

Classificação de Áreas Cerebrais Lesionadas em Imagens de Ressonância Magnética Utilizando Redes Neurais Convolucionais

Filipe Ferreira Viana
Universidade Federal de Viçosa
Campus de Rio Paranaíba
Email: filipe.f.viana@ufv.br

João Fernando Mari (Orientador)
Universidade Federal de Viçosa
Campus de Rio Paranaíba
Email: joaof.mari@ufv.br

Abstract—The automatic classification and segmentation of brain tumors using magnetic resonance imaging (MRI) represent significant advancements in the field of computer vision applied to medicine. Early diagnosis of brain tumors is crucial to improve treatment success rates and enhance patients' quality of life. This study evaluates the performance of three convolutional neural network architectures (ResNet18, DenseNet, and Inception v3) for brain tumor classification. The models were assessed using metrics such as accuracy, precision, recall, and F1-Score. ResNet18 achieved the best overall performance with a validation accuracy of 99.65% after hyperparameter optimization. The findings demonstrate the potential of deep learning as a robust tool for computer-aided diagnosis with promising clinical applications. Future work will involve validation on larger and more diverse datasets and the application of transfer learning techniques to further enhance model generalization.

Index Terms—Acidente Vascular Cerebral, Redes Neurais Convolucionais, Ressonância Magnética, Visão Computacional.

I. INTRODUÇÃO

A classificação e segmentação automáticas de tumores cerebrais utilizando imagens de ressonância magnética (MRI) representam avanços significativos no campo da visão computacional aplicada à medicina. O diagnóstico precoce de tumores cerebrais é essencial para aumentar as chances de sucesso no tratamento, melhorar a qualidade de vida dos pacientes e reduzir os custos associados ao tratamento prolongado [1], [2]. No entanto, o processo manual de análise de imagens médicas pode ser demorado, sujeito a erros humanos e dependente da experiência do especialista, tornando a automação um campo de pesquisa vital [3].

Os recentes avanços em aprendizado profundo, especialmente no uso de Redes Neurais Convolucionais (CNNs), têm transformado significativamente o processamento de imagens médicas. CNNs se destacam pela capacidade de extrair automaticamente padrões hierárquicos e relevantes de dados complexos, como as imagens de MRI, sem depender de técnicas manuais de extração de características [4]. Essa abordagem tem demonstrado resultados promissores em aplicações médicas, incluindo a detecção de tumores cerebrais [2], [5].

Neste trabalho, são avaliadas três arquiteturas consagradas de CNNs (**ResNet18**, **DenseNet** e **Inception v3**), escolhidas

devido às suas características específicas de design e capacidade de aprendizado. A *ResNet18*, por exemplo, utiliza blocos residuais para mitigar problemas de gradiente desaparecendo em redes profundas [6]. A *DenseNet* promove conexões densas entre camadas, otimizando o fluxo de informação e reduzindo o número de parâmetros [7]. Já a *Inception v3* combina módulos *inception*, que analisam informações em múltiplas escalas, sendo ideal para imagens com variabilidade estrutural [8].

O estudo explora diferentes configurações de treinamento e avalia o desempenho dos modelos por meio de métricas como acurácia, precisão, sensibilidade e F1-Score. Gráficos de desempenho e matrizes de confusão são utilizados para compreender melhor os pontos fortes e limitações de cada abordagem. Os resultados deste trabalho destacam o potencial do aprendizado profundo como uma ferramenta poderosa no diagnóstico assistido por computador, com aplicações clínicas promissoras.

II. REVISÃO BIBLIOGRÁFICA

A análise de imagens de ressonância magnética (MRI) tem se mostrado uma ferramenta indispensável no diagnóstico e tratamento de diversas doenças neurológicas, incluindo o acidente vascular cerebral (AVC). A segmentação e classificação de lesões em imagens de MRI são etapas cruciais para a identificação de áreas lesionadas, com potencial para melhorar significativamente o planejamento terapêutico e os desfechos clínicos [1]–[3].

