

SCC0270 - Redes Neurais e Aprendizado Profundo

Detecção e Análise de Tumores Cerebrais (CT + MRI)

Bruno Neves Boa Sorte
Caio Uramoto Evangelista
Matheus Rodrigues Dos Santos
Murillo Domingos de Almeida
Thiago Aragão da Silva

¹Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo (USP)

Resumo. *O diagnóstico preciso de tumores cerebrais depende de múltiplas fontes de informação. A Ressonância Magnética (MRI) oferece detalhes sobre tecidos moles, enquanto a Tomografia Computadorizada (CT) destaca estruturas ósseas. Este projeto visa desenvolver um modelo de aprendizado profundo que explora a sinergia entre imagens de MRI e CT para a detecção e segmentação de tumores, buscando uma precisão diagnóstica superior à de abordagens unimodais. O trabalho segue a proposta da disciplina de explorar problemas multimodais, estabelecendo inicialmente modelos de base (baseline) unimodais (apenas MRI) para, em seguida, compará-los com um modelo de fusão multimodal.*

1. Metodologia

1.1. Conjunto de Dados e Análise

Utilizamos o dataset público Brain Tumor Multimodal Image CT and MRI do Kaggle, que contém imagens de ambas as modalidades, categorizadas como Healthy ou Tumor. A análise exploratória inicial no conjunto de MRI revelou um desbalanceamento de classes, com 2.495 imagens de Tumor e 1.429 de Healthy, um fator a ser considerado no treinamento.

1.2. Pré-processamento

Uma pipeline de pré-processamento padronizada foi aplicada a ambas as modalidades (MRI e CT): Divisão dos Dados: 80% para treinamento e 20% para teste, garantindo o pareamento das amostras de pacientes na abordagem multimodal. Transformações: Todas as imagens foram redimensionadas para 224x224 pixels e convertidas para tensores do PyTorch para compatibilidade com arquiteturas pré-treinadas.

1.3. Arquitetura dos Modelos

Modelo Unimodal (Baseline): Foram implementadas duas Redes Neurais Convolucionais (CNNs), ResNet e EfficientNet, utilizando aprendizado por transferência com pesos pré-treinados na ImageNet. A camada de classificação final de ambas foi adaptada para o nosso problema binário (Healthy/Tumor). Estes modelos foram treinados exclusivamente com imagens de MRI para a tarefa de classificação.

Modelo Multimodal: A arquitetura planejada visava um duplo objetivo: classificação e segmentação. Para isso, adotávamos uma abordagem de fusão tardia com

dois fluxos paralelos (ResNet para MRI e ResNet para CT). Os vetores de características de alto nível seriam concatenados. Para a tarefa de **classificação**, esse vetor alimentaria uma camada MLP. Para a tarefa de **segmentação**, as características fundidas seriam passadas para um "decodificador" inspirado na arquitetura U-Net, permitindo a reconstrução de uma máscara de segmentação precisa do tumor.

1.4. Métricas de Avaliação

O desempenho dos modelos de classificação foi medido utilizando Acurácia, Precisão, Recall, F1-Score e a Matriz de Confusão. Para a tarefa de segmentação, as métricas planejadas eram o Coeficiente de Dice e a IoU (Intersection over Union).

2. Resultados e Discussões

2.1. Desempenho dos Modelos Unimodais (Classificação)

Os modelos unimodais de classificação, treinados apenas com imagens de MRI, serviram como nossa referência de desempenho. Os resultados obtidos com as arquiteturas ResNet e EfficientNet demonstram a eficácia da análise de uma única modalidade, alcançando métricas de performance extremamente altas, como pode ser visto abaixo.

— Relatório de Classificação Final para ResNetModel —				
	precision	recall	f1-score	support
Healthy	0.99	1.00	1.00	365
Tumor	1.00	1.00	1.00	635
accuracy			1.00	1000
macro avg	1.00	1.00	1.00	1000
weighted avg	1.00	1.00	1.00	1000

— Relatório de Classificação Final para EfficientNetModel —				
	precision	recall	f1-score	support
Healthy	1.00	1.00	1.00	365
Tumor	1.00	1.00	1.00	635
accuracy			1.00	1000
macro avg	1.00	1.00	1.00	1000
weighted avg	1.00	1.00	1.00	1000

2.2. Limitações e Resultados Esperados do Modelo Multimodal

Durante o desenvolvimento do projeto, encontramos limitações significativas de hardware que nos impediram de treinar os modelos multimodais avançados, tanto para a tarefa de classificação quanto para a de segmentação. A arquitetura, que processa simultaneamente múltiplos fluxos de dados e, no caso da segmentação, reconstrói uma saída com as dimensões da imagem, demanda uma capacidade de cômputo (memória e processamento de GPU) superior à que tínhamos disponível.

No entanto, a hipótese central do projeto permanece válida. Para a **classificação**, esperávamos que o modelo multimodal apresentasse um desempenho ainda mais robusto,

reduzindo casos de falsos positivos e negativos. Para a **segmentação**, a expectativa era ainda maior. Acreditamos que a fusão das informações permitiria uma delimitação muito mais precisa dos contornos do tumor. A MRI, com seu excelente contraste em tecidos moles, definiria o corpo principal do tumor, enquanto a CT ajudaria a diferenciar o tumor de estruturas adjacentes como ossos ou áreas de calcificação. Isso resultaria em um aumento significativo das métricas de sobreposição, como o Coeficiente de Dice e a IoU (Intersection over Union), e geraria máscaras de segmentação mais confiáveis para o planejamento cirúrgico.

3. Conclusão

Este projeto estabeleceu com sucesso uma metodologia para a análise de tumores cerebrais, implementando e avaliando modelos unimodais de classificação de alto desempenho baseados em imagens de MRI. As arquiteturas ResNet e EfficientNet se mostraram extremamente eficazes, criando uma base sólida para o problema.

Apesar de a fase de experimentação multimodal não ter sido executada por limitações de hardware, o estudo teórico e a arquitetura planejada para as tarefas de classificação e segmentação reforçam a validade da hipótese de que a fusão de dados de CT e MRI pode levar a um sistema de IA mais robusto e preciso. O trabalho realizado serve como um ponto de partida valioso, e a implementação do modelo multimodal para ambas as tarefas permanece como um passo futuro claro e promissor, justificando a exploração de arquiteturas como a U-Net para a segmentação. O sucesso dos modelos de base demonstra o grande potencial clínico e técnico das abordagens de aprendizado profundo, e a exploração multimodal é a próxima fronteira para avançar na precisão diagnóstica.

4. Referencias

- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. CVPR.
- Likhon, Murtoza. (2024). Brain Tumor Multimodal Image CT and MRI. Kaggle.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI.