Journal of Health Informatics

J Health Inform. 2026;18:e1891
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1891

ARTIGO COMPLETO
CBIS 2026

Detecção e classificação de tumores cerebrais em imagens de ressonância magnética, utilizando rede neural convolucional "leve"

Detection and classification of brain tumors in magnetic resonance images using a lightweight convolutional neural network

Leon Pereira Pinto Fagundes[1], Marcia Ito[2], Waldemar Bonventi Jr.[2] ✉

ORCID: 0009-0009-7205-0350 · 0000-0003-4799-2433 · 0000-0003-0720-499X

[1] Faculdade de Tecnologia de Votorantim, Votorantim (SP), Brasil.

[2] Centro Estadual de Educação Tecnológica Paula Souza, São Paulo (SP), Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1891

DOI
10.59681/2175-4411.v18.iEspecial.2026.1891

RECEBIDO
17/05/2026

APROVADO
20/07/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
Waldemar Bonventi Jr.
waldemar.bonventi@cps.sp.gov.br

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

Introdução: A correta classificação dos tumores cerebrais é fundamental para a definição do plano terapêutico e do prognóstico. **Objetivo:** Mostrar a viabilidade do uso de Redes Neurais Convolucionais, em especial a MobileNetV2, para classificar tumores cerebrais em imagens de ressonância magnética obtidas em vários planos. **Método:** A rede foi treinada com 5.619 imagens e testada em outras 1.404. **Resultados:** Obteve-se precisão muito alta para glioma, tumor pituitário e sem tumor (>0,95) e mais baixa para meningioma (0,80), que se confunde mais com o glioma. Foi publicada uma interface para apreciação e aperfeiçoamento do sistema. **Conclusão:** Arquiteturas leves, como a MobileNetV2, podem alcançar desempenho competitivo na detecção multiclasse de tumores cerebrais.

Descritores: neoplasias encefálicas; processamento de imagem assistida por computador; redes neurais de computação.

ABSTRACT

Introduction: Accurate classification of brain tumors is fundamental to determining the therapeutic plan and prognosis. **Objective:** To demonstrate the feasibility of using Convolutional Neural Networks, especially MobileNetV2, to classify brain tumors in magnetic resonance images obtained in various planes. **Method:** The network was trained with 5,619 images and tested on another 1,404. **Results:** Precision was very high for glioma, pituitary tumor, and no tumor (>0.95), and lower for meningioma (0.80), which is more easily confused with glioma. An interface for system evaluation and improvement was published. **Conclusion:** Lightweight architectures, such as MobileNetV2, can achieve competitive performance in multiclass brain tumor detection.

Keywords: brain neoplasms; image processing, computer-assisted; neural networks, computer.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Artigo apresentado no XXI Congresso Brasileiro de Informática em Saúde (CBIS 2026), realizado de 23 a 25 de setembro de 2026, em Brasília (DF), e publicado na seção CBIS 2026 do Journal of Health Informatics.

COMO CITAR ESTE ARTIGO: Fagundes LPP, Ito M, Bonventi Jr. W. Detecção e classificação de tumores cerebrais em imagens de ressonância magnética, utilizando rede neural convolucional "leve". J Health Inform 2026;18:e1891. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1891

1 INTRODUÇÃO

Os tumores cerebrais representam um importante problema de saúde pública, com impacto funcional e socioeconômico expressivo [1]. A correta classificação da lesão é fundamental para a definição do plano terapêutico e do prognóstico. Embora a ressonância magnética (RM) seja o método de referência para avaliação não invasiva de massas intracranianas, sua interpretação depende fortemente da experiência do radiologista e pode apresentar variabilidade interobservador [2].

Nas últimas décadas, o aprendizado profundo (*deep learning*), notadamente as redes neurais convolucionais (CNNs), mostrou desempenho de ponta em tarefas de visão computacional, incluindo a análise de imagens médicas. Estudos prévios evidenciam acurácia elevada na detecção e classificação de tumores cerebrais por RM. Persistem, contudo, desafios: bases restritas, cenários binários (tumor versus sem tumor), heterogeneidade de aquisição e a dependência de hiperparâmetros/arquiteturas [3].

Neste contexto, o presente trabalho investiga não apenas a aplicação de um modelo baseado em MobileNetV2 para classificação de tumores cerebrais, mas sobretudo busca responder a uma questão central: até que ponto uma arquitetura leve, otimizada para dispositivos de baixa capacidade computacional, pode atingir desempenho clinicamente relevante em um problema complexo como a detecção de tumores intracranianos? Assim, além de apresentar um *pipeline* completo de previsão, avaliamos a viabilidade técnica e prática do uso de CNNs compactas como ferramentas de triagem e apoio ao diagnóstico. Para isto, foi desenvolvida uma API (*Application Programming Interface*, Interface de Programação de Aplicação) e disponibilizada publicamente para avaliação do classificador de tumores cerebrais.

