

DETECÇÃO E CLASSIFICAÇÃO DE TUMORES CEREBRAIS EM IMAGENS DE RESSONÂNCIA MAGNÉTICA UTILIZANDO REDES NEURAIS CONVOLUCIONAIS

Leon Pereira Pinto Fagundes¹
Waldemar Bonventi Júnior²

RESUMO

A identificação precoce de tumores cerebrais é essencial para o adequado planejamento terapêutico e definição prognóstica. Neste estudo, propomos um modelo automático de classificação de imagens de ressonância magnética cerebral em quatro categorias — glioma, meningioma, pituitária e ausência de tumor — utilizando a arquitetura MobileNetV2 com transferência de aprendizado e fine-tuning. O modelo foi treinado sobre o Brain Tumor MRI Dataset (Kaggle) a partir do método supervisionado, dividido em conjuntos de treino com 5.619 amostras e validação com 1.404 amostras. A solução alcançou acurácia global de 92,5% e F1-score macro de 92,4%, com AUC-ROC $\geq 0,98$ para todas as classes. A classe “sem tumor” obteve o melhor desempenho com 97,9% de precisão, enquanto meningioma apresentou maior variabilidade, apresentando F1-score de 86,9%. A análise da distribuição de confiança evidenciou valores predominantemente acima de 0,97, indicando estabilidade decisória. Observou-se, contudo, confusão cruzada entre glioma e meningioma, possivelmente relacionada à sobreposição morfológica ou à heterogeneidade de aquisição. Os resultados demonstram a viabilidade de CNNs leves para apoio ao diagnóstico em cenários clínicos, ressaltando a necessidade de validação externa e ferramentas de interpretabilidade para uso seguro.

PALAVRAS-CHAVE: CNN; tumores cerebrais; deep learning; MobileNetV2.

INTRODUÇÃO

Os tumores cerebrais constituem relevante problema de saúde pública e no sistema nervoso central do portador, com impacto funcional e socioeconômico expressivo na vida do paciente. A classificação correta do tumor orienta manejo e prognóstico. A ressonância magnética (RM) é o padrão de excelência para avaliação não invasiva de lesões intracranianas, porém a interpretação depende da experiência do radiologista e pode sofrer variações interobservador.

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 em detecção e classificação de tumores cerebrais por RM. Persistem, contudo, desafios: bases restritas, cenários binários (tumor versus não tumor), heterogeneidade de aquisição e a dependência de hiperparâmetros/arquiteturas.

Este trabalho propõe, portanto, uma classificação multiclasse (4 classes) em RM cerebral empregando MobileNetV2 com transfer learning e fine-tuning, avaliando

¹ Desenvolvimento de Software Multiplataforma, Fatec Votorantim Benedicto Pagliato, Votorantim, Brasil, leon.fagundes@fatec.sp.gov.br, contato.leonfagundes@gmail.com.

² Doutor, Linha de pesquisa: Aprendizagem de Máquina, Fatec Votorantim Benedicto Pagliato, Votorantim, Brasil. E-mail: waldemar.junior@fatec.sp.gov.br, ORCID: 0000-0003-0720-499X.

rigorosamente métricas globais e por classe, além de analisar overfitting, distribuição de confiança e curvas ROC. Contribuímos com um pipeline reproduzível, leve e compatível com implantação em cenários com recursos computacionais moderados.

2 CLASSIFICAÇÃO MULTICLASSE DE TUMORES CEREBRAIS EM RM UTILIZANDO MOBILENETV2 COM FINE-TUNING

2.1 Tumores Cerebrais

Gliomas aparecem no meio do cérebro (região intra-axial), com bordas irregulares e mal definidas. Em T1 (imagem de mais realce anatômico) são mais escuros e em T2 (imagem que realça mais endemas) mais claros, cercados por uma área ainda mais branca (edema, que indica inchaço). Quando recebem contraste, podem realçar de forma desigual (partes internas ficam mais brancas), especialmente nos tipos mais agressivos. O aspecto geral é desorganizado, com sinais espalhados e mistura com o tecido cerebral.

