

Segmentação e Classificação de Tumores Cerebrais a partir de MRI utilizando Redes Neurais Convolucionais

Guilherme S.G. Brotto

Departamento de Informática

Universidade Federal do Espírito Santo

Vitória, ES, Brasil

guilherme.brotto@edu.ufes.br

Marlon Moratti do Amaral

Departamento de Informática

Universidade Federal do Espírito Santo

Vitória, ES, Brasil

marlon.amaral@edu.ufes.br

Nilo Garcia Monteiro

Departamento de Informática

Universidade Federal do Espírito Santo

Vitória, ES, Brasil

nilo.monteiro@edu.ufes.br

Abstract—Brain tumor is caused by fast and uncontrolled cell proliferation in the brain. It is one of the deadliest diseases that exist and its treatment depends strongly on an early detection in order to save many lives. A specialist with a thorough understanding of brain illnesses is needed to manually identify the proper type of brain tumor. Additionally, processing many images takes time and it is tiresome. Therefore, automatic segmentation and classification techniques are required to speed up and enhance the diagnosis of brain tumors. This work focus on applying these techniques on a real problem to identify the kind of brain cancer from images of Magnetic Resonance Imaging. The results show that in both classification and segmentation tasks, the neural network detected correctly the kind of tumor in most of the cases.

Index Terms—deep learning, convolutional neural networks, brain tumor, classification, segmentation, deep learning, convolutional neural networks, brain tumor, classification, segmentation

I. INTRODUÇÃO

Segundo informações do National Cancer Institute, apenas 33,4% dos indivíduos com algum tipo de câncer no cérebro sobreviveram 5 anos ou mais após o diagnóstico. Além disso, dos sobreviventes, 77% foram diagnosticados em um estágio inicial, em que o tumor ainda não se espalhou e continua localizado no cérebro. Desse modo, fica evidente a importância de identificar a presença de células malignas com antecedência a fim de evitar agravamentos irreversíveis. Todavia, em grande parte dos casos o diagnóstico manual é complicado, dependendo fortemente da disponibilidade de um profissional experiente, o qual ainda está sujeito a cometer equívocos. Logo, o estudo de tecnologias que permitam automatizar a detecção de diferentes tipos de câncer é muito importante, visando diminuir a quantidade de erros de diagnose.

Diversos estudos vêm sendo realizados nessa área. Em Lundervold et al. [1], foi estudado de modo genérico como o uso de aprendizado profundo pode e deve ser aplicado na área da saúde para detecção de doenças a partir de MRI. Ademais, com foco no estudo de doenças no cérebro, Yamanakkanavar et al. [2] discute como arquiteturas de redes neurais convolucionais podem ser usadas para analisar a estrutura anatômica do

cérebro e identificar traços de Alzheimer, realizando uma classificação e segmentação também baseado em MRI. Por fim, em Ghassemi et al. [3], ırmak [4] e Aamir et al. [5], são feitos estudos com objetivo de detectar e prever diferentes tumores cerebrais a partir de aprendizado profundo.

Com isso, esse trabalho propõe técnicas para identificar tumores cerebrais a partir de imagens geradas por ressonância magnética. Para isso, serão utilizadas técnicas de aprendizado profundo, utilizando de Redes Neurais Convolucionais (CNNs) para classificar e segmentar as imagens, ou seja, identificar o tipo de tumor e sua localização na imagem.

II. REFERENCIAL TEÓRICO

Esta seção discute conceitos importantes que definem as características gerais e exploram as dificuldades no universo do problema.

O processo de identificar a existência de tumores na região cerebral de um indivíduo envolve, em grande parte, o diagnóstico de imagens. Dessa forma, técnicas não-invasivas são aplicadas para gerar representações do corpo para análise posterior. Em regiões cerebrais, imagens de ressonância magnética (MRI) são utilizadas para aumentar a visibilidade do local de operação. Portanto, o equipamento utilizado no processo não utiliza de radiação, mas sim campos magnéticos.

