

Redes neurais convolucionais na classificação de tumores

Maria Yessenia Alvarez Gil

Data: 01 de diciembre de 2024

1. Introdução

O câncer é um conjunto de doenças caracterizadas pelo crescimento descontrolado e pela proliferação de células anormais no corpo. Como mencionado por Ayadi et al. (2021), entre as formas mais graves dessa doença está o tumor cerebral, considerado um dos cânceres mais letais em todo o mundo, tanto em adultos quanto em crianças. Esse tipo de câncer pode se manifestar de diversas formas, dependendo de sua localização, textura e estrutura. Além disso, apresenta uma baixa taxa de sobrevivência, tornando a classificação incorreta de um tumor cerebral uma questão crítica, já que pode levar a consequências graves. Isso destaca a importância de uma identificação precisa e precoce para a escolha de um tratamento adequado.

A ressonância magnética (RM) é uma técnica eficaz para identificar e distinguir tumores cerebrais. No entanto, a análise manual dessas imagens é uma tarefa desafiadora, que exige experiência e conhecimento especializados do radiologista, nem sempre disponíveis, além de estar sujeita a erros humanos, especialmente devido ao grande volume de dados e à complexidade das características dos tumores. Nesse cenário, torna-se essencial o uso de um sistema automatizado de diagnóstico assistido por computador. Entre as técnicas avançadas de classificação de imagens, as redes neurais convolucionais (CNNs) se destacam por sua notável eficácia e precisão nessa área.

Neste trabalho, desenvolvido no âmbito do curso de *Tópicos especiais em estatística computacional*, exploramos a implementação de uma arquitetura baseada em redes neurais convolucionais para resolver o problema de classificação de tumores cerebrais em imagens de RM. Utilizamos o conjunto de dados público Brain Tumor Classification (MRI) para treinar e avaliar o modelo. Além disso, nossa proposta inclui o tratamento diferenciado dos dados, permitindo comparar o desempenho do modelo em diferentes pré-processamentos dos dados.

A estrutura deste relatório abrange desde os fundamentos teóricos e metodológicos utilizados, passando pelos resultados experimentais obtidos, até uma discussão sobre as implicações dos achados. Mostramos como a classificação automatizada de tumores cerebrais pode oferecer resultados convincentes e contribuir significativamente para o diagnóstico precoce e preciso dessa doença.

2. Conjunto de dados

Neste trabalho, utilizamos o conjunto de dados Brain Tumor Classification (MRI), disponibilizado na plataforma Kaggle por Bhuvaji et al. (2020). Este conjunto de dados foi acessado em 23 de novembro de 2024 e contém imagens de ressonância magnética organizadas para a classificação de tumores cerebrais.

O dataset contém um total de 3264 imagens, distribuídas em pastas de treinamento e teste. Cada uma dessas pastas está subdividida em quatro categorias de tumores: Tumor Glioma, Tumor Meningioma, Tumor Pituitary, No Tumor. Os tumores são definidos de acordo com Louis et al. (2016):

- **Tumor Glioma:** Tumor cerebral originado em células gliais, que pode ser de baixo ou alto grau, variando em agressividade e prognóstico.
- **Tumor Meningioma:** Tumor geralmente benigno que se desenvolve nas meninges, com sintomas que dependem do tamanho e da localização.

- **Tumor Pituitary:** Tumor benigno da glândula pituitária que pode causar disfunções hormonais e problemas visuais ou neurológicos.

No conjunto de dados, foram encontradas 181 imagens duplicadas, as quais foram removidas. A Figura 1 apresenta a distribuição do número de imagens por classe antes e após o ajuste realizado. As barras em azul claro representam os valores iniciais, enquanto as barras em azul escuro mostram os valores finais, destacando as proporções relativas de cada categoria. Observa-se uma redução na quantidade de imagens para a classe “Não tumor”, que passa de 15,3% para 13,2%, enquanto as demais categorias (“Tumor Pituitário”, “Tumor Meningioma” e “Tumor Glioma”) mantêm proporções semelhantes com pequenos ajustes.