Meningiomas ficam na parte externa do cérebro (região extra-axial), com contornos arredondados bem definidos e empurram o cérebro ao invés de invadi-lo. Após o contraste, ficam intensamente brancos de forma uniforme. Um sinal comum é uma “cauda” branca saindo da base do tumor (realce da membrana que reveste o crânio). A imagem é limpa, com margens visíveis separando o tumor do cérebro.

Adenomas da hipófise (tumores pituitários) aparecem centralizados na base do cérebro, dentro da sela túrcica (estrutura óssea no meio do crânio). São redondos ou ovais, escuros em T1 e discretamente mais claros em T2. Após o contraste, algumas partes realçam (ficam brancas) e outras não, indicando áreas com líquido ou tecido degenerado. Quando crescem, podem parecer um “8” achatado ou um boneco de neve.

Nas imagens sem tumor, não há áreas que se destacam. O cérebro aparece simétrico, com coloração uniforme e nenhuma massa empurrando ou invadindo estruturas. Após contraste, só regiões normais ficam levemente mais claras, sem alterações focais.

Resumindo: gliomas são internos, borrados e irregulares; meningiomas são externos, bem definidos e muito brancos com contraste; tumores pituitários são centrais e têm realce parcial; imagens normais não têm realce ou deformações. Esses padrões visuais ajudam especialistas — e algoritmos — a classificar corretamente as imagens.

2.2 A Inteligência Artificial Na Saúde: Potencial E Desafios

A aplicação de técnicas de IA na área da saúde tem ampliado significativamente as possibilidades de suporte ao diagnóstico médico, especialmente em contextos com alta demanda de exames e escassez de especialistas. Modelos de deep learning, especialmente CNNs, são capazes de detectar padrões em imagens médicas com desempenho comparável ao de profissionais experientes, podendo atuar como ferramentas auxiliares na triagem e priorização de casos.

Contudo, mesmo modelos com acurácia global superior a 90% estão sujeitos a falsos positivos e falsos negativos, como evidenciado por suas respectivas matrizes de confusão ainda nesse artigo. Em problemas multiclasse como a classificação de tumores cerebrais, erros de classificação podem ocorrer com maior frequência entre classes com características similares, como gliomas e meningiomas de alto volume.

Além de outras métricas estatísticas, o Teorema de Bayes ajuda a interpretar a confiabilidade do diagnóstico fornecido por modelos de IA ao considerar a probabilidade da doença e a taxa de erro do classificador. Mesmo com um classificador preciso, a probabilidade pós-teste (isto é, a confiança no resultado fornecido) pode ser baixa em populações com baixa prevalência da condição.

Por isso, é fundamental que os resultados fornecidos por modelos de IA sejam interpretados por profissionais de saúde, dentro de um contexto clínico amplo. O uso indevido por leigos, sem apoio médico, pode levar a autodiagnósticos equivocados, com riscos psicológicos e clínicos relevantes.

Assim, destaca-se a IA como um instrumento de apoio ao especialista ou de cunho educacional, e não como substituto ao diagnóstico de um profissional, sendo necessária transparência nos indicadores de desempenho, como sensibilidade, especificidade, curva ROC e análise de classe dominante, especialmente em modelos com desequilíbrios na distribuição das classes ou performance desigual entre elas.

2.3 Redes Neurais Convolucionais e MobileNetV2

Inspiradas nas redes neurais convolucionais humanas, as CNNs utilizam camadas convolucionais para extrair padrões espaciais com compartilhamento de pesos, o que as torna eficientes e invariantes a translações locais (ou seja, reconhecem algo mesmo se estiver um pouco deslocado). Arquiteturas modernas (p.ex., MobileNetV2 utilizada nesse projeto) incorporam blocos otimizados que expandem e comprimem informações das imagens de forma eficiente, além de camadas finais simplificadas que preservam detalhes importantes. Essas estratégias permitem alto desempenho com baixo custo computacional, atributo crucial para aplicações clínicas e dispositivos na borda, como celulares, sensores ou totêmicos de atendimento hospitalar.

A MobileNetV2 surgiu como uma evolução da família MobileNet, desenvolvida para atender cenários com restrições de recursos computacionais (dispositivos móveis ou embarcados).

No seu artigo original, os autores apresentam uma nova estrutura de blocos de camadas que mantém bom poder de representação mesmo com baixo custo — o que reduz drasticamente a quantidade de operações e memória necessárias sem sacrificar a acurácia.