O método de extração de imagens inicia-se com o alinhamento dos prótons com o campo magnético do equipamento. Em seguida, é executado um distúrbio artificial por meio de pulsos de radiofrequência, deixando os prótons contra o campo alinhado. Tal turbulência incentiva o retorno dos prótons à sua configuração anterior por meio do relaxamento, liberando energia em forma de ondas de rádio, a qual será captada pelos sensores do equipamento, gerando as imagens [6].

Para fornecer melhor visão ao problema tratado, podem ser configurados diferentes métodos de aplicação e análise dos pulsos, e alguns ponderamentos são utilizados com base nas necessidades de cada problema: *T1*, *T2* e *Flair*. A diferença entre eles vêm da frequência entre a emissão dos pulsos, o que gera alterações no contraste dos tecidos analisados. Resumidamente, imagens ponderadas em *T1* realçam tecidos

que são rapidamente magnetizados, enquanto as imagens ponderadas em T_2 apresentam foco em tecidos que perdem sua magnetização mais rapidamente. E o Flair, por si, tende a anular sinais de fluídos, ignorando o líquido cefalorraquidiano [7].

Outra possível adição no processo de ressonância magnética é o uso de agentes de contraste, os quais são utilizados para oferecer melhor definição de imagem. Considerando tumores cerebrais, um tipo de ponderamento comumente utilizado é o T_{1ce} o qual usa o metal gadolínio como agente de contraste, facilitando a visualização de tumores por se tornarem tecidos mais facilmente magnetizados com sua adição [8].

Por fim, com as informações obtidas, o equipamento pode retornar três tipos de cortes (axial, sagital e coronal) em que cada tipo de corte têm como objetivo fornecer diferentes visualizações em planos do corpo.

III. METODOLOGIA

Esta seção descreve a abordagem proposta para a segmentação de tumores cerebrais em imagens de ressonância magnética.

A. Dados de entrada e saída

Os dados de entrada consistem em imagens 2D de ressonância magnética do cérebro de pacientes, capturadas em um dos três planos anatômicos: axial, sagital ou coronal.

Os dados de saída são máscaras de segmentação de um único canal que segmentam o tumor, e um rótulo que pode assumir os seguintes valores:

- 0) Glioma
- 1) Meningioma
- 2) Tumor Pituitário

Uma visualização mais detalhada pode ser observada na figura 1, onde apresentamos exemplos representativos das imagens de entrada com destaque para os tumores.

B. Pré-processamento e aumento de Dados

O conjunto de dados passou por uma série de transformações. Primeiramente, as imagens, que inicialmente estavam em escala de cinza, foram convertidas para o espaço de cores RGB no início do processo de pré-processamento. Esta conversão foi realizada replicando o canal de intensidade para os canais de vermelho, verde e azul, garantindo que todas as imagens estejam no formato esperado pela rede neural convolucional.

Em seguida, todas as imagens foram redimensionadas para uma resolução de 224×224 pixels. Este redimensionamento é fundamental para padronizar o tamanho das imagens e garantir compatibilidade com modelos pré-treinados na ImageNet, que frequentemente utilizam imagens nesta resolução durante o treinamento. Além do redimensionamento, as imagens foram normalizadas utilizando médias de $[0.485, 0.456, 0.406]$ e desvios padrão de $[0.229, 0.224, 0.225]$, conforme padrões estabelecidos pelo conjunto de dados ImageNet [9].

