

Aplicação de Deep Learning na Detecção de Câncer de Mama em Imagens Histológicas

Murilo Vital Rondina

UTFPR

Apucarana, Brasil

murilorondina@alunos.utfpr.edu.br

Abstract—This study investigates the application of deep learning techniques for detecting invasive ductal carcinoma (IDC) in histopathological breast tissue images. The Breast Histopathology Images dataset, consisting of 300,000 50×50 pixel RGB image patches, was used to train and evaluate two convolutional neural network architectures: LeNet and a custom VGG variant. No data augmentation or class balancing techniques were applied. Results show that the VGG-Custom model outperformed LeNet in all metrics, achieving 88.75% accuracy and an F1-score of 0.7963 for the positive class. The lower false negative rate of the VGG-Custom model highlights its potential for supporting medical diagnosis. The study reinforces the effectiveness of deep learning for low-resolution medical image classification and proposes future improvements using data augmentation and class balancing techniques.

Index Terms—Breast cancer, Deep learning, Histopathology, Convolutional Neural Networks, IDC detection.

I. INTRODUÇÃO

O câncer de mama é uma das doenças mais incidentes e letais entre as mulheres em todo o mundo, representando uma das principais causas de mortalidade feminina. No Brasil, segundo dados do Instituto Nacional de Câncer [1], são estimados mais de 73 mil novos casos da doença para cada ano do triênio 2023-2025, evidenciando a urgência de estratégias mais eficazes para o diagnóstico precoce. A detecção precoce está diretamente associada a maiores taxas de cura e melhor qualidade de vida para as pacientes, sendo a análise histopatológica uma das principais ferramentas diagnósticas utilizadas na prática clínica.

A análise histopatológica, que envolve o exame microscópico de tecidos mamários, é um dos métodos mais utilizados na detecção do câncer de mama. No entanto, esse processo depende de interpretação humana, o que pode ser demorado, subjetivo e suscetível a erros [2]. Nesse contexto, técnicas baseadas em Inteligência Artificial (IA), especialmente o aprendizado profundo (deep learning), têm se destacado como ferramentas promissoras no apoio ao diagnóstico médico, permitindo a identificação automática de padrões em imagens médicas com alto grau de acurácia [3].

Em particular, redes neurais convolucionais (CNNs) vêm sendo amplamente aplicadas em tarefas de classificação de imagens histológicas. Contudo, um dos desafios encontrados nesses conjuntos de dados é a baixa resolução das imagens, como no caso do conjunto IDC Breast Histopathology Dataset [4], que possui amostras de apenas 50×50 pixels. Essa

limitação pode comprometer a identificação de estruturas celulares sutis, afetando negativamente o desempenho do modelo.

II. TRABALHOS RELACIONADOS

A detecção de câncer de mama por meio de imagens histopatológicas é um dos principais desafios enfrentados na área de diagnóstico auxiliado por computador. Diversos estudos têm explorado o uso de algoritmos de aprendizado profundo para identificar padrões malignos em imagens de biópsias, contribuindo para diagnósticos mais rápidos e precisos.

Janowczyk e Madabhushi apresentam um estudo abrangente sobre o uso de aprendizado profundo na análise de imagens histológicas, destacando a aplicabilidade de redes neurais convolucionais (CNNs) em tarefas de segmentação e classificação de tecidos tumorais [5]. O trabalho enfatiza a importância do pré-processamento de imagens e do uso de anotações especializadas, ressaltando que a combinação entre arquitetura de rede e qualidade do dado é essencial para modelos robustos.

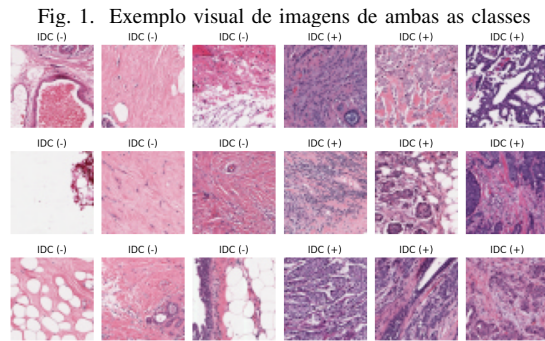
Adicionalmente, Mooney publicou um projeto na plataforma Kaggle que demonstra a viabilidade de se aplicar CNNs simples ao conjunto de dados IDC Breast Histopathology. O estudo reforça que mesmo redes com arquiteturas compactas podem alcançar bons resultados quando combinadas com técnicas de balanceamento de classes e normalização de dados [6].