Essa combinação de leveza e desempenho torna a MobileNetV2 particularmente adequada para aplicações em saúde, como a construção de um programa que receba imagens de tomografia ou ressonância e classifique automaticamente a presença ou ausência de tumor cerebral — pois ela permite que o sistema seja executado em máquinas com recursos limitados ou mesmo em dispositivos de borda, mantendo rapidez, eficiência e boa precisão.

2.4 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.

Apesar dos avanços, ainda persistem desafios importantes, como o desequilíbrio de classes, o risco de viés de generalização e a escassez de validação clínica. A seguir são apresentados os principais trabalhos que fundamentam esta pesquisa.

Ahmmmed et al. (2023) 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. (2024) 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 overfitting em bases públicas.

Alemayehu (2025) 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.

3 METODOLOGIA

A estratégia de treinamento adotada foi cuidadosamente estruturada em duas fases para mitigar sobreajuste e estabilizar gradientes. A primeira fase, com o backbone congelado, permitiu o ajuste fino do cabeçalho de classificação sem interferir nas representações convolucionais de baixo nível. A segunda fase, com fine-tuning a partir da camada 100 e learning rate significativamente reduzido, assegurou refinamento progressivo das características discriminativas de alto nível. Essa abordagem híbrida, combinada com o Early Stopping e o ReduceLROnPlateau, conferiu ao processo de treinamento maior estabilidade e melhor generalização.

3.1 Base De Dados

Utilizou-se o Brain Tumor MRI Dataset (Kaggle), composto por imagens provenientes de Figshare, SARTAJ e Br35H, organizadas em quatro classes (glioma, meningioma, pituitary, sem tumor). O repositório disponibiliza pastas já estratificadas em treino e teste que foram

unificados durante o treinamento e teste desse trabalho, a fim de retirar qualquer possíveis vieses impostos pelo autor do dataset na separação das imagens em pastas. No presente estudo, configurou-se 20% de separação do conjunto de dados para teste, resultando em 5.619 amostras para treino e 1.404 para validação, todas no formato JPG.

A distribuição por classe na validação foi:

- Glioma (324; 23,08%)
- Meningioma (329; 23,43%)
- Pituitária (351; 25,00%).
- Sem tumor (400; 28,49%)

3.2 Pré-Processamento Das Imagens

O pré-processamento seguiu etapas padronizadas para modelos baseados em MobileNetV2, garantindo compatibilidade com o backbone pré-treinado e estabilidade no treinamento. As imagens foram inicialmente convertidas para o formato RGB, redimensionadas para 224×224 pixels — dimensão de entrada requerida pela arquitetura — e posteriormente normalizadas para o intervalo [0,1] por meio da operação $\text{rescale} = 1./255$.

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

3.3 Arquitetura Do Modelo

A arquitetura é baseada na MobileNetV2, utilizando pesos pré-treinados no 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, notumor e pituitary.

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.

3.4 Ambiente De Desenvolvimento

- Linguagem: Python 3.10.0
- Biblioteca: TensorFlow 2.20 com Keras
- Treinamento sem utilização de GPU
- Dependências principais: Numpy, pandas, scikit-learn, matplotlib, pyyaml, pathlib, python-multipart, pillow

3.5 Processo De Treinamento E Validação

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

3.5.1 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.

Configurações principais:

- Otimizador Adam, learning rate = 5×10^{-4}
- EarlyStopping: monitor='val_loss', patience=10
- ReduceLROnPlateau: factor=0.5, patience=3
- Acurácia inicial: 0.8653 → final: 1.0000
- Acurácia de validação: 0.8654 → final: 0.9615
- Menor perda na validação: 0.1232 (época 11)
- Melhor acurácia: 0.9651 (época 17)
- Menor loss de validação: 0.1232 (época 11)

Esses resultados indicam rápida convergência e efetivo aprendizado inicial.

3.5.2 Fase 2 – Fine-tuning (liberação parcial do backbone)

Na segunda fase, o modelo passou a treinar as camadas superiores da MobileNetV2, mantendo congeladas as primeiras 100 camadas. O fine-tuning foi configurado para 20 épocas, sendo executadas 10 até estabilização do desempenho.