1.1 Objetivos

Geral: Treinar uma rede neural leve com imagens de tumores cerebrais e disponibilizar uma interface aos usuários (API) do sistema que a contém, para apoio ao diagnóstico médico.

Específicos:

• Mostrar que arquiteturas leves de redes neurais convolucionais, como a MobileNetV2, são capazes de alcançar desempenho competitivo na classificação multiclasse de tumores cerebrais em imagens de ressonância magnética, mesmo quando treinadas em ambientes computacionais sem uso de unidade de processamento gráfico (GPU);

• Desenvolver um modelo de CNN baseado na arquitetura MobileNetV2 utilizando técnicas de transferência de aprendizado e *fine-tuning*;

• Avaliar o desempenho do modelo por meio de métricas quantitativas, como acurácia, precisão, *recall*, F1-*score* e área sob a curva ROC (AUC-ROC);

• Analisar qualitativamente os principais padrões de acerto e erro do modelo, à luz de dificuldades clínicas reais;

• Investigar a estabilidade do treinamento e a viabilidade do uso do modelo em ambientes com recursos computacionais limitados.

1.2 Justificativa

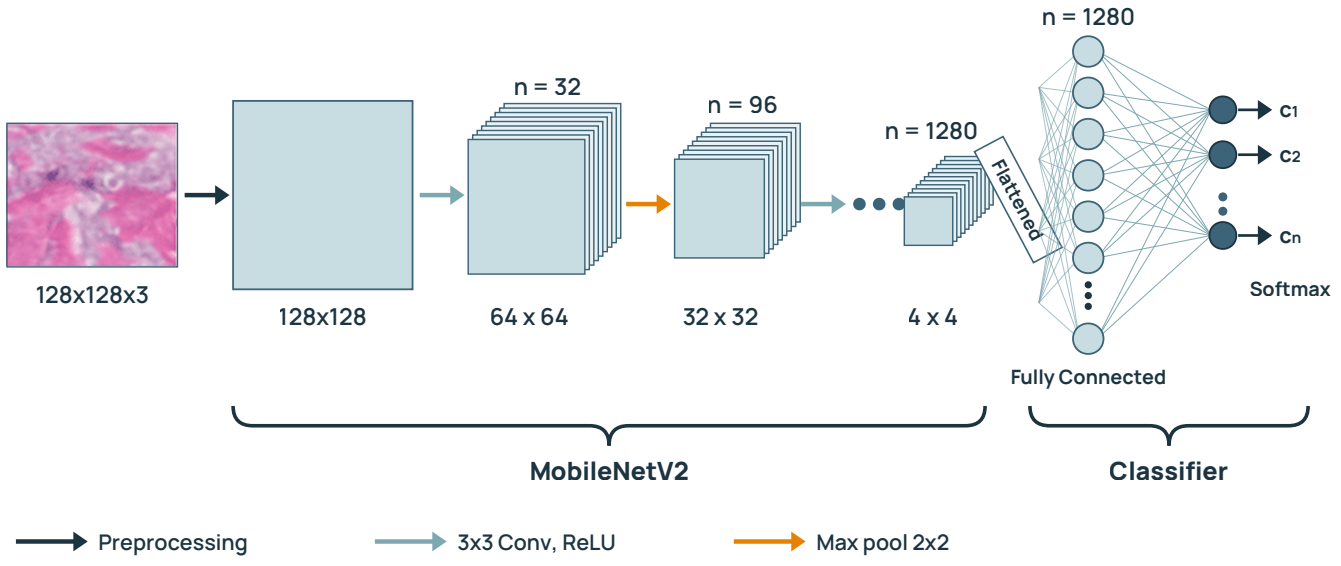
A aplicação de técnicas de Inteligência Artificial na área da saúde tem potencial para ampliar o acesso a ferramentas de apoio ao diagnóstico, sobretudo em contextos caracterizados por alta demanda de exames e escassez de especialistas. No caso específico dos tumores cerebrais, a complexidade morfológica das lesões e a variabilidade dos exames tornam o diagnóstico por imagem uma tarefa desafiadora.