Esses estudos reforçam a relevância das redes neurais profundas na detecção automatizada de câncer de mama, bem como apontam caminhos promissores para o uso combinado de super-resolução e classificação em imagens médicas de baixa resolução.

III. METODOLOGIA

A. Dataset

O conjunto de dados utilizado neste estudo é o Breast Histopathology Images, disponível na plataforma Kaggle, composto originalmente por 555.048 imagens histológicas de 50×50 pixels. Estas imagens são divididas entre amostras negativas (ausência de carcinoma ductal invasivo) e positivas (presença de carcinoma ductal invasivo - IDC), Figura 1.

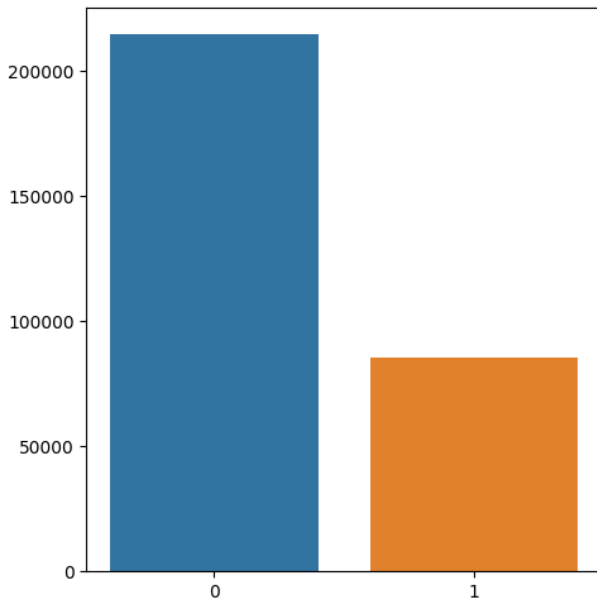


B. Pré-processamento

As imagens foram mantidas no tamanho original (50x50 pixels) com três canais RGB. Durante o pré-processamento, foi realizada a divisão entre treino e teste com preservação da proporção de classes, garantindo a manutenção do desbalanceamento característico do conjunto. Não foram aplicadas técnicas adicionais de normalização ou balanceamento de classes. O foco foi avaliar o desempenho das arquiteturas considerando a base reduzida, mas ainda desbalanceada.

Devido ao grande volume de dados, foi realizada uma redução no número de amostras para viabilizar o treinamento em tempo hábil e com menor custo computacional. O dataset foi reduzido para 300.000 imagens, mantendo a seguinte proporção: 214.696 imagens negativas e 85.304 imagens positivas, Figura 2.

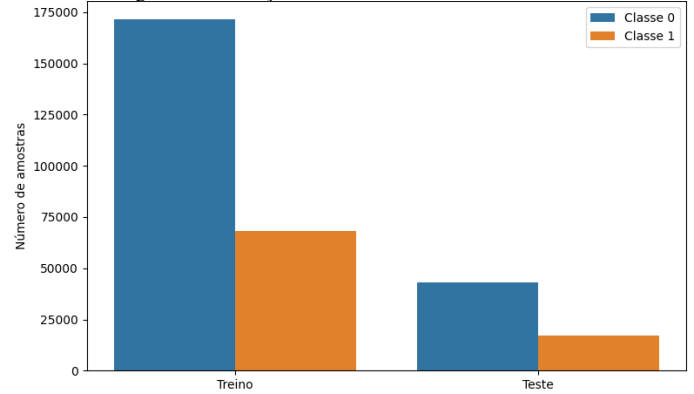
Fig. 2. Distribuição das classes no dataset



Para o treinamento e avaliação, os dados foram divididos da seguinte forma, Figura 3:

- Treinamento: 240.000 imagens (171.757 negativas e 68.243 positivas);
- Teste: 60.000 imagens (42.939 negativas e 17.061 positivas).

Fig. 3. Distribuição das classes entre teste e treino



C. Arquiteturas Testadas

Para uma melhor compreensão estrutural das redes utilizadas, a Tabela I e a Tabela II apresentam, respectivamente, o detalhamento camada por camada das arquiteturas LeNet e VGG-Custom implementadas neste estudo. Foram incluídas informações sobre o tipo de camada, a dimensão da saída (output shape) e o número total de parâmetros treináveis.

A LeNet, composta por 8 camadas ocultas (2 convolucionais, 2 de pooling, 2 densas e 2 dropout), possui 204 neurônios nas camadas densas ocultas e aproximadamente 204 mil parâmetros treináveis.

