

Introdução

Este projeto visa desenvolver um sistema de segmentação automática de tumores cerebrais em imagens de ressonância magnética (MRI) utilizando a arquitetura U-Net, uma rede neural convolucional (CNN) projetada especificamente para tarefas de segmentação biomédica. A segmentação precisa de tumores cerebrais é fundamental para diagnóstico clínico, planeamento cirúrgico, monitorização do tratamento e estudos prognósticos. A automação deste processo através de técnicas de deep learning promete aumentar a eficiência, solidez e reprodutibilidade em comparação com métodos manuais. O presente trabalho implementa e valida uma solução completa baseada em U-Net, abordando desde o pré-processamento de dados até a avaliação quantitativa do desempenho.

O código fonte está disponibilizado no seguinte repositório público: [sdas103/atividade-5-pdi](https://github.com/sdas103/atividade-5-pdi)

Tema

Segmentação semântica de tumores cerebrais em imagens de ressonância magnética multimodais utilizando uma arquitetura U-Net 2D otimizada para eficiência computacional.

Objetivos

Implementar uma U-Net para a segmentação de tumores cerebrais, desenvolver um pipeline completo de pré-processamento e aumento de dados para imagens médicas, treinar o modelo utilizando dados do desafio BraTS 2020, validar o desempenho com métricas médicas relevantes.

Metodologia

Dataset e Pré-processamento

O dataset utilizado é o **BraTS 2020** (Brain Tumor Segmentation), contendo 369 casos de pacientes com gliomas. Cada caso inclui quatro sequências de MRI: T1, T1ce (T1 com contraste), T2 e FLAIR, com resolução espacial de $240 \times 240 \times 155$ voxels. As anotações manuais por especialistas definem quatro classes: 0 (background), 1 (necrose tumoral), 2 (edema peritumoral) e 4 (tecido tumoral realçado). Entretanto por limitações de memória para este projeto foram utilizadas apenas 5 pacientes para treinamento e 5 pacientes para validação.

O pipeline de pré-processamento implementado inclui carregamento otimizado com memory mapping para processar grandes volumes 3D (≈ 140 MB por paciente) em sistemas com memória limitada, extração de fatias 2D contendo tecido tumoral, data augmentation durante

o treinamento com rotação aleatória ($\pm 15^\circ$), flip horizontal/vertical e ajuste de intensidade, e separação rigorosa entre conjuntos de treino (5 pacientes) e validação (5 pacientes distintos).

Treinamento e Otimização

O modelo foi treinado com os seguintes parâmetros: batch size de 2 (balanceando memória e estabilidade), 15 épocas com early stopping (paciência: 10 épocas), learning rate inicial de 0.0005 com redução para 0.00025 na época 10, função de perda Focal Loss com pesos por classe [0.006, 2.288, 1.411, 0.295] otimizada para desbalanceamento severo, e otimizador Adam ($\beta_1=0.9$, $\beta_2=0.999$).

Foram implementadas estratégias de otimização de memória incluindo carregamento preguiçoso (lazy loading) de fatias individuais, memory mapping para arquivos grandes, cache LRU limitado a 2GB com limpeza periódica, e garbage collection explícito entre batches.

Desempenho Quantitativo

O modelo alcançou os seguintes resultados após 15 épocas de treinamento (com early stopping na época 12): Melhor Dice Score de 0.1252 (época 2), Dice Score Final de 0.0868 ± 0.0927 , Loss de Treino Final de 0.0020 (redução de 83.6% desde época 1), Loss de Validação Final de 0.0065, Tempo Total de Treino de 83.8 minutos (CPU), Uso Máximo de Memória de 1.16 GB, e processamento de 309 slices de treino + 321 de validação = 630 slices no total.

Análise por Classe Tumoral:

A análise por classe revelou os seguintes resultados: Necrose tumoral (classe 1) com Dice = 0.0529 ± 0.0903 , Tumor realçado (classe 3) com Dice = 0.1353 ± 0.1783 , Edema peritumoral (classe 2) com Dice = 0.0723 ± 0.0480 , e Background (classe 0) com segmentação consistente (97.6% dos voxels).