A maior parte dos estudos existentes utiliza arquiteturas profundas e computacionalmente onerosas, o que dificulta sua adoção em ambientes institucionais com infraestrutura limitada. A escolha da MobileNetV2 decorre de seu baixo custo computacional e da possibilidade de ser treinada e implantada em equipamentos de baixa potência. Tal característica é fundamental para validar a hipótese central deste estudo: que modelos compactos são suficientes para demandas clínicas reais, desde que estruturados com boas práticas de treinamento, como transferência de aprendizado e *fine-tuning* controlado. Assim, justifica-se a investigação de modelos leves, como a MobileNetV2, que conciliem eficiência computacional e desempenho robusto. Este projeto contribui para a área ao demonstrar que soluções compactas podem ser eficazes, desde que acompanhadas de boas práticas de treinamento e validação.

1.3 Trabalhos Correlatos

Diversos estudos têm explorado o uso de técnicas de aprendizado profundo na classificação automática de imagens de ressonância magnética de tumores cerebrais. Essas abordagens, majoritariamente baseadas em redes neurais convolucionais, têm demonstrado elevada eficácia na distinção entre diferentes tipos tumorais, como gliomas, meningiomas, adenomas pituitários e cérebros saudáveis. Estratégias como *transfer learning*, uso de arquiteturas pré-treinadas (como VGG16 e ResNet) e mecanismos de atenção têm contribuído para aprimorar a acurácia e a robustez dos modelos, mesmo em cenários com conjuntos de dados limitados [3-6].

Figura 1. Diagrama mostrando a arquitetura da rede convolucional MobileNetV2.



Fonte: Elaborado pelos autores (2026), a partir de [10].

Apesar dos avanços, ainda persistem desafios importantes, como o desequilíbrio de classes, o risco de viés de generalização e a escassez de validação clínica. Ahmmed *et al.* [7] testaram modelos ResNet50 e InceptionV3 por meio de *transfer learning*, obtendo cerca de 97,7% de acurácia, especialmente em gliomas e tumores pituitários. Entre as limitações, destacam-se a dependência do domínio experimental e da base utilizada. Wong *et al.* [8] aplicaram a arquitetura VGG16 com *fine-tuning* a um conjunto maior, contendo 17.136 imagens distribuídas em quatro classes, obtendo 99,24% de acurácia. Mesmo assim, os autores enfatizam a necessidade de validação clínica e o risco de *sobreajuste* em bases públicas. Alemayehu [9] propôs uma CNN leve, otimizada com Keras Tuner, alcançando 98,78% de acurácia. O foco do estudo está na redução da complexidade computacional, embora permaneça a necessidade de testes com bases mais variadas e provenientes de diferentes equipamentos de aquisição de imagem.

1.4 Rede Neural Convolucional MobileNetV2

A MobileNetV2 é uma arquitetura de rede neural convolucional projetada para eficiência computacional e baixo consumo de recursos em dispositivos móveis. Ela introduz o bloco *inverted residual* com conexões de atalho (*skip connections*) e expansão linear: primeiro expande a dimensionalidade dos canais por meio de uma camada 1×1, aplica convoluções *depthwise* 3×3, e então projeta de volta a dimensionalidade original com outra camada 1×1 sem ativação não-linear. Essa estrutura reduz a quantidade de parâmetros e operações FLOPs em comparação com versões anteriores, mantendo alta acurácia em tarefas de classificação de imagens (≈ 71 % top-1 no ImageNet). A rede utiliza *Batch Normalization* e ReLU6 como funções de ativação, favorecendo a robustez em faixas de precisão reduzida (por exemplo, quantização para 8 bits) [10]. A figura 1 apresenta a arquitetura da rede, com uma das possíveis combinações de dimensões da imagem de entrada e tamanho da camada densa (*fully connected*).

2 MÉTODOS

Utilizou-se o Brain Tumor MRI Dataset, composto por imagens provenientes de Figshare, SARTAJ e Br35H [11], organizadas em quatro classes (glioma, meningioma, tumor pituitário, sem tumor). O repositório disponibiliza pastas já estratificadas em treino e teste, aleatoriamente, a fim de retirar quaisquer possíveis vieses impostos pelo autor do *dataset* na separação das imagens em pastas. No presente estudo, configurou-se 20% do conjunto de dados (1.404 imagens) para teste, restando 5.619 amostras para treino e 1.404 para testes, todas no formato JPG.

Importante destacar que, por ser uma base de imagens de pacientes anônima, sem metadados e pública, não houve a necessidade de consulta a Comitê de Ética, assim como as imagens submetidas por usuários ao testar a API são voláteis, não ficando armazenadas após a identificação da classe de tumor.

A distribuição por classe no conjunto de testes foi: Glioma (324; 23,08%), Meningioma (329; 23,43%), Pituitário (351; 25,00%), sem tumor (400; 28,49%).

