNN, mamografia, dual-branch, CBAM, classificação de imagens médicas.*

Abstract— Mammography is the primary breast cancer screening method, yet manual interpretation yields relevant false-negative rates. CNNs have shown promise for computer-aided diagnosis, but the isolated contribution of ROI extraction, attention, and wavelet pooling remains underexplored. This paper compares ten architectures on the mini-MIAS dataset via *Stratified 5-Fold Cross-Validation*, evaluating ROI extraction, CBAM, four wavelet families as pooling replacements, a hybrid strategy, and a dual-branch model (ResNet18 + Wavelet Coiflet). ROI extraction is the most critical component (+35.6 p.p.); Daubechies-4 is the most stable wavelet for pooling (Acc 92.81%, $\sigma=0.77$); and the dual-branch achieves the best performance (Acc 94.06%, F1 0.939, AUC 0.936, sensitivity 99.5%).

Keywords— *Wavelet pooling, CNN, mammography, dual-branch, CBAM, medical image classification.*

I. INTRODUÇÃO

O câncer de mama é a neoplasia maligna mais frequente entre mulheres no mundo, com cerca de 2,3 milhões de novos casos anuais [1]. Para prevenção, a mamografia permanece como o principal método de rastreamento, porém a interpretação manual está sujeita a taxas consideráveis de falsos negativos e positivos, o que pode comprometer tanto o diagnóstico precoce quanto a eficiência dos serviços de saúde [2]. Nesse contexto, sistemas de diagnóstico auxiliado por computador (CAD) baseados em aprendizado profundo têm demonstrado grande potencial para apoiar radiologistas na leitura dos exames [3].

Entre as técnicas empregadas no processamento dessas imagens, a Transformada Wavelet destaca-se por sua capacidade de decompor sinais em componentes de frequência localizados espacialmente, permitindo analisar detalhes em diferentes escalas [4]. Trabalhos recentes propuseram substituir operações de pooling tradicionais por decomposições wavelet no interior de Redes Neurais Convolucionais (CNNs) [5], [6], buscando preservar mais informação durante o downsampling. Em paralelo, módulos de

atenção como o CBAM (*Convolutional Block Attention Module*) [7] têm sido incorporados a CNNs para realçar automaticamente regiões e canais mais relevantes. Entretanto, a comparação sistemática de múltiplas famílias wavelet nesse papel, bem como a análise de sua interação com mecanismos de atenção e com a extração de regiões de interesse (*Region of Interest* - ROI), permanece pouco explorada em classificação de mamografias.

Este trabalho investiga sistematicamente dez arquiteturas organizadas em três grupos: (I) baselines com variações de ROI e CBAM; (II) quatro famílias wavelet como pooling e uma estratégia híbrida; e (III) uma arquitetura dual-branch combinando ResNet18 [8] e wavelet Coiflet. Todos os modelos são avaliados no dataset mini-MIAS [9] com validação cruzada estratificada de 5 folds, permitindo quantificar o impacto isolado de cada componente da pipeline.

II. TRABALHOS RELACIONADOS

Williams e Li [5] propuseram substituir o pooling por wavelet pooling utilizando Haar, demonstrando melhor preservação de informação durante o downsampling. Na mesma direção, Liu *et al.* [6] introduziram a Multi-level Wavelet CNN (MWCNN), explorando decomposições em múltiplas escalas para restauração de imagens. Em paralelo, o módulo CBAM [7] combina atenção por canal (*channel attention*) e espacial (*spatial attention*), tendo sido aplicado com sucesso em diversas tarefas de visão computacional. Li *et al.* [10] integraram wavelets em CNNs para classificação mais robusta a ruído.