Interpretação dos Resultados

Aspectos Positivos: O modelo demonstrou capacidade de aprendizado detectável, com melhoria inicial do Dice Score de 0.0911 para 0.1252 (+37.4%). A utilização de pacientes completamente diferentes para treino (BraTS001-005) e validação (BraTS006-010) preveniu data leakage e garantiu generalização real. A gestão eficiente de memória permitiu processar 10 pacientes (≈ 1.4 GB de dados) em CPU com apenas 1.16GB de uso máximo. O early stopping funcionou corretamente com detecção automática de overfitting após 10 épocas sem

melhoria, e a infraestrutura reproduzível com configuração serializada em JSON e resultados estruturados em múltiplos formatos.

Limitações Identificadas: O desempenho quantitativo foi limitado com Dice score máximo de 0.1252, abaixo do ideal para aplicação clínica (>0.70). O desbalanceamento extremo de classes, com background representando 97.6% dos voxels, criou forte viés preditivo. Observou-se overfitting precoce com melhor performance na época 2 seguida de plateau/declínio. A alta variabilidade (desvio padrão de ± 0.0927) indica inconsistência entre diferentes fatias/pacientes, e o dataset limitado com apenas 10 pacientes (5+5) mostrou-se insuficiente para generalização robusta.

Análise da Curva de Aprendizado: Nas épocas 1-2 houve aprendizado rápido com melhoria significativa (Dice: 0.0911 \rightarrow 0.1252). Nas épocas 3-12 observou-se plateau com oscilações (Dice: 0.08-0.09), indicando overfitting. Na época 10 ocorreu redução automática do learning rate (0.0005 \rightarrow 0.00025), e na época 12 o early stopping foi acionado após 10 épocas sem melhoria.

Comparação com a Literatura

Os resultados obtidos (Dice máximo 0.1252) ficam abaixo dos valores típicos reportados na literatura para o desafio BraTS, onde modelos state-of-the-art frequentemente alcançam Dice scores >0.85 . Esta discrepância pode ser atribuída a múltiplos fatores: escala do dataset reduzida (10 pacientes vs. centenas na literatura), arquitetura simplificada (U-Net 2D vs. arquiteturas mais complexas como nnU-Net, U-Net 3D ou Transformers), desbalanceamento extremo não totalmente otimizado (distribuição 97.6% background vs. 2.4% tumor), e recursos computacionais limitados (treinamento em CPU vs. GPUs especializadas com batch sizes maiores).

No entanto, comparado com o baseline de sempre prever background (Dice = 0.0), o modelo demonstrou aprendizado real (0.1252 representa melhoria infinita). A comparação com previsão aleatória uniforme (Dice esperado = 0.25 para 4 classes) mostra que há espaço significativo para melhoria. Apesar das limitações quantitativas, o projeto demonstrou viabilidade conceitual e estabeleceu infraestrutura metodológica sólida para desenvolvimento futuro, com validação rigorosa em pacientes não vistos durante o treinamento.

Conclusões

Este trabalho demonstrou a aplicação prática de redes neurais profundas para segmentação automática de tumores cerebrais, implementando um pipeline completo desde o pré-processamento de dados até a avaliação quantitativa. Apesar do desempenho quantitativo limitado (Dice máximo 0.1252), o projeto alcançou vários objetivos importantes: implementação técnica bem-sucedida da arquitetura U-Net com otimizações específicas para imagens médicas, desenvolvimento de pipeline robusto incluindo pré-processamento, aumento de dados e avaliação, e metodologia rigorosa com separação estrita entre conjuntos de treino e validação usando pacientes diferentes.

O projeto estabelece uma base para pesquisa futura em segmentação automática de tumores cerebrais, destacando tanto o potencial quanto os desafios práticos da aplicação de deep learning em medicina, particularmente a importância de dados adequados e metodologia rigorosa de validação. Perspectivas futuras incluem a expansão do dataset, experimentação com arquiteturas mais avançadas, implementação de técnicas especializadas para lidar com desbalanceamento extremo, e validação clínica mais abrangente.

Referencias

Ronneberger, O., Fischer, P., Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Medical Image Computing and Computer-Assisted Intervention.

Menze, B. H., et al. (2014). *The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)*. IEEE Transactions on Medical Imaging.