Parâmetros principais:

- Melhor acurácia: 0.9544 (época 8)
- Otimizador Adam, learning rate = 1×10^{-5}
- Camadas liberadas: da camada 100 em diante
- Data augmentation: rescale = 1./255
- Acurácia de validação: manteve-se entre 0.9259–0.9544
- Menor perda de validação: 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 validação.

4 RESULTADOS E DISCUSSÃO

4.1 Resultados Quantitativos

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

4.1.1 Métricas globais de desempenho

- 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

Esses valores demonstram desempenho robusto do modelo, com coerência entre métricas macro e weighted, indicando ausência de viés significativo em relação ao tamanho das classes.

4.1.2 Métricas por classe na validação

| Classe | Precisão | Recall | F1-Score | AUC-ROC | Acerto (%) |
|------------|----------|--------|----------|---------|------------|
| Glioma | 0.9786 | 0.8457 | 0.9073 | 0.9898 | 84,57% |
| Meningioma | 0.8036 | 0.9453 | 0.8687 | 0.9797 | 94,53% |
| Pituitária | 0.9876 | 0.9060 | 0.9450 | 0.9970 | 90,60% |
| Sem Tumor | 0.9566 | 0.9925 | 0.9742 | 0.9992 | 99,25% |

Os maiores índices de F1 estão em sem tumor (0.9742) e pituitária (0.9450), refletindo alta sensibilidade e especificidade nessas categorias.

4.1.3 Matriz de Confusão

| | VALORES VERDADEIROS | | | | |
|---|---------------------|------------|---------|-----------|--|
| P | Glioma | Meningioma | NoTumor | Pituitary | |
| R | 274 | 50 | 0 | 0 | |
| E | 1 | 311 | 13 | 4 | |
| D | | | | | |

| | | | | | |
|---|-----------|---|----|-----|-----|
| I | NoTumor | 0 | 3 | 397 | 0 |
| T | Pituitary | 5 | 23 | 5 | 318 |

A matriz de confusão revela que as principais confusões ocorreram entre glioma e meningoíoma, sobretudo no sentido glioma → meningoíoma.

4.2 Resultados Qualitativos

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

Para as confusões mais frequentes entre glioma ↔ meningoíoma, observou-se:

- Sobreposição de padrões morfológicos;
- Heterogeneidade de localização (intra-axial vs. extra-axial);
- Presença variável de realce pós-contraste e edema;
- Diferenças no plano de corte que podem reduzir a discriminação espacial.

Tais fatores são conhecidos obstáculos clínicos e refletem a dificuldade mesma 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 notumor e pituitária, ambas com recall acima de 0,90, sugere que padrões estruturais bem definidos são prontamente capturados pelo modelo, demonstrando sua capacidade de generalizar adequadamente mesmo em situações de variabilidade anatômica.

4.3 Discussão Dos Resultados

Os resultados evidenciam que uma arquitetura leve baseada na MobileNetV2, com entrada 224×224 e cabeça densa reduzida, 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.

Outro aspecto relevante é a preservação da estabilidade durante o fine-tuning. A ausência de aumento significativo de val-loss e o leve crescimento da acurácia sugerem que

liberar parcialmente o backbone a partir da camada 100 foi uma escolha adequada, permitindo refinar representações de alto nível sem comprometer o aprendizado previamente consolidado.

4.3.1 Desempenho geral

O modelo atingiu:

- AUC $\geq 0,97$ em todas as classes
- F1 $\geq 0,94$ para notumor e pituitary
- Acurácia global $> 90\%$, sem sinais de overfitting

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

4.3.2 Desempenho inferior em meningioma

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

1. Heterogeneidade morfológica (graus de agressividade, calcificações, variantes).
2. Variações de protocolo entre os exames, especialmente em CE-T1.
3. Possível leve desbalanceamento na composição multi-institucional do dataset.

4.3.3 Possíveis melhorias

Para aumentar o desempenho nessa classe:

- Data augmentation mais direcionada (rotação, zoom, deformações, elastic transformations);
- 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.

4.3.4 Implicações práticas

Os altos valores de F1 e AUC em notumor e pituitary indicam potencial para:

- Triagem assistida,
- Redução de carga de trabalho,
- Priorização de casos suspeitos.