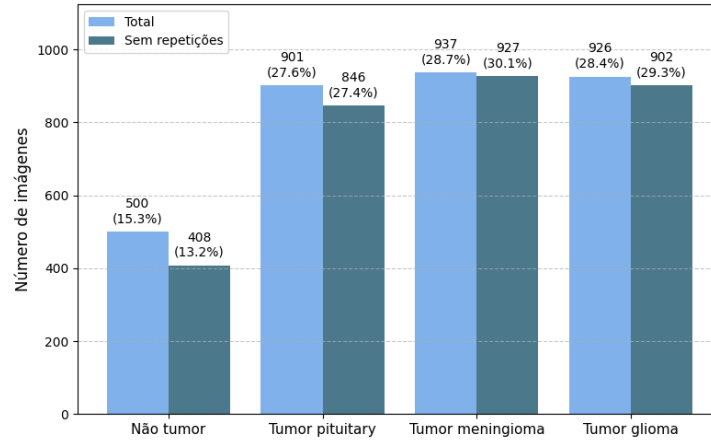


Figura 1: Distribuição do número de imagens por classe antes e após o ajuste no conjunto de dados, destacando as proporções relativas de cada categoria.

Abaixo, apresentamos um exemplo de imagem para cada classe:

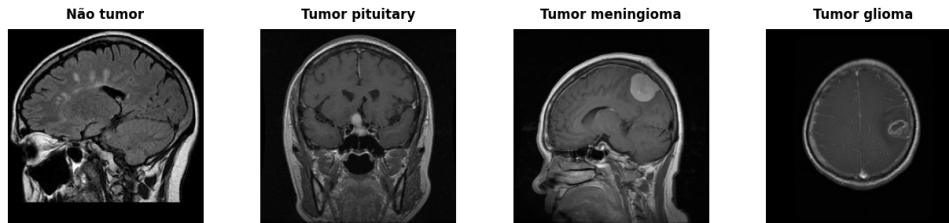


Figura 2: Exemplos de imagens para cada classe no conjunto de dados.

As imagens apresentam diferentes tamanhos e são compostas por três canais de cores (RGB). Os tamanhos mais frequentes no conjunto de dados são: (512, 512, 3) com 2302 imagens, (236, 236, 3) com 52 imagens e (225, 225, 3) com 37 imagens.

Na divisão dos dados, O conjunto de dados foi dividido da seguinte forma: 20% para o conjunto de teste. 80% para o conjunto de treinamento. Do total de treinamento, 20% foi reservado para validação.

3. Fundamentos teóricos e metodológicos

Nesta seção, detalhamos as principais etapas e conceitos utilizados no processamento e modelagem das imagens para a classificação de tumores cerebrais. As metodologias incluem o pré-processamento das imagens, a arquitetura de redes neurais convolucionais (CNNs) e estratégias de treinamento e otimização.

Pré-processamento dos dados

O pré-processamento é uma etapa essencial para garantir consistência nos dados de entrada e realçar características relevantes para o modelo. As etapas aplicadas foram:

- **Recorte das imagens:** Remoção de bordas desnecessárias para focar exclusivamente nas regiões de interesse. Isso reduz ruídos e facilita a aprendizagem do modelo.
- **Redimensionamento:** Todas as imagens foram ajustadas para um tamanho fixo, garantindo uniformidade na entrada do modelo. Isso é importante, pois as CNNs requerem tamanhos consistentes para realizar operações matriciais de forma eficiente.
- **Normalização:** Os valores dos pixels foram escalados para o intervalo $[0, 1]$ dividindo-os por 255 (valor máximo do RGB). Isso acelera o treinamento, melhora a estabilidade numérica e reduz o impacto de diferenças de brilho ou contraste entre imagens.

Redes neurais convolucionais (CNNs)

As redes neurais convolucionais (CNNs) são a base do modelo de classificação desenvolvido, sendo projetadas para processar dados estruturados em grades, como imagens, e extrair automaticamente informações relevantes nos dados. Um elemento fundamental dessas redes é a operação de convolução, que permite identificar padrões relevantes nas entradas. Matematicamente, a convolução entre uma imagem de entrada I e um filtro K de tamanho $M \times N$ é definida como:

$$S(i, j) = (I * K)(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i+m, j+n) \cdot K(m, n).$$

Em que $S(i, j)$ é o valor resultante da convolução no ponto (i, j) . $I(i+m, j+n)$ são os valores dos pixels da imagem de entrada. $K(m, n)$ filtro ou kernel. E $*$ é o operador de convolução. Essa operação consiste na multiplicação elemento a elemento entre uma região específica da imagem e o kernel, seguida por uma soma. O resultado é uma nova matriz (ou imagem) que destaca características extraídas, como bordas, texturas ou outros padrões fundamentais para o aprendizado da rede.

Estrutura de uma CNN

Camadas convolucionais: Extraem características locais da imagem. Quanto mais profundas forem as camadas, mais abstratas são as características aprendidas.