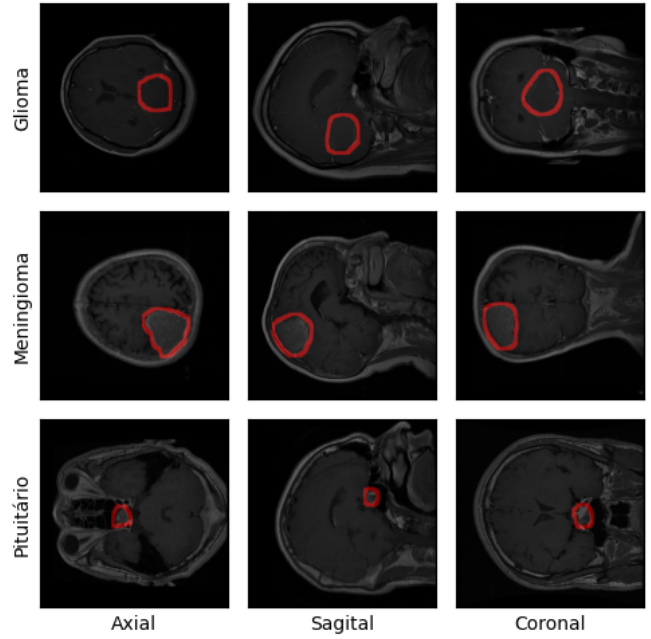


Fig. 1. Exemplos dos três tipos de tumores cerebrais (glioma, meningioma e pituitário) destacados nos três cortes anatômicos: axial, sagital e coronal.

Para enriquecer a diversidade do conjunto de treinamento, foram utilizadas técnicas de aumento de dados. Foram aplicadas diferentes transformações nas imagens, incluindo ajustes aleatórios de brilho, contraste, escala e translação. Estas transformações se mostraram adequadas para o contexto em questão.

A tabela I resume as transformações aplicadas com as funções da biblioteca de código aberto Albumentations [10], juntamente com os seus respectivos parâmetros.

TABLE I
TRANSFORMAÇÕES DA BIBLIOTECA ALBUMENTATIONS APLICADAS NO CONJUNTO DE DADOS

Função	Parâmetros
Affine	scale=(0.8, 1.2), translate_percent=0.125
ColorJitter	brightness=(0.6, 1.4), contrast=(0.6, 1.4)
Normalize	mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]
Resize	width=224, height=224

C. Rede Neural

Para a realização da tarefa proposta, utilizamos a rede DeepLabV3 com backbone ResNet-101, implementada na biblioteca de código aberto PyTorch [11]. A DeepLabV3 é uma arquitetura de rede neural convolucional projetada especificamente para segmentação semântica, que aproveita convoluções

dilatadas (convoluções atours) para capturar contexto em diversas escalas [12]. Utilizamos a cabeça de segmentação da DeepLabV3 e a cabeça padrão de classificação da ResNet-101, esta última conhecida por suas conexões de atalho (skip connections) que mitigam o desaparecimento de gradientes em redes profundas [13]. A rede foi inicialmente pré-treinada no conjunto de dados ImageNet, e todos os seus parâmetros foram mantidos como treináveis durante o ajuste para a tarefa específica de segmentação de tumores cerebrais.

D. Função de Loss e Otimizador

Afim de aprimorar a precisão na tarefa de segmentação, foi realizada uma composição de funções, combinamos a *DiceLoss*, que utiliza o coeficiente de *Dice* para medir o grau de sobreposição das distribuições, com a *BinaryCrossEntropyLoss*, que quantifica a diferença entre as distribuições de probabilidade. Um estudo completo sobre funções de loss aplicadas no contexto de segmentação de imagens médicas pode ser encontrada em Ma et al. [14]. Para a tarefa de classificação, foi utilizada a *CrossEntropyLoss*. Juntas, essas funções compõem a loss utilizada no método proposto.

Além disso, o otimizador escolhido foi o *Adaptive Moment Estimation (Adam)* [15], que combina características dos algoritmos *AdaGrad* e *RMSProp*. Este método se mostra eficaz ao ajustar dinamicamente as taxas de aprendizado durante o treinamento da rede neural convolucional, contribuindo significativamente para a otimização do processo.