Porém, para classes críticas (glioma/meningioma), recomenda-se cuidado:

- Validação externa em bases multi-institucionais;
- Avaliação multimodal por sequência (T1, T2, FLAIR, CE-T1);

- Análise sistemática de interpretabilidade;
- Estudos prospectivos clínicos.

4.3.5 Ameaças à validade

Embora o modelo apresente desempenho robusto, algumas limitações devem ser reconhecidas. A separação do conjunto de validação por imagens, e não por paciente, pode introduzir dependências que inflacionem as métricas caso múltiplos cortes do mesmo exame estejam presentes em partições distintas.

Adicionalmente, 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.

- Partição por imagens (e não por paciente) pode inflar métricas quando há múltiplos cortes do mesmo exame;
- Heterogeneidade de scanners e protocolos;
- Metadados limitados impedem ajustes finos por sequência.

4.3.6 Considerações éticas

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

5 CONCLUSÃO

Os resultados obtidos neste estudo demonstram que uma arquitetura leve, baseada na MobileNetV2 com fine-tuning seletivo, é capaz de alcançar desempenho altamente competitivo na classificação de tumores cerebrais em imagens de RM. Com acurácia global de 92,59%, F1 macro de 0,9238 e AUC-ROC acima de 0,97 em todas as classes, o modelo apresenta maturidade técnica suficiente para aplicações de triagem e apoio ao diagnóstico.

A elevada performance evidencia que, mesmo com um cabeçalho raso e entrada em baixa resolução (224×224), 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 predições falsas positivas. Esse comportamento é relevante clinicamente, 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 pituitary e notumor, 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 validação e o baixo gap treino–validação durante o fine-tuning reforçam 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.

REFERÊNCIAS

Artigos científicos / Periódicos

CHENG, J. et al. Enhanced performance of brain tumor classification via tumor region augmentation and partition. PLoS ONE, v.10, n.10, 2015, e0140381. DOI: 10.1371/journal.pone.0140381.

DEEPAK, S.; AMEER, P. M. Brain tumor classification using deep CNN features via transfer learning. Computers in Biology and Medicine, v.111, 2019, p.103345. DOI: 10.1016/j.combiomed.2019.103345.

EL-DAHSHAN, E.-S. A. et al. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. Expert Systems with Applications, v.41, n.11, 2014, p.5526–5545. DOI: 10.1016/j.eswa.2014.01.021.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. Nature, Londres, v.521, p.436–444, 2015. DOI: 10.1038/nature14539.

REHMAN, A. et al. A deep learning-based framework for automatic brain tumors classification using transfer learning. Circuits, Systems, and Signal Processing, v.39, 2020, p.757–775. DOI: 10.1007/s00034-019-01246-3.

SULTAN, H. H. et al. Multi-classification of brain tumor images using deep neural network. IEEE Access, v.7, 2019, p.69215–69225. DOI: 10.1109/ACCESS.2019.2919122.

Livros

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. Cambridge: MIT Press, 2016.

LOUIS, D. N. et al. WHO Classification of Tumors of the Central Nervous System. 5. ed. Lyon: IARC/WHO, 2021.

Dataset / Repositório

NICKPARVAR, M. Brain Tumor MRI Dataset. Kaggle, 2020. Disponível em: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>. Acesso em: 6 nov. 2025.

Trabalhos correlatos

SANDLER, M. et al. MobileNetV2: inverted residuals and linear bottlenecks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. p.4510–4520. DOI: 10.1109/CVPR.2018.00474.

AHMMED, S. et al. Enhancing brain tumor classification with transfer learning across multiple classes: an in-depth analysis. BioMedInformatics, v.3, n.4, 2023, p.1124–1144. DOI: 10.3390/biomedinformatics3040068.

ALEMAYEHU, N. Light weight CNN for classification of brain tumors from MRI images. Pré-print (arXiv:2504.21188), 2025. DOI: 10.48550/arXiv.2504.21188.

WONG, Y.; SU, E. L. M.; YEONG, C. F.; HOLDERBAUM, W.; YANG, C. Brain tumor classification using MRI images and deep learning techniques. PLoS ONE, v.20, n.5, 2025, e0322624. DOI: 10.1371/journal.pone.0322624.

Documentação / Guias Técnicos

KERAS. Keras API Reference. Disponível em: <https://keras.io>. Acesso em: 6 nov. 2025.