Já a arquitetura VGG-Custom apresenta 10 camadas ocultas, com 6 camadas convolucionais organizadas em blocos duplos, 3 camadas de pooling, uma camada densa oculta com 512 neurônios e uma camada dropout, totalizando cerca de 2,63 milhões de parâmetros treináveis.

A escolha por utilizar blocos com convoluções duplas na VGG-Custom segue o padrão da arquitetura VGG original, permitindo maior profundidade e capacidade de extração de padrões, mesmo utilizando filtros pequenos (3x3). Essa abordagem favorece a generalização do modelo com menor custo computacional em comparação a filtros maiores, além de introduzir mais não-linearidades úteis no processo de aprendizagem.

As tabelas a seguir detalham cada arquitetura Tabela I, II:

TABLE I
ARQUITETURA LENET

Layer	Acurácia	Precisão
convolucional	(None, 50, 50, 6)	456
pooling	(None, 25, 25, 6)	0
convolucional	(None, 21, 21, 16)	2,416
pooling	(None, 10, 10, 16)	0
flatten	(None, 1600)	0
dense	(None, 120)	192,120
dropout	(None, 120)	0
dense	(None, 84)	10,164
dropout	(None, 84)	0
dense	(None, 1)	85

TABLE II
ARQUITETURA VGG-CUSTOM

Layer	Acurácia	Precisão
convolucional	(None, 50, 50, 32)	896
convolucional	(None, 50, 50, 32)	9,248
pooling	(None, 25, 25, 32)	0
convolucional	(None, 25, 25, 64)	18,496
convolucional	(None, 25, 25, 64)	36,928
pooling	(None, 12, 12, 64)	0
convolucional	(None, 12, 12, 128)	73,856
convolucional	(None, 12, 12, 128)	147,584
pooling	(None, 6, 6, 128)	0
flatten	(None, 4608)	0
dense	(None, 512)	2,359,808
dropout	(None, 512)	0
dense	(None, 1)	513

D. Treinamento e Avaliação

Ambos os modelos foram treinados utilizando a divisão 80/20 (treino/teste). As métricas avaliadas foram Acurácia, Precisão, Recall, F1-score. Também foram geradas as matrizes de confusão para análise detalhada dos erros de classificação.

Os treinamentos foram conduzidos utilizando o framework TensorFlow, com ajuste básico de hiperparâmetros, sem aplicação de técnicas avançadas como data augmentation ou balanceamento artificial. Os gráficos de Loss e Accuracy do treinamento apresentadas nas Figuras 4, 5.

Fig. 4. Accuracy e Loss - LeNet

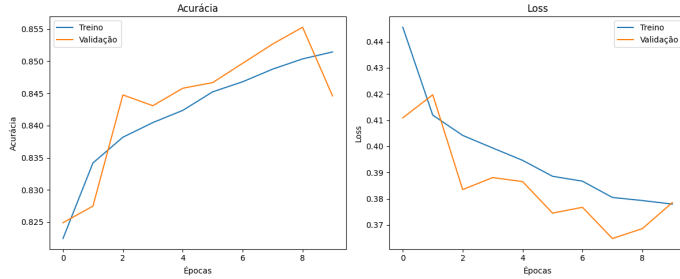
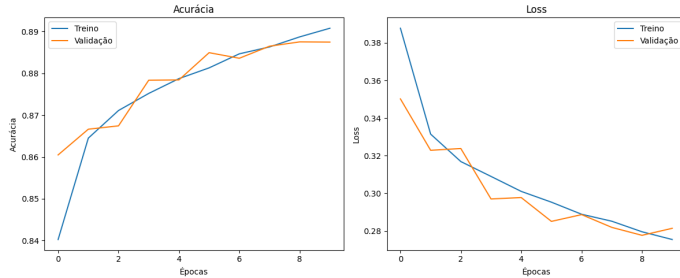


Fig. 5. Accuracy e Loss - VGG



IV. RESULTADOS E DISCUSSÃO

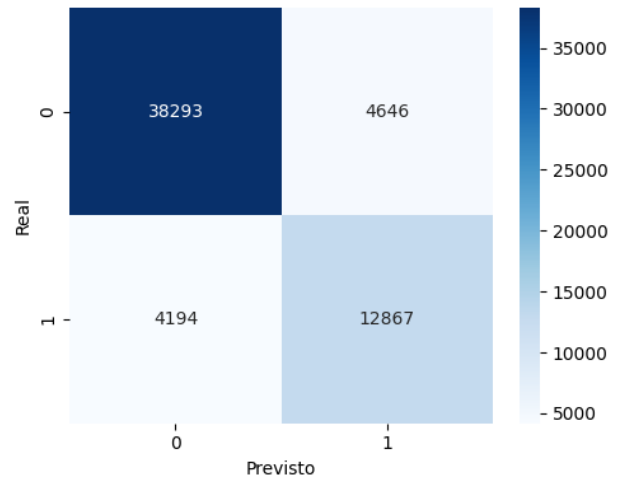
A. LeNet

O modelo LeNet apresentou uma acurácia de 85,27% na base de teste, com um F1-score de 0,744 para a classe positiva.