No contexto específico de mamografia, Ribli *et al.* [11] aplicaram *Deep learning* para detecção e classificação de lesões, enquanto Shen *et al.* [12] demonstraram a eficácia de estratégias de transfer learning na melhoria da detecção em rastreamento mamográfico. Focado no dataset mini-MIAS, Hepsağ *et al.* [13] reportaram 87% de acurácia com CNN aplicada a ROIs, ao passo que Aslan *et al.* [14] alcançaram 98,56% com uma arquitetura híbrida CNN+BiLSTM. Apesar desses avanços, não foi identificado um estudo que compare sistematicamente múltiplas famílias wavelet atuando como pooling em conjunto com mecanismos de atenção em um framework unificado, nem que quantifique o impacto isolado de cada componente da pipeline (ROI, CBAM, wavelet) na classificação de mamografias.

III. MATERIAIS E MÉTODOS

A. Dataset e Extração de ROIs

O dataset mini-MIAS [9] contém 322 mamografias (1024×1024, 8-bit escala de cinza), originalmente rotuladas como normal, benigno ou maligno. Nesta pesquisa, adotou-se uma formulação binária normal vs anormal (benigno +

maligno), resultando em 320 amostras utilizáveis, distribuídas em 5 folds estratificados (115 normais e 205 anormais).

Para imagens com lesão anotada, foi extraída uma ROI centrada nas coordenadas fornecidas pelo dataset, utilizando uma janela quadrada com margem de $2\times$ o raio da lesão, de modo a incluir contexto relevante ao redor do nódulo. Para imagens normais, empregou-se a região da mama após segmentação, removendo fundo e artefatos. Em seguida, como pré-processamento, todas as ROIs foram padronizadas via equalização de histograma adaptativa limitada por contraste (*Contrast Limited Adaptive Histogram Equalization* – CLAHE) [15], com o objetivo de realçar contraste local e reduzir variações de iluminação entre exames. Cabe destacar que a extração de ROI foi aplicada exclusivamente às amostras de treinamento; na etapa de validação, utilizou-se a imagem mamográfica completa, sem recorte prévio. Essa escolha visa simular de forma mais fiel o cenário clínico real, no qual o modelo deve ser capaz de generalizar a partir da imagem integral, sem depender de uma localização prévia da lesão.

B. Arquiteturas Comparadas

Dez arquiteturas são comparadas em três grupos:

Grupo I — Baselines: Para avaliar a contribuição dos componentes individuais, quatro arquiteturas de referência são consideradas. O modelo (A) corresponde à CNN base, sem utilização de ROI e CBAM. O modelo (B) adiciona o módulo CBAM à arquitetura base, sem aplicação de ROI, permitindo avaliar o efeito do mecanismo de atenção diretamente sobre a imagem completa. O modelo (C) incorpora a estratégia de ROI à CNN base, sem o uso de CBAM, possibilitando analisar o impacto da restrição espacial na região de interesse. Por fim, o modelo (D) combina ROI e CBAM, permitindo avaliar o ganho conjunto obtido pela utilização da região de interesse associada ao mecanismo de atenção.

A arquitetura base consiste em $\text{Conv}3\times3 \rightarrow \text{BN} \rightarrow \text{ReLU}$ (onde BN representa a normalização em lote, ou *Batch Normalization*) seguida por $\text{MaxPool}2d$, com progressão de canais $1 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 512$. O módulo CBAM [7] é inserido após cada bloco convolucional. O classificador consiste em $\text{AdaptiveAvgPool}2d \rightarrow \text{Linear}(512 \rightarrow 256) \rightarrow \text{ReLU} \rightarrow \text{Dropout}(0,4) \rightarrow \text{Linear}(256 \rightarrow 2)$.

Grupo II — Wavelet como Pooling: (E) Haar; (F) Biorthogonal 1.3 [16]; (G) Daubechies dB4; (H) Coiflet; (I) Híbrida (Haar+dB4+Coiflet+Bior sequenciais). Em todos os casos, utiliza-se ROI+CBAM. O $\text{MaxPool}2d$ é substituído por camada wavelet: convolução depthwise com filtros 2D (LL, LH, HL, HH) obtidos via produto externo dos coeficientes, $\text{stride}=2$, gerando $4\times$ canais. Os filtros são inicializados com a wavelet e tratados como parâmetros treináveis.