Camadas de pooling: Reduzem a dimensionalidade das representações intermediárias. O max-pooling, por exemplo, seleciona o valor máximo em uma região, preservando as informações mais relevantes.

Função de ativação ReLU: Introduce não linearidade ao modelo, substituindo valores negativos por zero. A ReLU acelera o treinamento e melhora a capacidade do modelo de aprender relações complexas.

Camadas totalmente conectadas: Integram as características aprendidas pelas camadas convolucionais e produzem uma representação compacta para a tomada de decisão final.

Camada de saída: A camada de saída possui tantos neurônios quanto o número de classes do problema. Para tarefas de classificação multiclasse, uma função de ativação frequentemente empregada é a softmax, que converte os valores de saída em probabilidades interpretáveis, garantindo que a soma seja 1, o que facilita a identificação da classe mais provável.

Otimização e métricas

Para o ajuste dos pesos do modelo, foi utilizado o otimizador **Adam**, amplamente reconhecido por sua eficiência e adaptabilidade. Ele combina os benefícios do momento e da adaptação dinâmica da taxa de aprendizado, facilitando a convergência durante o treinamento (Kingma & Ba, 2014).

As métricas utilizadas para avaliação foram:

- **Acurácia:** Mede a proporção de predições corretas em relação ao total, fornecendo uma visão geral do desempenho do modelo.

- **Loss (categorical cross-entropy):** Representa o erro médio durante o treinamento e teste, avaliando a discrepância entre as probabilidades previstas pelo modelo e os rótulos reais. Essa função é especialmente adequada para problemas de classificação multiclasse, pois mede a incerteza do modelo em relação às previsões corretas.

Estratégias para melhorar o desempenho

Dropout: Introduzido nas camadas totalmente conectadas, desativa aleatoriamente neurônios durante o treinamento. Isso reduz o risco de sobreajuste e melhora a generalização (Srivastava et al., 2014).

Aumento de dados: Realizou-se a geração de novas instâncias a partir de imagens existentes, aplicando transformações como rotação, inversão e translação. Essa técnica amplia o conjunto de treinamento e melhora a robustez do modelo (Shorten & Khoshgoftaar, 2019).

Para finalizar esta seção, vale ressaltar que o TensorFlow e o Keras foram as ferramentas fundamentais utilizadas para a implementação e treinamento do modelo. O TensorFlow é uma das bibliotecas de código aberto mais populares para a criação e treinamento de modelos de aprendizado de máquina, oferecendo alta flexibilidade e eficiência no processamento de dados complexos. O Keras, que funciona como uma interface de alto nível sobre o TensorFlow, facilita a construção e treinamento de redes neurais de maneira intuitiva, oferecendo funções predefinidas para arquiteturas de redes profundas, otimização de hiperparâmetros e avaliação de modelos. Essas ferramentas foram escolhidas devido ao amplo suporte da comunidade e à capacidade de lidar com grandes volumes de dados, sendo uma opção ideal para a classificação de imagens médicas (Abadi et al., 2016).

4. Aplicação

Para a implementação do modelo de classificação, iniciou-se com o processamento das imagens disponíveis. O conjunto de dados inicial continha um total de 3083 imagens após a eliminação de duplicatas. As seguintes etapas de pré-processamento foram realizadas:

Redimensionamento e conversão para escala de cinza: Todas as imagens foram redimensionadas para dimensões fixas de 225x225 pixels, baseando-se em Khan et al. (2021). Além disso, as imagens foram convertidas para escala de cinza, resultando em dimensões uniformes de (225, 225, 1).

Normalização: Os valores dos pixels das imagens foram normalizados dividindo-os por 255, para que os valores finais estivessem no intervalo de 0 a 1, facilitando o treinamento da rede.

Conjuntos de dados para treinamento

Foram criados três conjuntos de dados diferentes para avaliar o desempenho do modelo em diferentes situações de pré-processamento:

Conjunto 1: Incluiu as imagens preprocesadas como descrito acima.

Conjunto 2: Além das etapas do Conjunto 1, as bordas desnecessárias das imagens foram removidas para focar nas regiões relevantes.

Conjunto 3: A partir do Conjunto 2 (sem bordas), foi aplicado aumento de dados adicionando 150 imagens geradas aleatoriamente. Essas imagens adicionais foram criadas a partir de rotações e transformações leves para aumentar a variabilidade.