2.1 Pré-Processamento das imagens

O pré-processamento seguiu etapas padronizadas para modelos baseados em MobileNetV2, garantindo compatibilidade com o *backbone* da rede pré-treinado e estabilidade no treinamento. As imagens foram inicialmente convertidas para o formato RGB, redimensionadas para 224×224 *pixels* (dimensão original das imagens obtidas) e, posteriormente,

normalizadas para o intervalo [0,1] por meio da operação *rescale* = 1./255.

Além disso, as amostras foram organizadas em lotes (*batches*) de 8 imagens, conforme a configuração de treinamento definida. Como dito anteriormente, o particionamento dos dados utilizou 20% das amostras para testes, porém, mantendo embaralhamento durante a geração dos lotes de treino para reduzir viés na ordem das amostras.

As imagens para treinamento e teste apresentam o padrão mostrado na figura 2. Procurou-se obter amostras em aquisições nos eixos axial, sagital e coronal, sendo o primeiro mais frequente (80%).

2.2 Backbone do Modelo

A MobileNetV2 foi carregada com pesos pré-treinados sobre o conjunto ImageNet para acelerar a convergência e melhorar a capacidade de generalização. A camada base permaneceu inicialmente congelada durante a primeira fase do treinamento e teve suas últimas camadas liberadas na etapa de *fine-tuning*.

Sobre o *backbone* foi acoplado um cabeçalho de classificação composto por:

- GlobalAveragePooling2D, responsável por condensar espacialmente os mapas de ativação;
- Camada densa com 256 neurônios, ativação ReLU, atuando como camada totalmente conectada intermediária;
- Camada densa final com 4 neurônios, ativação Softmax, representando as quatro classes: glioma, meningioma, sem tumor e pituitário.

O modelo completo possui 2.586.948 parâmetros, dos quais 2.190.404 foram treináveis após o início do *fine-tuning*, enquanto 396.544 permaneceram congelados (fig. 1). O número de parâmetros é a soma de todos os pesos das conexões da rede configurada neste trabalho.

2.3 Ambiente de desenvolvimento

A linguagem utilizada foi a Python 3.10.0. Para programar e treinar a rede neural utilizou-se a biblioteca TensorFlow 2.20.0 com API Keras [12].

O *frontend* (interface gráfica) foi desenvolvido com NextJS (um *framework* JS) utilizando ReactJS (uma biblioteca JavaScript), que ajudam a criar componentes e a integração com o *deploy* na Vercel. Neste é utilizada a biblioteca axios (arquivo: https://github.com/leonfagundes/IC-frontend/blob/main/lib/upload.ts) que facilita a comunicação com a API que está no servidor, em *deploy* na plataforma *Render*.

A chamada da API no *backend* é feita apenas no momento em que o usuário envia a imagem clicando no botão enviar, após ter carregado a foto previamente. Após fazer a requisição HTTP na API, ela processa a imagem e responde na página *HTTP*.

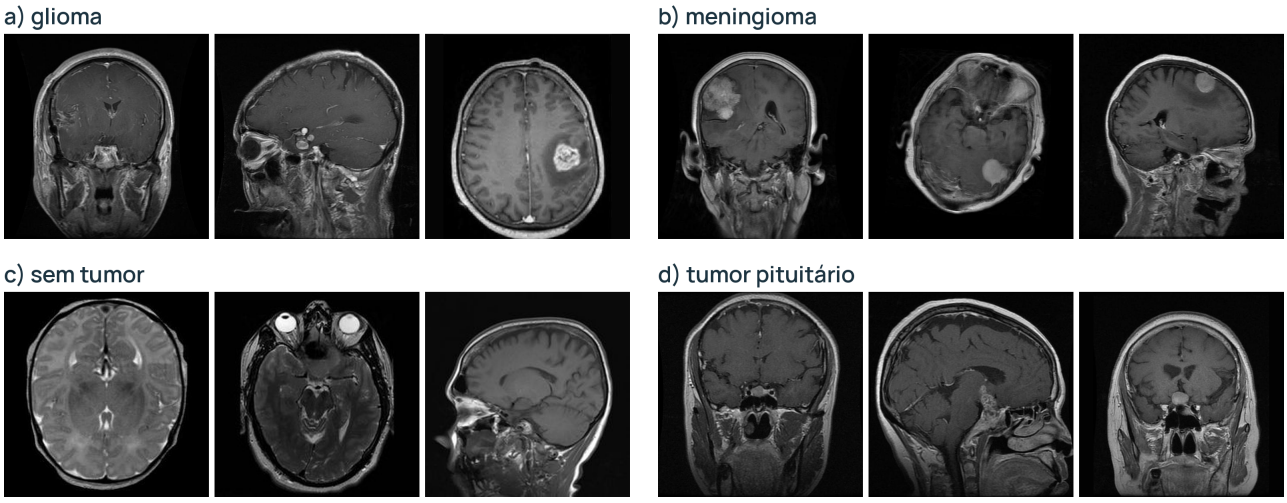
Existe uma verificação da API antes de fazer o *resizing* da imagem para ser corretamente processada pela rede neural, e o pré-processamento para identificar se a imagem está em tons de cinza (caso da ressonância magnética), ou se as imagens de RM contêm alguma anotação do usuário, como seta ou círculo colorido na imagem (sendo recusada).