A Figura 1 apresenta a arquitetura CNN com Wavelet Pooling + CBAM integrada.

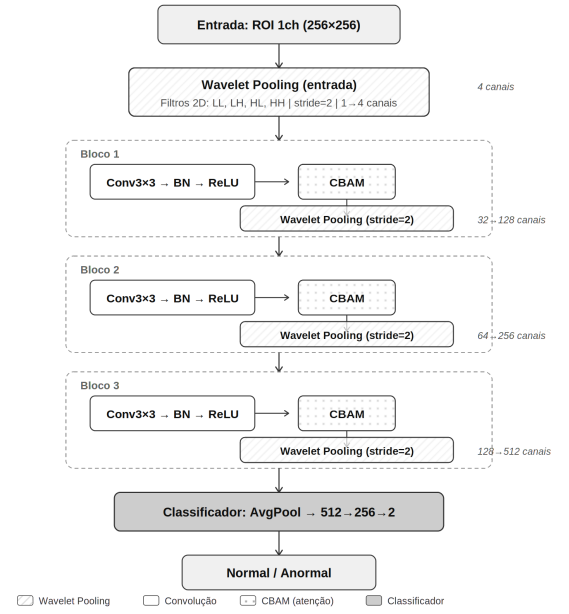


Fig.1. Arquitetura CNN com Wavelet Pooling + CBAM (Modelos B-E).

Grupo III — Dual-Branch: (J) Arquitetura com dois ramos paralelos: um ramo ResNet18 [8] pré-treinado no ImageNet (camada conv1 adaptada para 1 canal, produzindo 512 features) e um ramo wavelet Coiflet com projeção linear ($4 \rightarrow 512$). Os vetores de ambos os ramos são concatenados (1024 dimensões) e processados por classificador totalmente conectado com *dropout* progressivo.

A Figura 2 apresenta a arquitetura Dual-Branch (Modelo J).

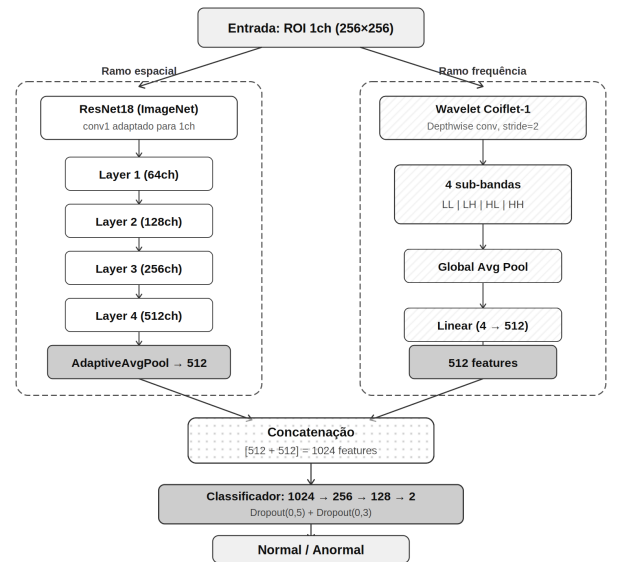


Fig.2. Arquitetura Dual-Branch (Modelo J).

C. Protocolo Experimental

A avaliação dos modelos foi realizada por meio de *Stratified 5-Fold Cross-Validation* ($\text{seed} = 42$), preservando a proporção entre classes em todos os folds. Cada experimento foi treinado por 50 épocas, com *batch size* 32 e taxa de aprendizado inicial de 3×10^{-4} , utilizando o otimizador Adam [17] e função de perda CrossEntropyLoss.

As imagens de entrada foram redimensionadas para 256×256 na fase de teste, enquanto, no treinamento, aplicou-se RandomResizedCrop 224 com escala entre 0,8 e 1,0, além de *data augmentation* composta por *flip* horizontal e rotação $\pm 15^\circ$. Todas as imagens foram normalizadas com média 0,5 e desvio padrão 0,5.

Como métricas de desempenho, foram computadas acurácia, F1-score e AUC-ROC por fold. Adicionalmente, matrizes de confusão globais (soma dos 5 folds, 320 amostras) foram utilizadas para derivar sensibilidade e especificidade dos principais modelos.

D. Ferramental de Desenvolvimento

Os experimentos foram conduzidos em uma CPU com processador AMD Ryzen 7 7735HS (8 núcleos, 16 threads), 16 GB de memória RAM DDR5 a 4800 MT/s e GPU NVIDIA GeForce RTX 4050 Mobile com 6 GB de VRAM, sob o sistema operacional Ubuntu 24.04.4 LTS. Todas as implementações foram desenvolvidas em Python 3.14.4, com o treinamento dos modelos realizado em GPU para viabilizar o processamento das arquiteturas de aprendizado profundo propostas. O treinamento foi conduzido por um número fixo de épocas, sem a utilização de mecanismo de *early stopping*, garantindo que todas as arquiteturas fossem avaliadas sob as mesmas condições experimentais. Não foram observados gargalos de memória durante a execução dos experimentos, incluindo a arquitetura Dual-Branch, que foi treinada integralmente na GPU disponível. O código-fonte dos dez modelos está arquivado publicamente no Zenodo [20].

IV. RESULTADOS

Os modelos A e B (sem ROI) apresentaram desempenho próximo ao acaso ($\sim 55\%$ e $\sim 51\%$), evidenciando que a extração de ROI é o componente mais crítico da pipeline, com ganho de +35,63 p.p. em acurácia ao comparar (A→C). Notavelmente, o modelo B (CBAM sem ROI) teve desempenho inferior ao modelo A, indicando que aplicar atenção sobre a imagem completa tende a realçar também regiões de ruído (fundo, músculo peitoral) que não são clinicamente relevantes. Já a inclusão de CBAM sobre a ROI (C→D) gerou ganho modesto em acurácia (+0,94 p.p.) e pequena melhora em AUC (+0,005), sugerindo que o módulo atua principalmente como amplificador de features quando aplicado a uma região previamente delimitada.

A Tabela I e Figura 3 apresentam os resultados dos dez modelos.

TABELA I

Desempenho dos dez modelos via *Stratified 5-Fold CV*.

Modelo	Acc% $\pm\sigma$	F1 $\pm\sigma$	AUC $\pm\sigma$
A: CNN simples	55,31 $\pm$ 9,25	0,499 $\pm$ 0,082	0,492 $\pm$ 0,022
B: CNN+CBAM	51,25 $\pm$ 8,75	0,507 $\pm$ 0,077	0,444 $\pm$ 0,094
C: CNN+ROI	90,94 $\pm$ 1,53	0,907 $\pm$ 0,016	0,923 $\pm$ 0,018
D: CNN+ROI+CBAM	91,88 $\pm$ 2,30	0,918 $\pm$ 0,023	0,928 $\pm$ 0,028
E: Haar pooling	91,25 $\pm$ 1,25	0,911 $\pm$ 0,013	0,915 $\pm$ 0,038
F: Bior 1.3 pooling	89,69 $\pm$ 2,12	0,896 $\pm$ 0,021	0,923 $\pm$ 0,025
G: dB4 pooling	92,81 $\pm$ 0,77	0,927 $\pm$ 0,008	0,913 $\pm$ 0,032
H: Coiflet pooling	92,50 $\pm$ 1,82	0,924 $\pm$ 0,019	0,911 $\pm$ 0,035
I: Híbrida pipeline	90,62 $\pm$ 1,98	0,905 $\pm$ 0,020	0,923 $\pm$ 0,026
J: Dual-Branch	94,06 $\pm$ 1,53	0,939 $\pm$ 0,016	0,936 $\pm$ 0,037

Entre as variantes com wavelet como pooling (modelos E–H), a Daubechies dB4 (G) destacou-se com a maior acurácia (92,81%) e o menor desvio padrão entre todos os dez modelos ($\sigma = 0,77$), indicando estabilidade excepcional. A Coiflet (H) obteve resultado muito próximo (92,50%), enquanto Haar (E) ficou abaixo do baseline e a Biorthogonal 1.3 (F) apresentou o pior desempenho do grupo (89,69%). A estratégia híbrida (I) não superou as melhores wavelets individuais, sugerindo que a combinação sequencial de múltiplas famílias não traz benefício adicional.

Por fim, o modelo J (Dual-Branch) atingiu o melhor desempenho geral, com Acc 94,06%, F1 0,939 e AUC 0,936, superando o segundo melhor modelo (G) em +1,25 p.p. de acurácia e +0,012 de F1.

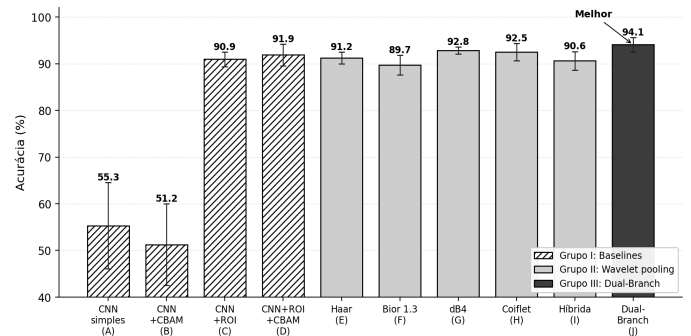


Fig.3. Comparação de acurácia dos dez modelos avaliados.

A Tabela II apresenta sensibilidade e especificidade para o baseline com ROI+CBAM (D), para dois modelos representativos com wavelet pooling (F, G) e para o modelo Dual-Branch (J), utilizando as matrizes de confusão globais (soma dos 5 folds, 320 amostras), destacando o total de falsos negativos (FN) e falsos positivos (FP).

TABELA II

Sensibilidade e especificidade (matrizes globais, 320 amostras).

Modelo	Sens.	Espec.	FN	FP
D: Baseline	0,966	0,835	7	19
F: Bior 1.3	0,937	0,826	13	20
G: dB4	0,985	0,826	3	20
J: Dual-Branch	0,995	0,843	1	18

Em termos absolutos, o modelo A cometeu 143 erros, o baseline D reduziu esse valor para 26, o melhor modelo com wavelet única (G, dB4) errou 23 amostras, e o Dual-Branch (J) atingiu apenas 19 erros. O modelo J alcançou sensibilidade de 99,5% (apenas 1 falso negativo em 205 casos anormais), resultado de grande relevância clínica para rastreamento mamográfico. A Daubechies dB4 (G) também obteve sensibilidade elevada (98,5%, 3 FN), enquanto a Biorthogonal 1.3 (F) apresentou a menor sensibilidade do conjunto (93,7%, 13 FN).

V. DISCUSSÃO

O achado mais relevante deste estudo é que a extração de ROI exerce o maior impacto sobre o desempenho, superando tanto o uso de atenção quanto a escolha da wavelet. Sem ROI, mesmo com CBAM, os modelos A e B operam próximos ao acaso, enquanto a simples adoção de ROI (modelo C) resulta em ganho de +35,63 p.p. em acurácia. Isso evidencia que a delimitação da região de interesse é pré-requisito fundamental

para qualquer estratégia de classificação em mamografia. Nesse contexto, o CBAM atua de forma essencialmente complementar: aplicado sobre a imagem inteira, tende a amplificar ruído; já quando restrito à ROI (C→D), produz um ganho modesto, mas consistente, em acurácia e AUC.