O recall de 75,42% indica que aproximadamente 24,58% das amostras positivas não foram corretamente identificadas, o que é um ponto crítico para aplicações médicas onde a sensibilidade é fundamental.

A matriz de confusão Figura 6 evidencia um número considerável de falsos negativos, o que pode comprometer a eficácia do modelo na prática clínica. Além disso, o modelo mostrou menor capacidade de generalização em relação ao VGG-Custom.

Fig. 6. Matriz de confusão - LeNet



B. VGG-Custom

O modelo VGG-Custom obteve um desempenho superior, alcançando uma acurácia de 88,75% e um F1-score de 0,796 para a classe positiva. O recall foi de 77,32%, indicando uma maior sensibilidade na detecção do carcinoma ductal invasivo.

Além disso, o VGG-Custom apresentou melhor equilíbrio entre precisão e recall, com menor quantidade de falsos negativos em comparação à LeNet Figura 7. Este resultado demonstra que arquiteturas mais profundas e complexas podem extrair melhor os padrões relevantes mesmo em imagens de baixa resolução.

C. Comparação

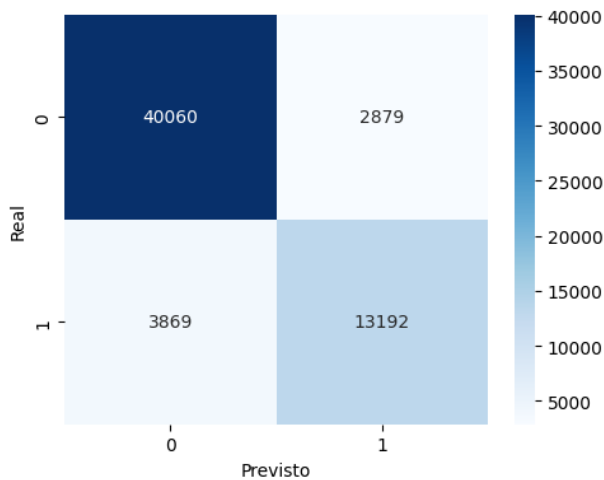
A Tabela III resume as métricas de desempenho de ambas as arquiteturas:

TABLE III
MÉTRICAS DE DESEMPENHO DOS MODELOS LeNET E VGG-CUSTOM

Arquitetura	Acurácia	Precisão	F1-score
LeNet	85,27%	0,7347	0,7443
VGG-Custom	88,75%	0,8209	0,7963

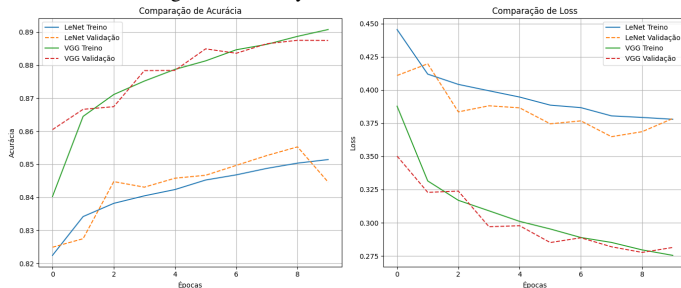
Os resultados confirmam que a arquitetura VGG-Custom foi mais eficiente na tarefa de classificação, sendo mais adequada para apoiar diagnósticos médicos baseados em imagens histológicas de baixa resolução.

Fig. 7. Matriz de confusão - VGG



Os gráficos de acurácia e loss ao longo das épocas de treinamento reforçam a melhor convergência do modelo VGG-Custom, que apresentou maior estabilidade e menor overfitting em comparação à LeNet Figura 8.

Fig. 8. Accuracy e Loss - LeNet vs VGG



D. Análise dos Resultados

Os resultados obtidos demonstram a eficácia de arquiteturas convolucionais na detecção de carcinoma ductal invasivo (IDC) em imagens histológicas. A VGG-Custom, por apresentar maior profundidade, superou a LeNet em todas as métricas, especialmente no F1-score (0,7963 contra 0,7443), o que reflete um melhor equilíbrio entre precisão e recall.

