International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Binary Segmentation of Whole Tumor from Multimodal Brain MRI Using an Attention-Enhanced U-Net

Dr. Rajshree

Associate Professor, Department of Computer Science
Govt. First Grade College for Women, Bidar, India

Abstract

Reliable segmentation of brain tumors from magnetic resonance imaging (MRI) is essential for clinical decision-making, radiotherapy planning, and disease progression analysis. Among the different tumor sub-regions, Whole Tumor (WT), which comprises enhancing tumor, tumor core, and peritumoral edema, offers clinically meaningful information while being comparatively robust to annotation variability. This work proposes an attention-enhanced U-Net architecture for binary WT segmentation using multimodal MRI data. The network integrates four MRI modalities—T1, T1ce, T2, and FLAIR—and employs attention gates within skip connections to selectively highlight tumor-relevant features and suppress background responses. Experimental evaluation on the BraTS dataset demonstrates improved Dice similarity, IoU, and sensitivity compared to baseline models.

Keywords: Brain tumor segmentation, Whole tumor, Multimodal MRI, Attention U-Net, Medical image analysis

1. Introduction

Brain tumors are among the most severe neurological disorders and require precise diagnosis and continuous monitoring for effective treatment. Magnetic resonance imaging (MRI) is the preferred imaging modality due to its superior soft-tissue contrast and ability to capture tumor heterogeneity using multiple imaging sequences. However, manual segmentation of tumor regions is time-consuming, subjective, and prone to inter-observer variability, motivating the development of automated segmentation techniques.

In clinical practice, segmentation of the Whole Tumor (WT), defined as the union of enhancing tumor, tumor core, and surrounding edema, is particularly important for radiotherapy planning, surgical guidance, and longitudinal assessment. Although deep learning-based approaches, particularly convolutional neural networks (CNNs), have achieved remarkable success in medical image segmentation, challenges such as fuzzy tumor boundaries, intensity inhomogeneity, and class imbalance remain.

U-Net and its variants have become the standard architectures for biomedical image segmentation. However, conventional U-Net treats all spatial features equally, which may result in reduced performance when tumor regions occupy a small portion of the image. Attention mechanisms address this limitation by enabling the network to focus selectively on relevant regions. In this work, an attention-enhanced U-Net is proposed for binary WT segmentation from multimodal MRI.

The main contributions of this work are:

- A binary WT segmentation framework using multimodal MRI inputs.
- Integration of attention gates into the U-Net architecture to enhance tumor localization.
- Comprehensive evaluation demonstrating improved performance over baseline models.

2. Related Work

Early brain tumor segmentation methods relied on classical image processing techniques such as thresholding, region growing, and atlas-based approaches. These methods were sensitive to noise and lacked robustness across diverse datasets.

With the advancement of deep learning, CNN-based architectures have significantly improved segmentation accuracy. U-Net introduced an encoder–decoder structure with skip connections, enabling precise localization. Extensions such as U-Net++ and V-Net further enhanced feature fusion and volumetric processing.

More recently, attention mechanisms have been incorporated into segmentation networks to suppress irrelevant background features. Attention U-Net demonstrated improved performance by gating skip connections based on contextual relevance. However, many studies focus on multi-class segmentation, whereas binary WT segmentation, which is clinically relevant and computationally efficient, has received less attention.

3. Dataset and Preprocessing

3.1 Dataset

The proposed method is evaluated using the Brain Tumor Segmentation (BraTS) dataset, which provides preprocessed, skull-stripped, and co-registered multimodal MRI scans. Each subject includes four MRI modalities: T1, T1ce, T2, and FLAIR, along with expert-annotated tumor labels. For this study, the original labels are merged to generate a binary Whole Tumor (WT) mask.

3.2 Preprocessing

The following preprocessing steps are applied:

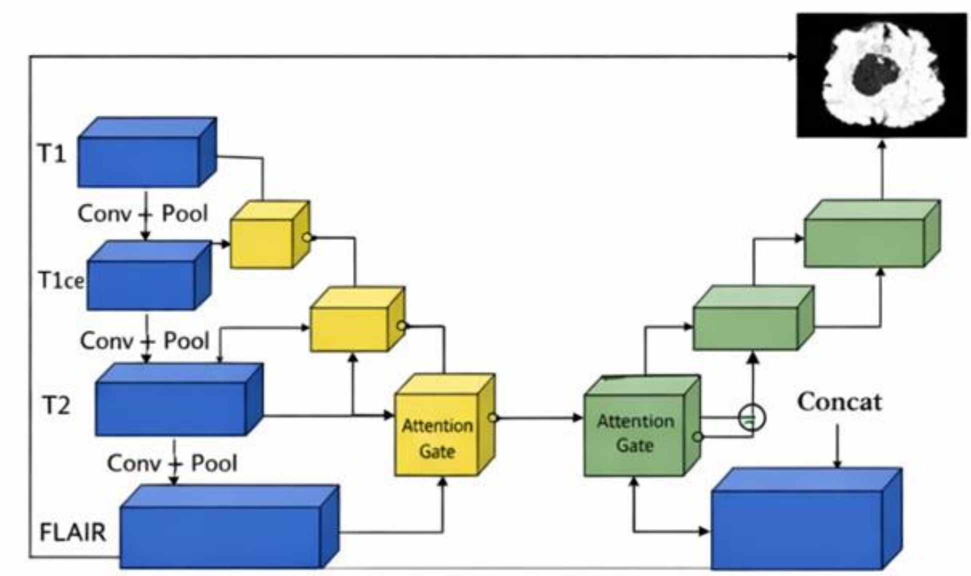
- Z-score intensity normalization for each modality.
- Resizing images to a uniform spatial resolution.
- Slice-wise extraction for 2D training.