Entre as famílias wavelet utilizadas como pooling, a Daubechies dB4 destaca-se como a opção mais adequada para mamografias, combinando a maior acurácia média (92,81%) com o menor desvio padrão ($\sigma = 0,77$), o que indica estabilidade superior ao longo dos folds. A Coiflet apresenta desempenho muito próximo, sugerindo que ambas as famílias capturam bem padrões texturais relevantes do parênquima mamário, enquanto a Biorthogonal 1.3 apresenta o pior desempenho, possivelmente devido a efeitos indesejados de seus filtros assimétricos no downsampling de ROIs.

O resultado da estratégia híbrida (Modelo I), combinando quatro famílias wavelet sequencialmente, não superou as melhores wavelets individuais, sugerindo que a diversidade de filtros gera inconsistências entre níveis. Isso contrasta com o dual-branch (J), onde cada ramo opera independentemente e a fusão ocorre apenas na classificação.

Por fim, o modelo J (Dual-Branch) alcançou sensibilidade de 99,5%, apenas 1 amostra anormal classificada como normal em 205, um resultado de grande relevância clínica para rastreamento mamográfico, onde falsos negativos têm consequências graves. A combinação de *transfer learning* (ResNet18/ImageNet) com features wavelet, processadas em ramos arquiteturalmente independentes até a etapa de concatenação, permite que cada ramo se especialize: o ramo ResNet captura a estruturas globais da lesão (formas e contornos), enquanto o ramo wavelet Coiflet atua para detectar texturas e padrões locais de frequência. A combinação produz uma representação mais rica que qualquer um dos ramos isoladamente.

Uma limitação importante deste estudo é o tamanho reduzido do mini-MIAS, o que restringe a variabilidade de casos, em particular, a diversidade de densidades mamárias, tipos de lesão e equipamentos de aquisição, e pode inflar a variância das estimativas de desempenho. Bases maiores e mais heterogêneas, como o DDSM [18] (com milhares de exames de múltiplas instituições) e VinDr-Mammo [19] (imagens digitais adquiridas em equipamentos de diferentes fabricantes, com anotações BI-RADS completas), permitiriam avaliar a robustez do dual-branch frente a variações de protocolo de aquisição e a distribuição de classes mais desbalanceadas, mais próximas do cenário real de rastreamento. Além disso, a formulação binária normal vs anormal adotada aqui simplifica o cenário clínico real, em que é fundamental distinguir não apenas a presença de anormalidade, mas também seu caráter benigno ou maligno, bem como o contexto BI-RADS associado.

VI. CONCLUSÕES

Os experimentos deste trabalho levam a quatro achados centrais: (1) a extração de ROI é o componente mais determinante da pipeline, indicando que “onde o modelo olha” importa tanto quanto “como ele aprende”; (2) o CBAM isolado degrada o desempenho, mas combinado com ROI produz ganho consistente, indicando que atenção espacial só é efetiva sobre entradas já focadas na região relevante; (3) entre as famílias wavelet usadas como pooling, a Daubechies dB4

mostrou-se a opção mais eficaz e estável (Acc 92,81%, $\sigma = 0,77$); e (4) a arquitetura dual-branch ResNet18 + Wavelet Coiflet apresentou o melhor desempenho global (Acc 94,06%, F1 0,939, AUC 0,936), alcançando sensibilidade de 99,5%.

A principal contribuição deste estudo consiste na comparação sistemática de dez arquiteturas em um framework unificado, permitindo quantificar o impacto individual de componentes como ROI, CBAM e diferentes estratégias de pooling wavelet na classificação de mamografias. Como desdobramento futuro, pretende-se avaliar a generalização das arquiteturas propostas em bases maiores e mais heterogêneas, mantendo a extração de ROI e a arquitetura dual-branch como componentes centrais da pipeline, devido ao impacto dominante observado neste estudo. Além disso, a wavelet Daubechies dB4 será investigada como candidata ao ramo de frequência, devido à sua maior estabilidade experimental, incluindo estratégias de combinação com a Wavelet Coiflet, que apresentou o melhor resultado no modelo dual-branch. Também serão exploradas estratégias de fusão baseadas em mecanismos de atenção e a extensão da tarefa para classificação multiclasse (normal/benigno/maligno).