A VGG-Custom apresentou uma taxa significativamente menor de falsos negativos. Esse resultado é especialmente relevante em contextos médicos, onde a sensibilidade diagnóstica é fundamental. Mesmo que a acurácia global da LeNet (85,27%) seja relativamente boa, o risco de não detectar um caso positivo pode inviabilizar seu uso clínico direto.

Observa-se ainda que ambas as arquiteturas foram treinadas em um dataset desbalanceado, o que impacta diretamente o recall das classes minoritárias. A ausência de técnicas de balanceamento, como oversampling, undersampling ou ponderação de classes, pode ter limitado o desempenho, principalmente da LeNet.

Além disso, os gráficos de loss e accuracy evidenciam uma maior estabilidade e menor overfitting no modelo VGG-Custom, sugerindo que a arquitetura mais profunda foi capaz de generalizar melhor os padrões de malignidade presentes nas imagens.

Outro ponto importante refere-se à simplicidade dos modelos utilizados, o que reforça a viabilidade de soluções eficientes mesmo com recursos computacionais limitados. No entanto, há margem para melhorias com técnicas como data augmentation e uso de super-resolução, conforme proposto na metodologia.

Por fim, os resultados obtidos validam o potencial do uso de CNNs como ferramentas auxiliares no diagnóstico histopatológico, destacando a importância de um projeto de arquitetura alinhado às características dos dados.

V. CONCLUSÃO E TRABALHOS FUTUROS

Este estudo demonstrou que o uso de redes neurais convolucionais pode contribuir significativamente para a detecção de carcinoma ductal invasivo em imagens histológicas de baixa resolução. Os resultados obtidos evidenciaram que arquiteturas mais profundas, como a VGG-Custom, são capazes de oferecer melhor desempenho, principalmente na redução de falsos negativos, o que é essencial para aplicações médicas.

Trabalhos futuros, suger-se:

- Aplicação de técnicas de data augmentation para aumentar a robustez dos modelos;
- Testes com arquiteturas mais complexas, que podem oferecer ganhos adicionais em desempenho.
- Implementação de métodos de balanceamento de classes para reduzir o impacto do desbalanceamento observado no dataset.

O desenvolvimento de modelos cada vez mais precisos e sensíveis é fundamental para que ferramentas baseadas em deep learning possam, no futuro, integrar fluxos de diagnóstico médico, contribuindo para o aumento da assertividade e a redução de diagnósticos tardios.

REFERÊNCIAS

- [1] INSTITUTO NACIONAL DE CÂNCER (Brasil). Estimativa 2023: incidência de câncer no Brasil. Rio de Janeiro: INCA, 2022. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/estimativa-2023.pdf>. Acesso em: 03 maio 2025.
- [2] ELMORE, J. G. et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA, v. 313, n. 11, p. 1122–1132, 2015. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25781441/>. Acesso em: 03 maio 2025.
- [3] LITJENS, G. et al. A survey on deep learning in medical image analysis. Medical Image Analysis, v. 42, p. 60–88, 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28778026/>. Acesso em: 03 maio 2025.
- [4] MOONEY, P. Breast Histopathology Images and IDC Detection with CNN. Kaggle, 2018. Disponível em: <https://www.kaggle.com/code/paultimothymooney/predict-idc-in-breast-cancer-histology-images/notebook>. Acesso em: 03 maio 2025.
- [5] JANOWCZYK, A.; MADABHUSHI, A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. Computers in Biology and Medicine, [S. l.], 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2153353922005478>. Acesso em: 03 maio 2025.
- [6] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016. Disponível em: <https://www.deeplearningbook.org>. Acesso em: 03 maio 2025.

- [7] CHOLLET, F. Deep Learning with Python. 2. ed. Shelter Island: Manning Publications, 2021.
- [8] GERON, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2. ed. Sebastopol: O'Reilly Media, 2019.
- [9] RUSSELL, S. J.; NORVIG, P. Inteligência Artificial. Uma Abordagem Moderna. 4. ed. Rio de Janeiro: LTC, 2022.
- [10] ROJAS, R. Neural Networks: A Systematic Introduction. Berlin: Springer-Verlag, 1996.
- [11] GONZÁLEZ, R. C.; WOODS, R. E. Digital Image Processing. 4. ed. New York: Pearson, 2018.