A API escrita em linguagem Python (https://github.com/leonfagundes/IC-backend/blob/main/api/main.py) carrega o modelo pré-treinado da rede neural [13] para então analisar a imagem enviada pelo usuário.

2.4 Processo de treinamento e testes

O treinamento foi realizado em duas fases distintas, seguindo o procedimento tradicional para modelos baseados em *transfer learning* com *fine-tuning*.

Figura 2. Exemplos de imagens utilizadas no sistema.



Fonte: Brain Tumor MRI Dataset (Kaggle) [11].

Fase 1: Treinamento inicial (*backbone* congelado)

Na primeira etapa, apenas o cabeçalho de classificação foi treinado, mantendo todo o *backbone* MobileNetV2 congelado. Foram configuradas até 30 épocas, mas o processo foi interrompido após 21 épocas pelo mecanismo de *Early Stopping*, monitorando a perda de validação. Nesta etapa, a rede foi configurada com os seguintes parâmetros: Otimizador Adam, *learning rate* = 5×10^{-4}, EarlyStopping: monitor='val_loss', patience=10, ReduceLROnPlateau: factor=0,5, patience=3.

Esses parâmetros resultaram no seguinte desempenho: Acurácia de testes: 0,8654 e final: 0,9615, com a menor perda nos testes de 0,1232 na época 11 e melhor acurácia de 0,9651 na época 17. Esses resultados indicam rápida convergência e efetivo aprendizado inicial, devido aos pesos já estarem configurados com a base ImageNet.

Fase 2: *Fine-tuning* (liberação parcial do *backbone*)

Na segunda fase, o modelo passou a incluir as camadas superiores da MobileNetV2, mantendo congeladas as primeiras 100 camadas. O *fine-tuning* foi configurado para 20 épocas, sendo executadas 10 até a estabilização do desempenho. Nesta etapa, a rede foi configurada com os seguintes parâmetros: Melhor acurácia: 0,9544 (época 8); Otimizador Adam, *learning rate* = 1×10^{-5}; Camadas desbloqueadas: da camada 100 em diante; *Data augmentation*: *rescale* = 1./255; Acurácia de testes: manteve-se entre 0,9259 e 0,9544; Menor perda de testes: 0,1569.

A estabilidade da acurácia e a redução progressiva da perda indicam que o *fine-tuning* não levou a sobreajuste, mantendo consistência entre treino e testes.

O sistema pode ser acessado via Web no endereço https://ic-frontend-six.vercel.app, onde o usuário pode carregar uma imagem no padrão requerido para a análise e classificação (figura 3).

Figura 3. Interface do usuário via navegador Web.



Fonte: Elaborado pelos autores (2026).

3 RESULTADOS

A avaliação do modelo foi realizada no conjunto de testes composto por 1.404 imagens distribuídas em quatro classes (glioma, meningioma, sem tumor e pituitário). O desempenho global atingiu acurácia de 92,59%, com métricas *macro* e *weighted* elevadas, indicando bom equilíbrio entre as classes.

As métricas globais de desempenho obtidas foram as seguintes: Acurácia: 0,9259; *Loss*: 0,3563; Precisão macro: 0,9316; *Recall* macro: 0,9224; F1-score macro: 0,9238; Precisão *weighted*: 0,9336; *Recall weighted*: 0,9259; F1-score *weighted*: 0,9268.

Tabela 1. Métricas de desempenho nos testes do modelo

Classe	Precisão	*Recall*	F1-Score	AUC-ROC	Acerto
Glioma	0,98	0,85	0,91	0,99	0,85
Meningioma	0,80	0,95	0,87	0,98	0,95
Pituitário	0,99	0,91	0,95	1,00	0,91
Sem Tumor	0,96	0,99	0,97	1,00	0,99

Fonte: Elaborado pelos autores (2026).

Tabela 2. Matriz de confusão na amostra de testes

Previsto \ Real	Glioma	Meningioma	Sem Tumor	Pituitário
Glioma	274	50	0	0
Meningioma	1	311	13	4
Sem Tumor	0	3	397	0
Pituitário	5	23	5	318

Fonte: Elaborado pelos autores (2026).