AGRADECIMENTOS

Os autores agradecem ao Instituto Federal de Santa Catarina (IFSC) pelo suporte institucional.

Referências

- [1] H. Sung *et al.*, “Global cancer statistics 2020,” *CA Cancer J. Clin.*, v.71, n.3, pp.209–249, 2021.
- [2] R. L. Siegel, K. D. Miller, e A. Jemal, “Cancer statistics, 2020,” *CA Cancer J. Clin.*, v.70, n.1, pp.7–30, 2020.
- [3] N. Wu *et al.*, “Deep neural networks improve radiologists’ performance in breast cancer screening,” *IEEE Trans. Med. Imaging*, v.39, n.4, pp.1184–1194, 2020.
- [4] S. G. Mallat, “A theory for multiresolution signal decomposition,” *IEEE Trans. PAMI*, v.11, n.7, pp.674–693, 1989.
- [5] T. Williams e R. Li, “Wavelet pooling for convolutional neural networks,” em *Proc. ICLR*, 2018.
- [6] P. Liu *et al.*, “Multi-level wavelet-CNN for image restoration,” em *Proc. IEEE/CVF CVPR Workshops*, 2018, pp.773–782.
- [7] S. Woo, J. Park, J.-Y. Lee, e I. S. Kweon, “CBAM: Convolutional block attention module,” em *Proc. ECCV*, 2018, pp.3–19.
- [8] K. He, X. Zhang, S. Ren, e J. Sun, “Deep residual learning for image recognition,” em *Proc. IEEE/CVF CVPR*, 2016, pp.770–778.
- [9] J. Suckling *et al.*, “The mammographic image analysis society digital mammogram database,” *Excerpta Medica, Int. Congress Series*, v.1069, pp.375–378, 1994.
- [10] Q. Li, L. Shen, S. Guo, e Z. Lai, “Wavelet integrated CNNs for noise-robust image classification,” em *Proc. IEEE/CVF CVPR*, 2020, pp.7245–7254.
- [11] D. Ribli, A. Horváth, Z. Unger, P. Pollner, e I. Csabai, “Detecting and classifying lesions in mammograms with deep learning,” *Scientific Reports*, v.8, art.4165, 2018.
- [12] L. Shen *et al.*, “Deep learning to improve breast cancer detection on screening mammography,” *Scientific Reports*, v.9, art.12495, 2019.
- [13] P. F. Hepsağ, Y. A. Özel, e A. Yazıcı, “Using deep learning for mammography classification,” em *Proc. IEEE UBMC*, 2017, pp.418–423.
- [14] M. F. Aslan, K. Sabancı, e A. Durdu, “A hybrid end-to-end learning approach for breast cancer diagnosis,” *Comput. Electr. Eng.*, v.105, art.108527, 2023.
- [15] K. Zuiderveld, “Contrast limited adaptive histogram equalization,” em *Graphics Gems IV*, Academic Press, 1994, pp.474–485.
- [16] A. Cohen, I. Daubechies, e J.-C. Feauveau, “Biorthogonal bases of compactly supported wavelets,” *Commun. Pure Appl. Math.*, v.45, n.5, pp.485–560, 1992.
- [17] I. Loshchilov e F. Hutter, “Decoupled weight decay regularization,” em *Proc. ICLR*, 2019.
- [18] M. Heath, K. Bowyer, D. Kopans, R. Moore, e W. P. Kegelmeyer, “The Digital Database for Screening Mammography,” em *Proc. Int. Workshop on Digital Mammography*, 2000, pp.212–218.
- [19] H. T. Nguyen *et al.*, “VinDr-Mammo: A large-scale benchmark dataset,” *Scientific Data*, v.10, art.277, 2023.
- [20] Paz Faria, B. (2026). Modelos WaveletCNN para Classificação de Imagens Mamográficas (1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.19583030>