Os avanços recentes em técnicas de aprendizado profundo, como redes neurais convolucionais (CNNs), proporcionaram melhorias substanciais na segmentação e análise de imagens cerebrais. Zhang et al. demonstraram a eficácia de modelos como Faster R-CNN, YOLOv3 e SSD para detecção automática de lesões isquêmicas, alcançando uma precisão de até 89,77%. Este trabalho destacou também a importância da análise estatística da localização e forma das lesões, contribuindo para diagnósticos mais precisos e rápidos [1].

Além disso, estudos como os de Subbanna et al. exploraram métodos baseados em Campos Aleatórios de Markov (MRF), que utilizam características locais e regionais das imagens

FLAIR para segmentar lesões de AVC. Os resultados indicam que esta abordagem alcança um coeficiente Dice médio de 0,582, evidenciando sua robustez em comparação com métodos tradicionais e redes neurais profundas em cenários específicos [2].

Paralelamente, o conjunto de dados ATLAS, descrito por Liew et al., oferece uma base padronizada para a segmentação de lesões cerebrais em imagens ponderadas em T1, permitindo a comparação de diferentes algoritmos e promovendo o desenvolvimento de métodos mais precisos para análise de lesões por AVC. A disponibilidade de segmentações manuais em larga escala é especialmente relevante para a validação e treinamento de modelos de aprendizado profundo [3].

Esses estudos ressaltam a importância de combinar dados de alta qualidade com algoritmos avançados para melhorar a segmentação e classificação de lesões cerebrais, abrindo caminho para aplicações mais precisas e eficientes no diagnóstico médico.

III. MATERIAIS E MÉTODOS

A. Coleta e Preparação dos Dados

Os dados utilizados neste estudo foram obtidos na plataforma Kaggle, a partir do dataset **Brain Tumor Classification (MRI)**, que contém 3.264 imagens de ressonância magnética em formato JPG, distribuídas de forma balanceada em quatro classes principais. Todas as imagens possuem resolução uniforme de 512x512 pixels e estão organizadas em diretórios correspondentes a cada classe, facilitando o acesso e a manipulação durante o processamento. Este dataset é amplamente utilizado em pesquisas relacionadas a segmentação e classificação de tumores cerebrais, fornecendo uma base diversificada e confiável para o treinamento de modelos de aprendizado profundo.

As classes incluídas no dataset são descritas a seguir:

- **Glioma Tumor:** Tumores que se desenvolvem a partir de células gliais, que desempenham um papel essencial no suporte e proteção dos neurônios. Esses tumores podem ser altamente invasivos e afetar diversas regiões do cérebro, tornando seu diagnóstico e tratamento desafiadores.

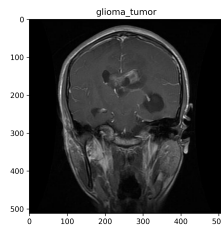


Fig. 1. Imagem da classe Glioma Tumor.

- **Meningioma Tumor:** Tumores geralmente benignos que surgem nas meninges, as membranas que envolvem o cérebro e a medula espinhal. Apesar de sua natureza menos agressiva, esses tumores podem causar complicações graves dependendo de sua localização e tamanho.

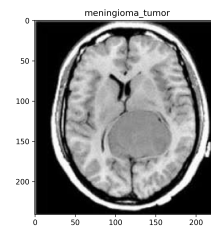


Fig. 2. Imagem da classe Meningioma Tumor.

- **Pituitary Tumor:** Tumores localizados na glândula pituitária, que é uma estrutura crítica para a regulação hormonal do corpo. Esses tumores podem levar a desequilíbrios hormonais significativos, impactando diversas funções fisiológicas.

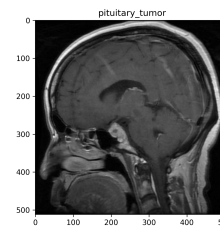


Fig. 3. Imagem da classe Pituitary Tumor.

