



MODELOS PREDITIVOS PARA DIAGNÓSTICO NEUROLÓGICO: CLASSIFICAÇÃO DE IMAGENS CEREBRAIS EM TUMOR E NORMAL POR VISÃO COMPUTACIONAL

Matheus Machado da Silva e Souza¹, Luis Fernando Bueno¹, Lucas Rafael Gonçalves Dolenkei¹, Gabriel Passos de Jesus^{1,2,3}

¹*Centro de Ensino Superior de Maringá, Ponta Grossa, Brasil
(matheusmachado4971@gmail.com)*

²*Pontifícia Universidade Católica do Paraná, Curitiba, Brasil*

³*Universidade Federal do Paraná, Curitiba, Brasil*

Resumo: Este trabalho investiga modelos preditivos para apoio ao diagnóstico neurológico por visão computacional, com foco na classificação binária de imagens de ressonância magnética cerebral em normal e tumor. Em experimento conduzido no Google Colab com base pública do Kaggle, compararam-se uma CNN basal e a EfficientNetB3 com transferência de aprendizado. O melhor modelo atingiu acurácia de 88,46% e ROC-AUC de 0,95, indicando maior potencial para triagem assistida por IA.

Palavras-chave: Visão computacional; neurologia; medicina; redes neurais; aprendizado de máquina

INTRODUÇÃO

A aplicação de inteligência artificial à saúde tem ampliado as possibilidades de apoio ao diagnóstico por imagem, sobretudo em tarefas repetitivas e sensíveis à variação interobservador. No contexto neurológico, exames de ressonância magnética (RM) oferecem alta riqueza anatômica, porém sua interpretação demanda tempo, experiência especializada e protocolos consistentes. Nesse cenário, a visão computacional surge como estratégia promissora para extrair padrões discriminativos e auxiliar a identificação precoce de alterações cerebrais.

Na literatura recente, métodos baseados em redes neurais convolucionais e transferência de aprendizado têm alcançado desempenho competitivo na classificação de tumores cerebrais em RM, especialmente quando há limitação de dados rotulados e necessidade de reaproveitar representações aprendidas em bases amplas (SWATI et al., 2019; DEEPAK; AMEER, 2019; RASA et al., 2024). Estudos mais recentes também apontam ganhos com arquiteturas EfficientNet, fusão multimodal e estratégias de interpretabilidade, reforçando a tendência de modelos mais robustos e mais próximos de cenários clínicos reais (MALEK et al., 2023; ZAHOR et al., 2024; ULLAH et al., 2024; FRIEDRICH et al., 2025).

Este trabalho apresenta um experimento preliminar de classificação binária de imagens cerebrais em tumor e normal, comparando uma CNN baseline com um

modelo EfficientNetB3 por transferência de aprendizado. O objetivo é analisar, em ambiente reprodutível de baixo custo computacional, o impacto da escolha da arquitetura sobre métricas clássicas de desempenho, discutindo o potencial e as limitações desse tipo de abordagem para apoio ao diagnóstico neurológico.

MATERIAL E MÉTODOS

O experimento foi conduzido no Google Colab, por oferecer ambiente gratuito com suporte a GPU, bibliotecas de aprendizado profundo e reprodutibilidade adequada para pesquisa acadêmica exploratória. Foi utilizada uma base pública de imagens de RM cerebral disponível no Kaggle, contendo exemplos de cérebros saudáveis e com presença de tumor, apropriada para tarefas de classificação supervisionada.

Para este resumo expandido, adotou-se um recorte binário das classes, reorganizando as imagens em dois grupos: Normal e Tumor. O fluxo metodológico envolveu: (i) carregamento e inspeção visual das imagens; (ii) redimensionamento e normalização; (iii) separação em conjuntos de treino e teste; (iv) treinamento de dois modelos; e (v) avaliação por métricas de classificação. O primeiro modelo foi uma CNN baseline, construída como referência inicial. O segundo foi a EfficientNetB3 com transferência de aprendizado, estratégia recomendada na literatura para melhorar generalização em bases médicas de tamanho moderado (SWATI et al., 2019; RASA et al., 2024).

As métricas analisadas foram acurácia, precisão, recall, F1-Score e ROC-AUC, além do relatório por classe. Tais medidas são relevantes porque, em aplicações em saúde, o acerto global não é suficiente: é necessário observar o equilíbrio entre sensibilidade à detecção de casos com tumor e especificidade para evitar falsos alarmes. Assim, a interpretação dos resultados considerou não apenas a acurácia, mas também a capacidade do modelo em recuperar corretamente os casos positivos.