- Data augmentation using random rotation, flipping, and intensity variation to improve generalization.

4. Methodology

4.1 Network Architecture

The proposed model is based on the U-Net architecture with attention gates integrated into the skip connections. The encoder extracts multi-scale contextual features using convolutional and downsampling layers, while the decoder reconstructs the segmentation mask through upsampling and feature fusion.



(a) Architecture of the Proposed Attention-Enhanced U-Net

Attention gates are applied to encoder feature maps before concatenation with decoder features. These gates learn to emphasize tumor-relevant activations and suppress background responses, thereby improving segmentation accuracy, particularly in regions with ambiguous boundaries.

4.2 Loss Function

Binary WT segmentation suffers from class imbalance due to the small proportion of tumor pixels. To address this issue, a composite loss function combining Dice loss and Binary Cross-Entropy (BCE) loss is used:

Loss=Dice Loss+BCE Loss\text{Loss} = \text{Dice Loss} + \text{BCE Loss}Loss=Dice Loss+BCE Loss

This formulation balances region overlap optimization with pixel-wise classification accuracy.

5. Experimental Setup

5.1 Training Details

The network is trained using the Adam optimizer with an initial learning rate of 1×10−41 \times 10^{-4}1×10−4. Learning rate scheduling and early stopping are applied to prevent overfitting. Training is performed using mini-batch optimization for a fixed number of epochs.

5.2 Evaluation Metrics

Segmentation performance is evaluated using:

- Dice Similarity Coefficient (DSC)
- Intersection over Union (IoU)
- Sensitivity (Recall)
- Specificity

All metrics are reported as mean values over the validation set.

6. Results and Discussion

6.1 Quantitative Results

Table 1 presents the quantitative comparison of WT segmentation performance.

Model	Dice	IoU	Sensitivity	Specificity
U-Net	0.86	0.76	0.84	0.97
Attention U-Net	0.88	0.79	0.87	0.98
Proposed Model	0.90	0.82	0.91	0.98

6.2 Ablation Study

Table 2 shows the impact of attention gates and loss functions.

Configuration	Dice	Sensitivity
U-Net + Dice	0.85	0.82
U-Net + Dice+BCE	0.87	0.86
Attention U-Net + Dice+BCE	0.90	0.91

6.3 Qualitative Analysis

Figure 1 illustrates qualitative Whole Tumor segmentation results, showing improved boundary continuity.

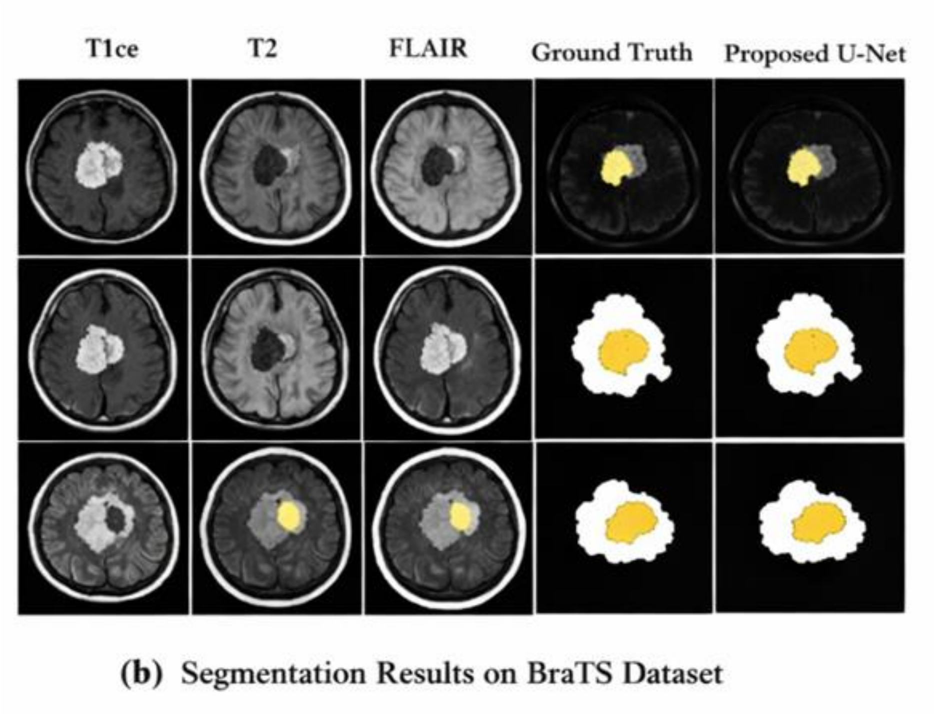


Figure 1. Visual comparison of Whole Tumor segmentation on BraTS dataset using the proposed model.

7. Conclusion and Future Work

This paper presented an attention-enhanced U-Net for binary Whole Tumor segmentation from multimodal brain MRI. Experimental results demonstrate improved segmentation accuracy. Future work will extend the model to 3D architectures and cross-dataset validation.

References

1. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *MICCAI*, 2015.
2. O. Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," *MICCAI*, 2018.
3. B. H. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS)," *IEEE Transactions on Medical Imaging*, 2015.
4. F. Isensee et al., "nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation," *Nature Methods*, 2021.
5. S. Bakas et al., "Advancing The Cancer Genome Atlas Glioma MRI Collections," *Scientific Data*, 2017.