Conforme a tabela 1, os maiores índices de F1 estão em sem tumor (0,97) e pituitário (0,95), refletindo alta sensibilidade e especificidade nessas categorias. Na tabela 2, a matriz de confusão revela que os maiores erros de classificação ocorreram entre glioma e meningioma. A acurácia da previsão pode ser avaliada somando-se os acertos (diagonal principal da matriz) em relação ao total.

3.1 Resultados Qualitativos

A análise das amostras corretamente classificadas mostra que o modelo capturou bem padrões texturais e morfológicos consistentes das classes sem tumor e pituitário, possivelmente devido à sua menor variabilidade intraclasse.

Para as confusões mais frequentes entre glioma e meningioma, observou-se: a) sobreposição de padrões morfológicos; b) heterogeneidade de localização (intra-axial vs. extra-axial); c) presença variável de realce pós-contraste e edema; d) diferenças no plano de corte que podem reduzir a discriminação espacial.

Tais fatores são obstáculos clínicos conhecidos e refletem a dificuldade enfrentada por radiologistas em exames isolados.

Os resultados apresentam não apenas elevada acurácia global, mas também consistência entre diferentes métricas de avaliação. O alinhamento entre precisão, *recall* e F1-*score,* tanto nas métricas *macro* quanto nas *weighted*, evidencia que o modelo se comporta de maneira equilibrada entre classes de tamanhos distintos. Além disso, o desempenho superior em sem tumor e pituitário, ambos com *recall* acima de 0,90, sugere que padrões estruturais bem definidos são revelados pelo modelo, demonstrando sua capacidade de generalizar adequadamente até em casos de variabilidade anatômica.

4 DISCUSSÃO

Os resultados evidenciam que uma arquitetura leve baseada na MobileNetV2, com entrada 224×224 e cabeçalho denso reduzido, oferece desempenho competitivo para classificação de tumores cerebrais.

A alta precisão observada na classe glioma, embora acompanhada de *recall* reduzido, indica que o modelo privilegia a redução de falsos positivos nessa categoria, o que pode ser clinicamente desejável em cenários de triagem, onde o custo de um diagnóstico excessivamente sensível pode resultar em sobrecarga do fluxo de trabalho. No entanto, essa assimetria revela também a necessidade de estratégias que aumentem a sensibilidade sem prejudicar a especificidade, como ajustes de limiar, reequilíbrio das classes ou ampliação do *dataset* com casos de glioma mais diversos.

Esses achados reforçam que o bom desempenho observado não depende de arquiteturas profundas ou modelos de grande porte. Ao contrário, demonstram que uma CNN leve como a MobileNetV2 é capaz de lidar com características morfológicas complexas presentes em tumores cerebrais, mesmo em condições de treinamento limitado. Isso é coerente com o objetivo central desta pesquisa: validar a eficácia de arquiteturas compactas para tarefas clínicas relevantes.

4.1 Desempenho geral

Os resultados globais permitem reforçar a hipótese central deste estudo: arquiteturas leves podem atingir desempenho competitivo em tarefas clínicas complexas. O modelo preditivo atingiu AUC ≥ 0,97 em todas as classes, F1 ≥ 0,94 para os casos “sem tumor” e “pituitário”, com acurácia global acima de 90%, sem sinais de *sobreajuste*.

Esses valores demonstram grande capacidade discriminativa, especialmente para classes com padrões bem definidos.

4.2 Desempenho inferior em meningioma

A classe meningioma apresentou os menores valores de precisão e F1, o que pode ser atribuído a:

- Heterogeneidade morfológica (graus de agressividade, calcificações, variantes).
- Variações de protocolo entre os exames, especialmente em T1 com contraste (CE-T1).
- Possível leve desbalanceamento na composição do *dataset*.

4.3 Possíveis melhorias

Para aumentar o desempenho nessa classe, há algumas iniciativas a considerar, dentre elas a técnica de *Data Augmentation* ser mais direcionada às características morfológicas das imagens, definindo ou limitando parâmetros da técnica tais como: rotação, zoom, deformações, transformações elásticas; uso de *Focal loss* ou *class weights* para favorecer classes difíceis; técnicas de mixagem (*mixup*, *cutmix*); entradas maiores (ex.: 256×256 ou 288×288); inclusão de mapas de atenção (Grad-CAM) para regularizar as regiões relevantes.

Ainda, a divisão dos dados foi feita por imagem e não por paciente, o que pode alterar sensivelmente as métricas, já que múltiplos cortes de um mesmo exame podem aparecer simultaneamente em treino e teste (vazamento de dados). No repositório há um comentário de que não há informação sobre pacientes [11]. Portanto, a identificação de quais cortes pertencem a quais pacientes é um próximo passo importante no desenvolvimento desta API.