RESULTADOS E DISCUSSÃO

A Tabela 1 resume os resultados preliminares obtidos. A CNN baseline apresentou desempenho moderado, com acurácia de 61,54% e ROC-AUC de 0,6243. Embora o recall para a classe Tumor tenha sido relativamente alto (0,8077), o modelo mostrou dificuldade em reconhecer adequadamente a classe Normal, com recall de apenas 0,42. Esse comportamento sugere viés para a identificação de casos positivos, o que pode ser útil em triagem inicial, mas ainda insuficiente para uso confiável.

Tabela 1. Resultados da CNN Baseline

Métrica	Resultado
Acurácia	0,61
Precisão	0,58
Recall	0,80
F1-Score	0,67
ROC-AUC	0,62

Elaboração. Os autores.

Já a EfficientNetB3 superou de forma consistente a abordagem basal em todas as métricas, alcançando acurácia de 88,46%, F1-Score de 0,8889 e ROC-AUC de 0,9556. Também houve maior equilíbrio entre as classes: para Normal, precisão de 0,92 e recall de 0,85; para Tumor, precisão de 0,86 e recall de 0,92. Esses achados estão alinhados com trabalhos recentes que mostram a superioridade de modelos pré-treinados e variantes EfficientNet em tarefas de classificação de RM cerebral, devido à melhor eficiência no aproveitamento de parâmetros e à maior capacidade de extração de atributos relevantes (MALEK et al., 2023; RASA et al., 2024).

Tabela 2. Resultados da CNN EfficientNetB3

Métrica	Resultado
Acurácia	0,88
Precisão	0,85
Recall	0,92
F1-Score	0,88
ROC-AUC	0,95

Elaboração. Os autores.

Do ponto de vista de pesquisa, os resultados indicam que a transferência de aprendizado é um caminho mais promissor do que uma CNN treinada do zero para o cenário investigado. Em termos aplicados, a elevada sensibilidade do modelo EfficientNetB3 é desejável em tarefas de apoio à triagem, pois reduz a chance de não detectar exames com tumor. Ainda assim, é necessário cautela: a base utilizada é pública, o conjunto de teste é reduzido e os resultados não substituem validação clínica. Revisões recentes na interface entre visão computacional e neurologia destacam justamente que desempenho experimental elevado não implica adoção imediata em ambiente assistencial, sendo indispensáveis validação externa, explicabilidade e avaliação multicêntrica (OZKARA et al., 2023; FRIEDRICH et al., 2025).

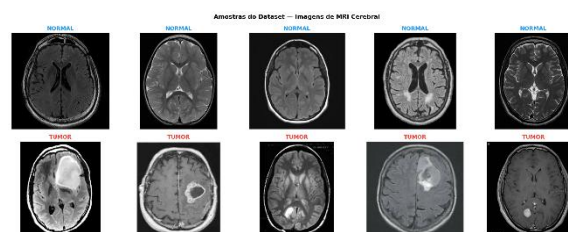


Figura 1. Comparativo entre as imagens na CNN Baseline e na CNN EfficientNetB3

Do ponto de vista de pesquisa, os resultados indicam que a transferência de aprendizado é um caminho mais promissor do que uma CNN treinada do zero para o cenário investigado. Em termos aplicados, a elevada sensibilidade do modelo EfficientNetB3 é desejável em tarefas de apoio à triagem, pois reduz a chance de não detectar exames com tumor. Ainda assim, é necessário cautela: a base utilizada é pública, o conjunto de teste é reduzido e os resultados não substituem validação clínica. Revisões recentes na interface entre visão computacional e neurologia destacam justamente que desempenho experimental elevado não implica adoção imediata em ambiente assistencial, sendo indispensáveis validação externa, explicabilidade e avaliação multicêntrica (OZKARA et al., 2023; FRIEDRICH et al., 2025).

CONCLUSÃO

Os resultados preliminares mostram que a EfficientNetB3 com transferência de aprendizado apresentou desempenho substancialmente superior ao da CNN baseline na classificação binária de imagens cerebrais em tumor e normal. O ganho observado em acurácia, F1-Score e ROC-AUC reforça o potencial de arquiteturas pré-treinadas para aplicações de apoio ao diagnóstico neurológico por imagem. Em termos clínicos, esse achado é relevante porque o erro médico em diagnóstico por imagem pode ocorrer tanto pela não identificação de lesões verdadeiras quanto pela interpretação excessiva de achados inespecíficos,



produzindo atrasos terapêuticos, ansiedade para pacientes e familiares, solicitação de exames adicionais e até intervenções desnecessárias. Nesse sentido, modelos computacionais mais estáveis podem funcionar como apoio à decisão e segunda leitura, contribuindo para reduzir falhas perceptuais e variabilidade entre observadores, sem eliminar a responsabilidade clínica do especialista. Assim, a principal contribuição deste estudo não é propor substituição da conduta médica, mas indicar que abordagens híbridas, nas quais a expertise do profissional é complementada por ferramentas de IA, podem aumentar a confiabilidade diagnóstica e favorecer decisões mais seguras, especialmente em contextos de triagem e alta demanda assistencial.