TENSORFLOW. TensorFlow Guide. Disponível em: <https://www.tensorflow.org>. Acesso em: 6 nov. 2025.

Fontes médico-radiológicas

RADIOPAEDIA. Pituitary macroadenoma – Radiology Case. Disponível em: <https://radiopaedia.org/cases/pituitary-macroadenoma>. Acesso em: 6 nov. 2025.

EMEDICINE.MEDSCAPE. Pituitary adenoma imaging: practice essentials, computed tomography, magnetic resonance imaging. Disponível em: <https://emedicine.medscape.com/article/337678-overview>. Acesso em: 6 nov. 2025.

ENDOCRINE ABSTRACTS. Diagnostic difficulty in differentiating between meningioma and pituitary adenoma. In: BES2004 – 23rd Joint Meeting of the British Endocrine Societies. Disponível em: <https://www.endocrine-abstracts.org/ea/0007/ea0007p268>. Acesso em: 6 nov. 2025.

MEDALITY. Meningioma (Brain) – MRI Online/Medality. Disponível em: <https://www.medality.com/case/meningioma-brain>. Acesso em: 6 nov. 2025.

RADIOLOGY ASSISTANT. Systematic approach to brain tumors. Disponível em: <https://radiologyassistant.nl/neuroradiology/brain-tumors/systematic-approach-to-brain-tumors>. Acesso em: 6 nov. 2025.

EMEDICINE.MEDSCAPE. Astrocytoma: practice essentials, background, pathophysiology. Disponível em: <https://emedicine.medscape.com/article/250284-overview>. Acesso em: 6 nov. 2025.

MEDIUM. A summary of the MobileNetV2: inverted residuals and linear bottlenecks paper. Disponível em: <https://medium.com/codex/a-summary-of-the-mobilenetv2-inverted-residuals-and-linear-bottlenecks-paper-e19b187cb78a>. Acesso em: 6 nov. 2025.

BARBOSA, T. P. L. et al. Identificação de tumores cerebrais (glioma, meningioma, pituitário) em imagens de ressonância magnética utilizando rede neural convolucional. João Pessoa: IFPB, 2025. Trabalho de Conclusão de Curso.

ESTRATÉGIA MED. Resumo de Neuroglioma: riscos, tratamento e mais! Disponível em: <https://www.estategiamed.com.br/blog/neurologia/neuroglioma/>. Acesso em: 2 dez. 2025.

ESTRATÉGIA MED. Resumo sobre Meningioma: definição, manifestações clínicas e mais! Disponível em: <https://www.estategiamed.com.br/blog/neurologia/meningioma/>. Acesso em: 2 dez. 2025.

FLEURY MEDICINA E SAÚDE. Diagnóstico por imagem das doenças hipotálamo-hipofisárias. Disponível em: <https://www.fleury.com.br>. Acesso em: 2 dez. 2025.

ONCOLOGIA D'OR. Como identificar um tumor na ressonância magnética? Disponível em: <https://www.oncologidor.com.br/noticias/tumor-na-ressonancia/>. Acesso em: 2 dez. 2025.

SMITHUIS, R.; MONTANERA, W. Systematic Approach to Brain Tumors. THE RADIOLOGY ASSISTANT, 2008. Disponível em: <https://radiologyassistant.nl/neuroradiology/brain-tumor/systematic-approach>. Acesso em: 2 dez. 2025.

Apêndice A

GITHUB. leonfagundes/IC-backend (código-fonte). Disponível em: <https://github.com/leonfagundes/IC-backend>. Acesso em: 6 nov. 2025.

GITHUB. Código de treinamento – train_mobilenetv2.py. Disponível em:
https://github.com/leonfagundes/IC-backend/blob/main/src/train_mobilenetv2.py. Acesso em: 6 nov. 2025.

GITHUB. Modelo treinado – modelo_mobilenetv2_224x224_finetuned.keras. Disponível em:
https://github.com/leonfagundes/IC-backend/blob/main/model/modelo_mobilenetv2_224x224_finetuned.keras. Acesso em: 6 nov. 2025.

Apêndice B

DETECTIA. Website integrado à API de predição. Disponível em: <http://72.61.33.18:4001> ou <https://ic-frontend-six.vercel.app>. Acesso em: 6 nov. 2025.