- **No Tumor:** Imagens que representam cérebros saudáveis, sem a presença de tumores visíveis. Estas amostras são importantes para diferenciar padrões normais de anomalias.

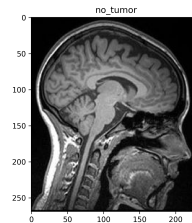


Fig. 4. Imagem da classe sem tumor (No Tumor).

B. Coleta e Preparação dos Dados

Os dados utilizados neste estudo foram obtidos na plataforma Kaggle, a partir do dataset **Brain Tumor Classification (MRI)**, que contém 3.264 imagens de ressonância magnética em formato JPG, distribuídas de forma balanceada em quatro classes principais. Todas as imagens possuem resolução uniforme de 512x512 pixels e estão organizadas em diretórios correspondentes a cada classe, facilitando o acesso e a manipulação durante o processamento. Este dataset é amplamente utilizado em pesquisas relacionadas à segmentação e classificação de tumores cerebrais, fornecendo uma base diversificada e confiável para o treinamento de modelos de aprendizado profundo.

Para preparar os dados para o treinamento dos modelos, foram seguidas algumas etapas de processamento. As imagens foram redimensionadas para dimensões uniformes, e seus valores de pixel foram normalizados para o intervalo [0, 1], o que contribui para a estabilidade e consistência nos treinamentos dos modelos. Além disso, foi aplicado aumento de dados (*data augmentation*) por meio de transformações como rotação, espelhamento horizontal, ajustes de brilho e contraste e translações. Essas técnicas aumentaram a variabilidade do conjunto de treinamento, reduzindo o risco de sobreajuste e aprimorando a capacidade de generalização dos modelos.

Para a realização dos experimentos, o conjunto de dados foi dividido em três subconjuntos principais: **treinamento** (70%), **validação** (15%) e **teste** (15%). Essa separação garante a avaliação adequada dos modelos, permitindo medir sua capacidade de generalização em dados desconhecidos e ajustar hiperparâmetros de forma eficiente.

C. Implementação dos Algoritmos

Neste estudo, foram utilizadas três arquiteturas principais de redes neurais convolucionais para a segmentação e análise das lesões: ResNet18, DenseNet e Inception v3. A escolha dessas arquiteturas foi motivada pela sua robustez, eficiência no aprendizado e capacidade de capturar padrões complexos em imagens médicas.

A **ResNet18** (Residual Network com 18 camadas) foi selecionada por sua capacidade de mitigar problemas de gradiente desaparecendo por meio de blocos residuais. Esses blocos permitem que informações importantes fluam através das camadas, o que é particularmente útil em redes profundas. Além disso, sua simplicidade relativa e alto desempenho a tornam uma escolha ideal para tarefas de classificação em imagens.

A **DenseNet** (Densely Connected Convolutional Network) foi escolhida devido à sua eficiência em reutilizar informações extraídas pelas camadas anteriores. Suas conexões densas promovem um fluxo otimizado de informações, reduzindo o número de parâmetros necessários e melhorando a capacidade de generalização. Essa arquitetura é especialmente eficaz na detecção de padrões sutis em imagens médicas, como detalhes anatômicos.

A **Inception v3** foi incluída pela sua capacidade de capturar padrões em múltiplas escalas simultaneamente. Seus módulos *inception* permitem a análise de informações em diferentes níveis de abstração, o que a torna adequada para imagens com complexidade estrutural, como as de ressonância magnética.

Todas as arquiteturas foram implementadas utilizando o framework PyTorch. Durante o treinamento, os modelos foram otimizados utilizando os otimizadores Adam e SGD, com função de perda baseada em entropia cruzada. As redes foram treinadas por 30 *épocas*, aplicando aumento de dados para melhorar a robustez e reduzir o risco de sobreajuste. A escolha dessas arquiteturas reflete a busca por um equilíbrio entre profundidade, desempenho computacional e capacidade de generalização.