Como continuidade do projeto, pretende-se ampliar o protocolo experimental com validação cruzada, técnicas de aumento de dados, ajuste de hiperparâmetros, mapas de ativação para interpretabilidade e eventual comparação com arquiteturas mais recentes. Também será importante discutir aspectos éticos, reprodutibilidade e limites de generalização, de modo que o estudo permaneça alinhado ao uso responsável de IA na saúde. Além disso, trabalhos futuros devem considerar de forma mais explícita como falsos positivos e falsos negativos se traduzem em risco clínico, uma vez que a utilidade real do modelo depende não apenas de métricas agregadas, mas de seu impacto potencial sobre o erro médico, o fluxo de atendimento e a segurança do paciente.

AGRADECIMENTOS

Os autores agradecem nominalmente à Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação da UniCesumar; ao Programa de Pós-Graduação em Informática da PUCPR; ao Programa de Pós-Graduação em Medicina Interna da UFPR; ao Programa de Pós-Graduação em Clínica Cirúrgica da UFPR; ao Programa de Pós-Graduação em Computação Aplicada da UEPG; ao Prof. Dr. Leonardo Camargo, do curso de Medicina da PUCPR - Londrina; à Dra. Maria de Fátima Domingos (FSP/USP); aos Prof. Dr. Felipe de Lara Janz, Prof. MSc. Josyel Olszewski, Prof. Dr. Wanderson Ribeiro e Prof. Dr. Fábio Brasil de Oliveira, à Dra. Danielly Augusta Ramos Felippi, à Dra. Camila Ceruti e ao Dr. Miguel Fabrin, e à acadêmica Ester Souza, do curso de Medicina da UFV, pelo apoio institucional, pelas contribuições acadêmicas e pelo incentivo ao desenvolvimento deste trabalho.

REFERÊNCIAS

BREMBILLA, G.; TAKWOINGI, Y.; KASIVISVANATHAN, V. Tackling interobserver variability in multiparametric magnetic resonance imaging (MRI): is MRI even better than we think for prostate cancer diagnosis? *European Urology*, v. 79, n. 1, p. 8--10, 2021.

DEEPAK, S.; AMEER, P. M. Brain tumor classification using deep CNN features via transfer learning. *Computerized Medical Imaging and Graphics*, v. 111, p. 103345, 2019.

FRIEDRICH, M. U.; RELTON, S.; WONG, D.; ALTY, J. Computer vision in clinical neurology: a review. *JAMA Neurology*, 2025.

HAJIAHMADI, S. et al. Interobserver and intraobserver variability in magnetic resonance imaging evaluation of patients with suspected disc herniation. *Heliyon*, v. 6, n. 11, e05201, 2020.

HUSSIEN, A. R. et al. Diagnostic errors in neuroradiology: a message to emergency radiologists and trainees. *Canadian Association of Radiologists Journal*, v. 73, n. 2, p. 384--395, 2022.

MALEK, N.; MOHAMED, M.; ELHALAWANY, B.; et al. Multi-class classification of brain tumor types from MR images using EfficientNets. *Biomedical Signal Processing and Control*, v. 84, p. 104777, 2023

OZKARA, B. B.; CHEN, M. M.; FEDERAU, C.; et al. Deep learning for detecting brain metastases on MRI: a systematic review and meta-analysis. *Cancers*, v. 15, n. 2, p. 334, 2023.

RASA, S. M. et al. Brain tumor classification using fine-tuned transfer learning models on magnetic resonance imaging (MRI) images. *DIGITAL HEALTH*, v. 10, 2024.

SWATI, Z. N. K. et al. Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, v. 75, p. 34--46, 2019.

SUBRAMANIAN, H. V.; CANFIELD, C.; SHANK, D. B. Designing explainable AI to improve human-AI team performance: a medical stakeholder-driven scoping review. *Artificial Intelligence in Medicine*, v. 149, p. 102780, 2024.

TOPFF, L. et al. Artificial intelligence-assisted double reading of chest radiographs to detect clinically relevant missed findings: a two-centre evaluation. *European Radiology*, v. 34, n. 9, p. 5876--5885, 2024.

ULLAH, M. S. et al. Multimodal brain tumor segmentation and classification from MRI scans based on optimized DeepLabV3+ and interpreted networks information fusion empowered with explainable AI. *Computers in Biology and Medicine*, v. 182, p. 109183, 2024.

ZAHOOR, M. M. et al. Brain tumor MRI classification using a novel deep residual and regional CNN. *Biomedicines*, v. 12, n. 7, p. 1395, 2024.