4.4 Implicações práticas

Os altos valores de F1 e AUC nos casos “sem tumor” e “pituitário” indicam potencial no auxílio ao diagnóstico para uma triagem assistida, proporcionando uma

redução de carga de trabalho e, principalmente, a priorização de casos suspeitos.

A eficiência computacional do modelo também tem implicações práticas importantes. O fato de que a MobileNetV2 foi treinada sem GPU, resultando em boas métricas, indica que soluções baseadas em CNNs leves podem ser implantadas em hospitais, clínicas e sistemas de triagem rápida com recursos computacionais modestos, ampliando o acesso a tecnologias de apoio ao diagnóstico.

Porém, para classes críticas (glioma/meningioma), recomenda-se cuidado ao realizar validação externa em bases multi-institucionais, com apoio das redes de saúde ao criar um repositório de imagens para estes casos. O uso dos resultados de estudos clínicos prospectivos também reforçaria o desempenho da rede neural.

4.5 Aspectos de validação

Embora o modelo apresente desempenho robusto, é importante considerar limitações que influenciam a interpretação da eficácia de arquiteturas leves. A partição por imagens (não por paciente) pode inflar métricas, mas não compromete o objetivo principal desta pesquisa, que é avaliar a capacidade de generalização de uma arquitetura compacta. A validação externa ainda é essencial.

A heterogeneidade de *scanners* e protocolos, conforme evidenciada pelas variações nas classes glioma e meningioma, reforça a necessidade de validação externa em ambientes multi-institucionais, assegurando a aplicabilidade clínica do modelo em cenários variados.

Outros vieses podem influenciar o desempenho do sistema, como realizar treinamento por quantidade de imagens (e não por paciente), podendo alterar métricas quando há múltiplos cortes do mesmo exame.

4.6 Governança dos dados

Sendo uma ferramenta com API pública que recebe imagens de usuários, cabe destacar que as imagens são anônimas, usadas apenas para executar as requisições e descartadas imediatamente após o processamento, sem novos treinamentos da rede neural. O formato da imagem é validado no envio, minimizando o risco de conteúdo malicioso e não são criadas contas para usuários neste protótipo.

5 CONCLUSÕES

Este estudo mostrou que arquiteturas leves de redes neurais convolucionais, como a MobileNetV2, podem alcançar desempenho competitivo na detecção multiclasse de tumores cerebrais. Com métricas elevadas e estabilidade de treinamento mesmo sem GPU, os resultados confirmam a hipótese central desta pesquisa: modelos compactos são capazes de capturar padrões morfológicos complexos e oferecer suporte confiável ao diagnóstico.

A elevada performance evidencia que, mesmo com um cabeçalho raso e entrada em baixa resolução (imagens 224×224 pixels), a rede preserva a capacidade de reconhecer padrões morfológicos complexos.

Entre as classes avaliadas, destaca-se o desempenho para glioma, que alcançou precisão de 0,9786 e AUC-ROC de 0,9898, indicando excelente especificidade. Apesar do *recall* moderado (0,8457), a combinação de métricas sugere que o modelo é particularmente eficiente em identificar corretamente amostras verdadeiramente representativas de glioma, minimizando previsões falsas positivas. Esse comportamento é clinicamente relevante, uma vez que gliomas estão entre as neoplasias intracranianas mais desafiadoras e que demandam alta confiabilidade na triagem automática.

O desempenho robusto também para as classes pituitário e sem tumor, ambas com F1 superiores a 0,94, evidencia a consistência do modelo para estruturas mais homogêneas ou padrões não patológicos, demonstrando estabilidade no reconhecimento de características anatômicas regulares. A estabilidade das curvas de testes durante o *fine-tuning* reforça que o modelo converge de forma controlada, com baixo risco de sobreajuste.

Do ponto de vista prático, os resultados apontam um caminho promissor para o uso de CNNs leves em cenários de triagem clínica, especialmente em serviços com limitação de *hardware*, já que o treinamento foi realizado sem GPU dedicada. O estudo confirma que modelos compactos, quando combinados a estratégias de treinamento adequadas, podem atingir desempenho comparável ao de arquiteturas mais pesadas. Assim, a presente abordagem mostra potencial para integração em fluxos de trabalho reais, desde que acompanhada de validação externa e avaliação multimodal em dados independentes.

Assim, esta pesquisa reforça que a eficiência computacional não precisa ser obtida às custas de desempenho, e que modelos leves representam uma solução promissora para sistemas de apoio ao diagnóstico em larga escala.

Muito importante enfatizar que o modelo não substitui o radiologista; funciona como suporte à decisão, podendo auxiliar na triagem de exames e na detecção de padrões relevantes.