Arquitetura do Modelo

Ele começa com uma camada convolucional com 32 filtros de tamanho (3,3), função de ativação ReLU e uma entrada de dimensão (225, 225, 1). Em seguida, aplica uma operação de MaxPooling com uma janela (2,2), seguida por um Dropout com taxa de 0,2. Depois, uma segunda camada convolucional com 64 filtros de tamanho (3,3) é adicionada, seguida por outra operação de MaxPooling e Dropout. Na sequência, é incorporada uma terceira camada convolucional com 128 filtros, seguida novamente por MaxPooling e Dropout. A quarta camada convolucional contém 512 filtros e é acompanhada por outra

operação de MaxPooling e Dropout. Após as camadas convolucionais, a saída é achatada por meio da camada Flatten, que é conectada a uma camada densa com 512 neurônios e ativação ReLU. Finalmente, a rede termina com uma camada densa de saída com 4 neurônios e ativação softmax para classificação multiclasse.

Avaliação e resultados

Os modelos foram treinados por 10 épocas devido às limitações de recursos computacionais e tempo disponível. Após o treinamento de cada modelo com os três conjuntos de dados, foram obtidos os seguintes resultados de desempenho:

Modelo	Acurácia no Teste	Loss no Teste
Modelo 1	0,81	0,69
Modelo 2	0,88	0,49
Modelo 3	0,86	0,59

Tabela 1: Desempenho dos modelos acurácia e o loss no conjunto de teste para cada modelo.

Os resultados indicam que o Modelo 2 apresentou o melhor desempenho, com a maior acurácia no teste (0,88) e o menor valor de loss (0,49). Na Figura 3, são exibidas as curvas de acurácia e perda (loss) durante o treinamento e a validação dos modelos. Observa-se que, apesar do número reduzido de épocas, os modelos demonstraram boa capacidade de generalização.

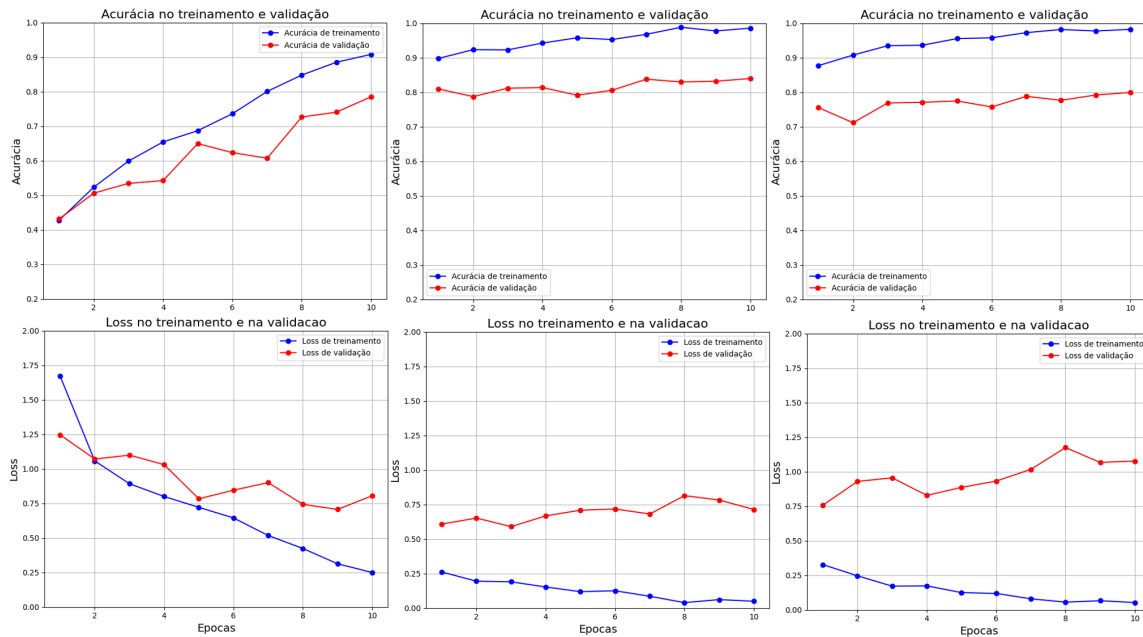


Figura 3: Curvas de acurácia (superior) e loss (inferior) ao longo do treinamento (linha azul) e validação (linha vermelha) dos três modelos. A primeira coluna corresponde ao Modelo 1, a segunda ao Modelo 2, e a terceira ao Modelo 3.

Adicionalmente, a Figura 4 apresenta as matrizes de confusão para os três modelos. A matriz do Modelo 2 destaca-se por apresentar o menor número de erros de classificação, evidenciando sua eficácia na tarefa de classificação.

