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

Abstract

Brain tumor segmentation is a pivotal challenge in medical image analysis, with substantial implications for diagnosis, treatment planning, and outcome prediction. Manual segmentation, being time-consuming and prone to inter-observer variability, demands robust automated approaches that are both precise and generalizable. In this study, we introduce a novel hybrid deep learning architecture, **AttUNet**, which integrates attention mechanisms into a Residual U-Net backbone, thereby enhancing the model's capability to localize and segment tumor regions accurately within magnetic resonance imaging (MRI) scans. Our approach is particularly focused on addressing the segmentation of **Lower Grade Glioma (LGG)** using the publicly available TCGA-LGG dataset, a benchmark dataset consisting of 110 patient cases curated from The Cancer Imaging Archive (TCIA).

The core architecture leverages a bottleneck with 1024 feature maps and residual blocks of 16×16, incorporating a dilation rate of 8. This configuration enables the model to learn multiscale contextual features, crucial for delineating complex tumor boundaries. The encoded features are propagated to the decoder through transpose convolution layers, while also serving as gating signals in the attention mechanism. This dual-role helps the model to bridge the semantic gap between the encoder and decoder efficiently, ensuring that salient tumor features are retained during upsampling. ReLU activation functions are applied across the network to introduce non-linearity and mitigate the vanishing gradient problem, enabling effective learning of hierarchical representations.

In the pursuit of improved model convergence and generalization, batch normalization is employed. It standardizes feature distributions within mini-batches, using computed means and variances per batch dimension. This normalization ensures training stability and accelerates convergence. The loss function is another significant innovation in our framework. We adopt the **Tversky Loss** and its advanced variant, the **Focal Tversky Loss**, both of which are highly effective in managing class imbalance—an inherent problem in medical imaging where the background class often dominates. By setting $\alpha = 0.7$ in Tversky Loss and $\gamma = 0.75$ in Focal Tversky Loss, we bias the model to penalize false negatives more heavily, thereby improving sensitivity to small tumor regions.

We evaluate our model using a threefold set of segmentation metrics—**Dice Similarity Coefficient (DSC)**, **Intersection over Union (IoU)**, and **Similarity Index (SI)**. These metrics provide a comprehensive analysis of both overlap accuracy and spatial consistency between the predicted segmentation maps and the expert-labeled ground truth. AttUNet achieves remarkable results, attaining a DSC of **0.94**, an IoU of **0.92**, and an SI of **0.94**, outperforming established models including 3D U-Net, Deep Res-UNet, Znet Deep Learning, and UNet++ in benchmark comparisons.

Experimentation was performed using five-fold cross-validation on patient clusters. Precision and recall scores further affirm the model's robustness: non-tumor regions achieved an **F1 score of 0.95** (precision = 0.92, recall = 0.96), while tumor regions attained an **F1 score of**

0.91 (precision = 0.95, recall = 0.85). The high recall for non-tumor areas reduces false positives, ensuring clinical safety, whereas high precision in tumor core identification supports confident diagnosis.

Beyond numerical evaluation, visual analysis of segmentation masks reveals that AttUNet is capable of fine-grained delineation of tumor boundaries, even in cases of morphological variability. Figure-based analyses demonstrate that the model successfully captures both the core and peritumoral regions, thus validating its potential for deployment in clinical decision support systems. Furthermore, the integration of genomic data such as IDH mutation status and DNA methylation patterns within the imaging pipeline paves the way for future multimodal learning approaches that could enhance personalized medicine.

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.

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.

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