E. Base de dados

Para a realização deste estudo, utilizamos um conjunto de dados de imagens de ressonância magnética do cérebro [16], que inclui imagens em três planos anatômicos distintos: axial, sagital e coronal. O conjunto de dados é composto por 3064 imagens do tipo *T1ce*, provenientes de 233 pacientes diagnosticados com três tipos de tumores cerebrais: glioma, meningioma e tumor pituitário. Especificamente, o conjunto inclui 1426 amostras de glioma, 708 amostras de meningioma e 930 amostras de tumor pituitário, cada uma acompanhada por sua respectiva máscara de segmentação do tumor. Para cada conjunto de imagem e máscara, um índice de identificação do paciente é anotado, visto que o mesmo paciente pode possuir uma quantidade variável de imagens no dataset.

F. Métricas Estatísticas

Para avaliar o desempenho dos modelos, foram utilizadas duas métricas estatísticas, acurácia balanceada e o coeficiente dice, para as tarefas de classificação e segmentação respectivamente.

A acurácia balanceada é interessante pois considera o desbalanceamento entre as classes. Sua aplicação é semelhante à média macro das revocações por classe. Para uma classe i , calculam-se seus verdadeiros positivos (TP_i) e falsos negativos (FN_i). Com isso, a acurácia balanceada para um problema com n classes é definida como

$$BA = \frac{1}{n} \sum_{j=1}^n \frac{TP_j}{TP_j + FN_j} \quad (1)$$

O coeficiente dice é frequentemente usado em problemas de segmentação. Ele é calculado pela média harmônica entre a precisão e a revocação de uma predição. Em outras palavras, equivale a duas vezes a intersecção entre a máscara de segmentação prevista e a real, dividido pelo número de pixels nas duas imagens. Com isso, como o problema de segmentação é binário, o coeficiente dice é definido como:

$$Dice = \frac{2 \times |X \cap Y|}{|X \cup Y|} \quad (2)$$

TABLE II
COEFICIENTE DE DICE POR FOLD, SEGUIDO DE MÉDIA E DESVIO PADRÃO

Score					Média	Desvio Padrão
90.66%	87.84%	90.81%	88.94%	84.04%	88.82%	2.25%
91.69%	89.67%	85.68%	89.11%	89.77%		

TABLE III
ACURÁCIA BALANCEADA POR FOLD, SEGUIDA DE MÉDIA E DESVIO PADRÃO

Score					Média	Desvio Padrão
100.00%	98.12%	100.00%	99.45%	99.19%	99.38%	0.65%
98.38%	99.05%	100.00%	100.00%	99.61%		

IV. EXPERIMENTOS E RESULTADOS

Com o intuito de treinar a rede para as tarefas de segmentação e classificação e conseguir obter um desempenho médio do modelo, foi feita uma validação cruzada com 10 folds. Ademais, com o intuito de evitar o viés introduzido pela existência de diversas imagens diferentes de um mesmo paciente, cada paciente aparece somente em um fold, ou seja, não haverá intersecção de pacientes entre treino e teste. Levando em consideração que dados de mesma origem entre treino e teste são mais comprometedores que um possível desbalanceamento de classes, tal configuração foi optada, o que deixa a proporção entre classes levemente desbalanceada.

Para assegurar-se que um superajuste não será configurado, os dados disponibilizados para treino são separados entre validação e treino. Tal segregação também leva em conta os índices dos pacientes, evitando o viés. A porcentagem de dados de validação é definida como aproximadamente 25% dos dados de treino.

Além disso, os experimentos foram realizados utilizando um *batch size* de 32, ao longo de 30 épocas, com uma taxa de aprendizagem de 0.001.