Embora o Modelo 3 treinado com o conjunto aumentado tenha obtido um desempenho similar ao Modelo 2, esperava-se que o aumento de dados trouxesse melhores resultados. Isso pode ser atribuído a fatores como a baixa representatividade das transformações aplicadas nos dados aumentados, que podem ter introduzido ruído, ou ao número limitado de épocas, que pode ter impedido o modelo de aproveitar completamente a variabilidade adicional. Mesmo assim, o desempenho geral foi consistente, reforçando a eficácia do pré-processamento aplicado.

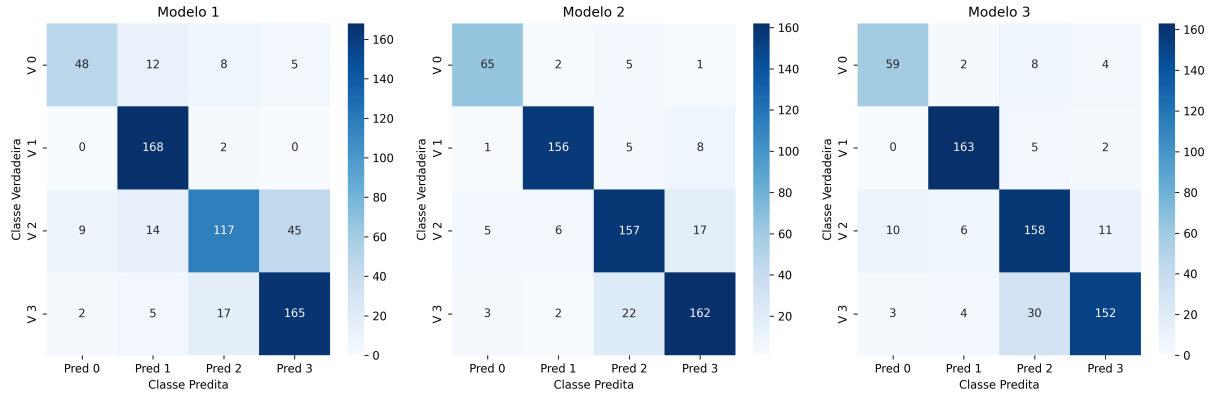


Figura 4: Matrizes de confusão dos três modelos avaliados.

Reprodutibilidade: Para mais detalhes, o código utilizado para o pré-processamento, treinamento, avaliação, geração dos gráficos e das matrizes de confusão está disponível no Google Colab, possibilitando a reprodução e análise completa dos experimentos.

5. Conclusões

As redes neurais convolucionais demonstraram ser ferramentas eficazes na classificação de imagens médicas, com destaque para o Modelo 2, que alcançou uma acurácia de 88% e o menor valor de perda (0,49). Esses resultados evidenciam seu potencial em problemas complexos, mas também ressaltam a importância de continuar aprimorando a precisão e reduzindo taxas de erro para garantir confiabilidade no setor médico.

Observou-se que o aumento de dados nem sempre melhora o desempenho, reforçando a necessidade de gerar dados sintéticos de alta qualidade. Além disso, a remoção de bordas irrelevantes foi essencial para otimizar os resultados, destacando o impacto do pré-processamento adequado na eficácia dos modelos. Estudos futuros devem focar em ajustes arquiteturais e no uso de mais recursos computacionais para alcançar melhores resultados.

Referências

- [1] Abadi, M. (2016). Tensorflow: learning functions at scale. In *Proceedings of the 21st ACM SIGPLAN international conference on functional programming*, pages 1–1.
- [2] Ayadi, W., Elhamzi, W., Charfi, I., and Atri, M. (2021). Deep CNN for brain tumor classification. *Neural processing letters*, 53:671–700.
- [3] Bengio, Y., Goodfellow, I., and Courville, A. (2017). Deep learning, volume 1. MIT press Cambridge, MA, USA.
- [4] Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., and Kanchan, S. (2020). Brain tumor classification (MRI). URL <https://www.kaggle.com/dsv/1183165>.
- [5] Khan, H. A., Jue, W., Mushtaq, M., and Mushtaq, M. U. (2021). Brain tumor classification in MRI image using convolutional neural network. *Mathematical Biosciences and Engineering*.
- [6] Kingma, D. P. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [7] Louis, D. N., Perry, A., Reifenberger, G., Von Deimling, A., Figarella-Branger, D., Cavenee, W. K., Ohgaki, H., Wiestler, O. D., Kleihues, P., and Ellison, D. W. (2016). The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131:803–820.
- [8] Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48.