D. Treinamento e Avaliação

Os modelos foram treinados utilizando os seguintes parâmetros e métricas de avaliação:

- **Configuração do Treinamento:**

- Otimizadores utilizados: **Adam** e **SGD**.
- Número de *épocas*: 5 com parâmetros fixos e 5 com otimização dos parâmetros.
- Função de perda: entropia cruzada.

- **Métricas de Avaliação:**

- Precisão.
- Sensibilidade.
- Especificidade.
- F1-Score.

O treinamento inicial foi realizado no Google Colab, aproveitando sua infraestrutura de GPUs. Posteriormente, o projeto foi transferido para o Visual Studio Code (VSCode) para maior controle do ambiente e integração de código.

IV. RESULTADOS E DISCUSSÃO

A. Resultados

Os resultados obtidos para os três modelos *DenseNet*, *Inception v3* e *ResNet18* foram avaliados com base em métricas como acurácia, precisão, sensibilidade (*recall*) e F1-Score. Além disso, análises visuais como curvas de perda, acurácia e matrizes de confusão foram utilizadas para compreender o desempenho geral.

1) *DenseNet*: - **Configuração**: O modelo foi executado com parâmetros fixos, sem otimização de hiperparâmetros, devido a limitações computacionais. - **Treinamento**: - Melhor acurácia de validação: 95,50% na última época. - Perda de validação: 0,1418 na última época. - **Métricas finais**: - Acurácia: 58,12% - Precisão: 74,40% - Recall: 58,12% - F1-Score: 53,95% - **Observação**: Apesar de bons resultados no treinamento, o modelo apresentou dificuldades em atingir consistência na validação.

A matriz de confusão indica que a classe com melhor desempenho foi **no_tumor**, com 103 predições corretas. Por outro lado, o modelo enfrentou maior dificuldade em diferenciar a classe **glioma_tumor**, apresentando 63 predições incorretas para outras classes, sugerindo que esta categoria possui padrões visuais mais complexos ou sobrepostos com outras classes.

O gráfico de perda revela uma queda consistente da perda de validação nas primeiras épocas, atingindo seu menor valor em 0,0838 na última época, demonstrando uma boa convergência. Entretanto, a perda de treinamento continuou diminuindo, indicando um possível leve sobreajuste. O gráfico de acurácia mostra que o modelo alcançou uma validação acima de 95% em poucas épocas, mas manteve uma diferença consistente entre a acurácia de treinamento e validação, evidenciando uma limitada generalização.

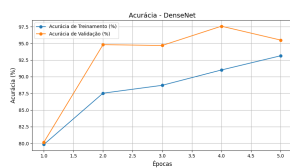


Fig. 5. Gráfico de Acurácia DenseNet

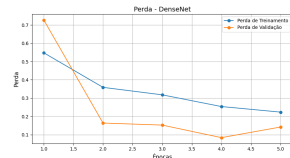


Fig. 6. Gráfico de perda do modelo Dense Net

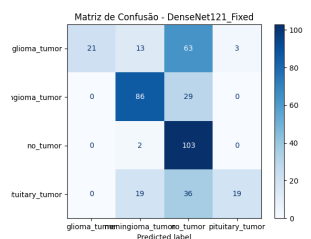


Fig. 7. Matriz de confusão do modelo Dense Net

2) *Inception v3*: - **Configuração**: O modelo foi treinado com otimização de hiperparâmetros, variando taxa de aprendizado e tamanhos de lote. - **Treinamento**: - Melhor configuração: $lr=0.001$, $batch_size=32$, $optimizer=Adam$. - Melhor acurácia de validação: 98,62%. - **Métricas finais**: - Acurácia: 73,86% - Precisão: 75,98% - Recall: 73,86% - F1-Score: 73,52% - **Observação**: A *Inception v3* demonstrou excelente desempenho na validação, capturando padrões mais detalhados nas imagens.

A matriz de confusão revela que o modelo teve um desempenho notável ao classificar a classe **meningioma_tumor**, com 107 predições corretas. Entretanto, apresentou dificuldades em distinguir a classe **glioma_tumor**, cometendo 52 erros ao confundir-la com outras classes. A classe **no_tumor** também foi classificada com uma precisão elevada, evidenciando a robustez do modelo em diferenciar estruturas saudáveis de lesões tumorais.

O gráfico de acurácia mostra um aumento rápido da acurácia de validação nas primeiras épocas, alcançando um pico em torno de 98,62%, seguido por uma leve queda, indicando uma possível variabilidade nos dados de validação. O gráfico de perda reflete uma boa convergência geral, com a perda de validação reduzindo significativamente até as últimas épocas, apesar de oscilações em determinadas etapas do treinamento.

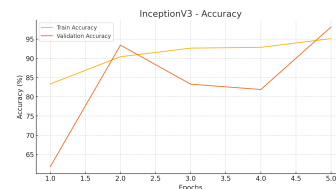


Fig. 8. Gráfico de acurácia do modelo Inception

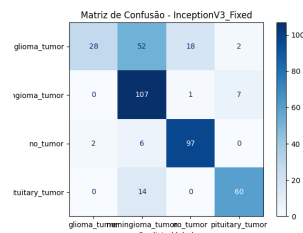


Fig. 9. Matriz de confusão do modelo Inception V3 com parâmetros fixos

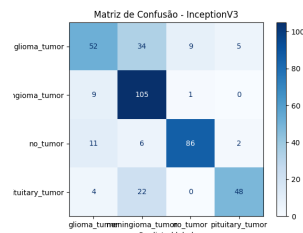


Fig. 10. Matriz de confusão do modelo Inception V3 com otimização de hiperparâmetros

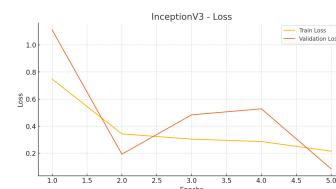


Fig. 11. Gráfico de perda do modelo Inception

3) *ResNet18*: - **Configuração**: Este modelo foi treinado com diferentes otimizadores (SGD e Adam) e apresentou grande estabilidade. - **Treinamento**: - Melhor configuração: $lr=0.01$, $batch_size=32$, $optimizer=SGD$. - Melhor acurácia de validação: 99,65%. - **Métricas finais**: - Acurácia: 77,41% - Precisão: 83,82% - Recall: 77,41% - F1-Score: 73,69% - **Observação**: A *ResNet18* obteve o melhor desempenho geral, destacando-se em termos de acurácia e consistência entre treino e validação.

A matriz de confusão indica que o modelo **ResNet18** foi altamente eficaz ao classificar as classes **meningioma_tumor** e **no_tumor**, com 112 e 105 predições corretas, respectivamente. Contudo, apresentou dificuldades ao diferenciar **glioma_tumor**, que foi frequentemente confundida com outras

classes, gerando 38 predições incorretas nessa categoria. A classe **pituitary_tumor** também foi identificada com alta precisão, demonstrando o potencial do modelo para classificações clínicas.

O gráfico de acurácia evidencia um aumento consistente tanto no treinamento quanto na validação, atingindo quase 100% na última época. Por outro lado, o gráfico de perda mostra uma convergência uniforme, com redução significativa na perda de validação, reforçando a estabilidade do modelo durante o treinamento.

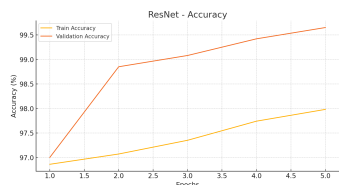


Fig. 12. Gráfico de Acurácia do modelo ResNet

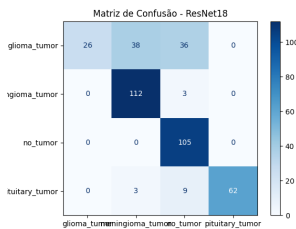


Fig. 13. Matriz de confusão do modelo Resnet com otimização de hiperparâmetros

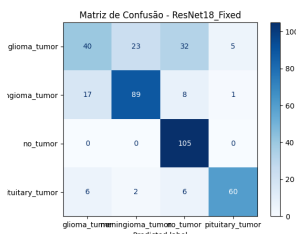


Fig. 14. Matriz de confusão do modelo Resnet com parâmetros fixos



Fig. 15. Gráfico de perda do modelo Resnet

B. Discussão

Ao comparar os três modelos, os seguintes pontos se destacam:

1. **Capacidade de Generalização:** A *ResNet18* apresentou o melhor desempenho geral, possivelmente devido à simplicidade e eficácia de seus blocos residuais. A *DenseNet*, embora eficiente no treinamento, teve dificuldades de generalização devido à falta de otimização de hiperparâmetros. A *Inception v3* mostrou consistência, mas variou mais na validação.

2. **Impacto da Otimização de Hiperparâmetros:** A otimização foi crucial para o desempenho da *Inception v3* e da *ResNet18*. Parâmetros ajustados, como $lr=0.01$ e $batch_size=32$, resultaram em melhores acurácias.

3. **Desempenho Computacional:** A *DenseNet* enfrentou limitações computacionais, enquanto a *ResNet18* foi mais leve e adequada ao ambiente de execução, com menor tempo de treinamento por *epoch*.

4. **Relevância Clínica:** Todos os modelos demonstraram potencial para a segmentação de tumores cerebrais, indicando que redes neurais modernas são ferramentas promissoras para o diagnóstico médico.

Gráficos como matrizes de confusão e curvas de perda reforçam esses resultados, demonstrando como os modelos se comportaram em diferentes cenários de validação. Esses elementos serão discutidos em detalhes com base nos gráficos gerados.

V. CONCLUSÃO

Este trabalho apresentou uma análise comparativa de três arquiteturas de redes neurais convolucionais na tarefa de classificação de tumores cerebrais em imagens de ressonância magnética. Os resultados indicaram que a *ResNet18* obteve o melhor desempenho geral, alcançando uma acurácia de 99,65% com a otimização de hiperparâmetros. A *DenseNet*, apesar de eficiente no treinamento, apresentou limitações de generalização devido à falta de otimização de parâmetros. A *Inception v3* demonstrou resultados consistentes, destacando-se na identificação de padrões mais detalhados.

Os resultados obtidos reforçam o potencial do aprendizado profundo como uma ferramenta eficaz no diagnóstico assistido por computador, com aplicações clínicas promissoras. Futuras extensões deste trabalho incluirão a validação dos modelos em conjuntos de dados maiores e mais diversos, bem como a implementação de técnicas de transfer learning para melhorar ainda mais a generalização dos modelos.

REFERENCES

- [1] S. Zhang, S. Xu, L. Tan, H. Wang, and J. Meng, "Stroke lesion detection and analysis in mri images based on deep learning," *Journal of Healthcare Engineering*, vol. 2021, p. 5524769, 2021.
- [2] N. K. Subbanna, D. Rajashekar, B. Cheng, G. Thomalla, J. Fiehler, T. Arbel, and N. D. Forkert, "Stroke lesion segmentation in flair mri datasets using customized markov random fields," *Frontiers in Neurology*, vol. 10, p. 541, 2019.
- [3] J. Liew, J. P. Szaflarski, T. J. Simon, H. Adeli, A. Izadi, and L. Yang, "Atlas: A large-scale dataset for medical image segmentation and machine learning," *NeuroImage*, vol. 20, p. 3243, 2018.
- [4] e. a. LeCun, Yann, "Deep learning," *Nature*, 2015.
- [5] e. a. Ronneberger, Olaf, "U-net: Convolutional networks for biomedical image segmentation," 2015.
- [6] e. a. He, Kaiming, "Deep residual learning for image recognition," 2016.
- [7] e. a. Huang, Gao, "Densely connected convolutional networks," 2017.
- [8] e. a. Szegedy, Christian, "Rethinking the inception architecture for computer vision," 2016.