Na tabela II e na tabela III, apresentamos os resultados agregados dos 10 folds. Para a segmentação, obtivemos uma média

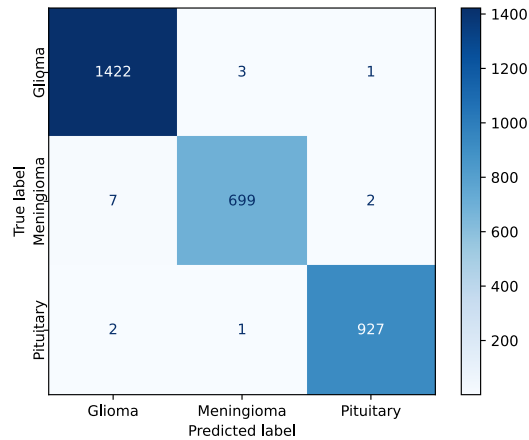


Fig. 2. Matriz de confusão que condensa o resultados dos 10 folds de validação cruzada e detalha a distribuição das predições para cada classe de tumor.

de $88.82\% \pm 2.25\%$ para o coeficiente de *Dice*, indicando uma boa precisão na delimitação dos contornos tumorais. Em termos de classificação, a acurácia média foi de $99.38\% \pm 0.65\%$, demonstrando a capacidade do modelo em distinguir entre diferentes classes de tumores cerebrais com alta precisão. Além disso, a matriz de confusão na Figura 2 condensa o resultados dos 10 folds de validação cruzada e detalha a distribuição das predições para cada classe de tumor.

V. CONCLUSÕES

Este trabalho, como um todo, apresentou uma abordagem para o problema de segmentação e classificação de tumores cerebrais com dados de imagens de ressonância magnética. Em geral, a rede apresentada localizou com considerável precisão o local dos tumores, assim como inferiu de forma excelente as classificações destes.

Durante o trabalho, escolhas foram feitas para auxiliar na prevenção de vies, como a separação do dataset, a qual foi feita levando em conta a existência de múltiplas imagens de um só paciente. Dessa forma, a busca de dados se torna mais complexa ao ponto que diferentes datasets tendem a ter seus objetivos mais distantes do que o proposto pelo trabalho, o qual evitou ao máximo a contaminação de dados. Assim, em trabalhos futuros, a adição de imagens com exemplos de ausência de tumor poderia ser relevante para determinar um desempenho realístico, considerando um ambiente real.

REFERENCES

- [1] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on mri," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, 2019. Special Issue: Deep Learning in Medical Physics.
- [2] N. Yamanakkanavar, J. Y. Choi, and B. Lee, "Mri segmentation and classification of human brain using deep learning for diagnosis of alzheimer's disease: A survey," *Sensors*, vol. 20, no. 11, 2020.

- [3] N. Ghassemi, A. Shoeibi, and M. Rouhani, "Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images," *Biomedical Signal Processing and Control*, vol. 57, p. 101678, 2020.
- [4] E. ırmak, "Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 45, pp. 1–22, 04 2021.
- [5] M. Aamir, Z. Rahman, Z. A. Dayo, W. A. Abro, M. I. Uddin, I. Khan, A. S. Imran, Z. Ali, M. Ishfaq, Y. Guan, and Z. Hu, "A deep learning approach for brain tumor classification using mri images," *Computers and Electrical Engineering*, vol. 101, p. 108105, 2022.
- [6] M. A. Mafraji, "Ressonância magnética."
- [7] D. R. J. Vieira, "Ressonância magnética (rm): Princípios básicos."
- [8] D. A. Ibrahim MA, Hazhirkarzar B, "Gadolinium magnetic resonance imaging."
- [9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.
- [10] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information*, vol. 11, no. 2, 2020.
- [11] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimselshin, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," 2019.
- [12] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015.
- [14] J. Ma, J. Chen, M. Ng, R. Huang, Y. Li, C. Li, X. Yang, and A. L. Martel, "Loss odyssey in medical image segmentation," *Medical Image Analysis*, vol. 71, p. 102035, 2021.
- [15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.
- [16] J. Cheng, "Brain Tumor Dataset," 4 2017.