AttUNet: Enhancing ResUNet with Attention Mechanisms for Accurate Brain Tumor Segmentation

Abstract

In medical image analysis, brain tumor segmentation is a crucial problem with significant ramifications for diagnosis, treatment planning, and outcome prediction. Due of its time-consuming nature and susceptibility to inter-observer variability, manual segmentation need reliable automated methods that are accurate and broadly applicable. In order to improve the model's ability to precisely locate and segment tumor locations inside magnetic resonance imaging (MRI) scans, we present a unique hybrid deep learning architecture called AttUNet. This design incorporates attention mechanisms into a Residual U-Net backbone. We employ the publically accessible TCGA-LGG dataset, a benchmark dataset of 110 patient cases selected from The Cancer Imaging Archive (TCIA), to specifically address the segmentation of Lower Grade Glioma (LGG).

The basic architecture employs a bottleneck with an 8 dilation rate and 1024 feature maps with 16x16 residual blocks. This configuration enables the model to learn multiscale contextual signals, which is crucial for illustrating complex tumor borders. Transpose convolution layers carry the encoded properties to the decoder, where they serve as gating signals in the attention mechanism. This dual-role helps the model bridge the semantic gap between the encoder and decoder, ensuring that crucial tumor features are maintained throughout upsampling. ReLU activation functions are employed across the network to add non-linearity and mitigate the vanishing gradient problem in order to efficiently train hierarchical representations.

Batch normalization is used in the quest for better model generalization and convergence. The feature distributions within mini-batches are standardized through the use of calculated means and variances for each batch dimension. In addition to speeding up convergence, this normalization guarantees training stability. In our framework, the loss function is still another important innovation. We use the Tversky Loss and its more sophisticated version, the Focal Tversky Loss, which are both very good at controlling class imbalance, which is a prevalent issue in medical imaging. We bias the model to penalize false negatives more severely by setting $\alpha = 0.7$ in Tversky Loss and $\gamma = 0.75$ in Focal Tversky Loss. This increases the sensitivity to small tumor areas.

We assess our model using three different segmentation metrics: Similarity Index (SI), Intersection over Union (IoU), and Dice Similarity Coefficient (DSC). These measures offer a thorough examination of the spatial consistency and overlap accuracy between the expert-labeled ground truth and the anticipated segmentation maps. In benchmark comparisons, AttUNet outperforms well-known models including 3D U-Net, Deep Res-UNet, Znet Deep Learning, and UNet++, achieving impressive results with a DSC of 0.94, an IoU of 0.92, and a SI of 0.94.

The experiments were conducted using five-fold cross-validation on patient clusters. The accuracy and recall scores further confirm the robustness of the model, demonstrating that the F1 score for tumor regions was 0.91 (precision = 0.95, recall = 0.85) and the F1 score for non-tumor regions was 0.95 (precision = 0.92, recall = 0.96). The high recall for non-tumor regions reduces false positives, ensuring clinical safety, while the excellent precision in tumor core detection encourages a trustworthy diagnosis.

Beyond numerical assessment, AttUNet can delineate tumor boundaries finely, even when there is morphological diversity, according to visual examination of segmentation masks. The model's ability to accurately represent both the core and peritumoral areas is validated by figure-based analysis, indicating that it has the potential to be implemented in clinical decision support systems. Additionally, the imaging pipeline's incorporation of genetic data, such as DNA methylation patterns and IDH mutation status, opens the door for future multimodal learning strategies that could improve tailored therapy.

In conclusion, AttUNet represents a significant advancement in medical image segmentation by integrating attention-aware residual networks with specialized loss functions and robust validation metrics. The results demonstrate the model's superiority in segmenting complex brain tumor structures from MRI scans, both quantitatively and qualitatively. This research not only lays the groundwork for clinical translation but also aligns with broader AI-driven healthcare initiatives, thus embodying the synthesis of academic excellence and societal impact.

From an academic perspective, this research aligns with cutting-edge topics in Artificial Intelligence, particularly deep learning for medical imaging, attention mechanisms, and imbalanced data learning. The work is grounded in a rigorous theoretical framework while addressing real-world healthcare challenges, making it highly relevant to the vision and research interests of the University of Passau's MSc in Artificial Intelligence. My motivation to pursue this program stems from its interdisciplinary curriculum that bridges AI theory and application, as well as its emphasis on ethical AI and health informatics, both of which are central to my long-term goals of contributing to intelligent diagnostic systems in medicine.

I look forward to building upon this foundation at the University of Passau through collaborative research, advanced coursework, and interdisciplinary innovation