A principal contribuição deste trabalho foi a criação da API/interface para estudos e treinamento de usuários da área médica, mais do que a metodologia de reconhecimento com redes neurais leves.

LINKS

A API desenvolvida poderá ser utilizada no seguinte site: https://ic-frontend-six.vercel.app

Contribuição dos autores. Os autores contribuíram igualmente em todas as etapas de construção do artigo: concepção e delineamento do estudo; coleta, análise e interpretação dos dados; redação e revisão crítica do manuscrito; e aprovação da versão final.

Agradecimentos. Ao Centro Estadual de Educação Tecnológica Paula Souza (CEETEPS) pela concessão de bolsa de Iniciação Científica ao autor deste trabalho.

Financiamento. Os autores não declararam fonte de financiamento.

Disponibilidade de dados e materiais. Os autores não declararam a disponibilidade dos dados e materiais da pesquisa.

Conflitos de interesse. Os autores não declararam conflitos de interesse.

Declaração de uso de IA. Os autores declaram que o uso de Inteligência Artificial Generativa contribuiu apenas para melhorar a clareza e a gramática do texto, sendo o conteúdo, principalmente a discussão e as conclusões, de inteira responsabilidade dos autores.

REFERÊNCIAS

1. Sampath SG, Manriquez AH, Haile H, Botkiss M, Izima C, Adapa AR, et al. Socioeconomic disparities in patient-reported outcomes, health literacy, and access to care among patients with primary brain tumors: findings from the All of Us Research Program. J Neurooncol. 2026;176(1):123. doi: https://doi.org/10.1007/s11060-025-05381-7
2. Mohebbi A, Mohammadzadeh S, Zare AH, Moradi Z, Ardakani AA, Mohammadi A. Assessing inter-rater reliability of MRI features in glioma: a multi-radiologist agreement study. BMC Med Imaging. 2025;25(1):480. doi: https://doi.org/10.1186/s12880-025-01941-5
3. Yadav RK, Kumar M, Nandi A. A scalable brain tumor diagnosis from large-scale MRI datasets using CNN-ViT and expert-attention fusions. Int J Comput Intell Syst. 2025;18(1):254. doi: https://doi.org/10.1007/s44196-025-00981-7
4. Naeem AB, Osman O, Alsubai S, Cevik T, Zaidi A, Rasheed J. Lightweight CNN for accurate brain tumor detection from MRI with limited training data. Front Med (Lausanne). 2025;12:1636059. doi: https://doi.org/10.3389/fmed.2025.1636059
5. Rasheed M, Iqbal S, Jaffar A, Akram S. Advanced deep learning-based brain tumor classification using a novel customized CNN and optimized residual network. PLoS One. 2025;20(10):e0334430. doi: https://doi.org/10.1371/journal.pone.0334430
6. Erukude ST, Marella VC, Veluru SR. Explainable deep learning in medical imaging: brain tumor and pneumonia detection. In: 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA); 2025 set 3-5; Tirupur, India. [s.l.]: IEEE; 2025. p. 906-11. doi: https://doi.org/10.1109/ICIMIA67127.2025.11200629
7. Ahmmed S, Podder P, Mondal M, Rahman SMA, Kannan S, Hasan MJ, et al. Enhancing brain tumor classification with transfer learning across multiple classes: an in-depth analysis. BioMedInformatics. 2023;3(4):1124-44. doi: https://doi.org/10.3390/biomedinformatics3040068
8. Wong Y, Su ELM, Yeong CF, Holderbaum W, Yang C. Brain tumor classification using MRI images and deep learning techniques. PLoS One. 2025;20(5):e0322624. doi: https://doi.org/10.1371/journal.pone.0322624
9. Alemayehu N. Light weight CNN for classification of brain tumors from MRI images. arXiv [Preprint]. 2025. doi: https://doi.org/10.48550/arXiv.2504.21188
10. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018 jun 18-23; Salt Lake City (UT). [s.l.]: IEEE; 2018. p. 4510-20. doi: https://doi.org/10.1109/CVPR.2018.00474
11. Nickparvar M. Brain Tumor MRI Dataset. [s.l.]: Kaggle; 2021. [citado 2025 nov 6]. doi: https://doi.org/10.34740/KAGGLE/DSV/2645886
12. Keras. Keras: deep learning for humans [Internet]. [citado 2025 nov 6]. Disponível em: https://keras.io
13. GitHub. Modelo treinado: modelo_mobile-netv2_224x224_finetuned.keras [Internet]. [citado 2025 nov 6]. Disponível em: https://github.com/leonfagundes/IC-backend/blob/main/model/modelo_mobilenetv2_224x224_finetuned.keras

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde