

Towards Precise Brain Tumor Segmentation in MRI with Boundary Delineation

*Sang Suh and Ozlem Deveci**

Dept of Computer Science, East Texas A&M University, U.S.A.

Abstract. Brain tumor segmentation in MRI scans is crucial for accurate diagnosis, treatment planning, and monitoring tumor progression. Manual segmentation is time-consuming and prone to errors, while traditional deep learning models often struggle with small or low-contrast tumor regions and precise boundary delineation [27]. In this research, we propose an enhanced brain tumor segmentation approach using a hybrid TransUNet architecture combined with a boundary-aware loss function adopted from [28]. The model leverages the convolutional encoder of U-Net for local features and a Transformer module for capturing global context. The boundary-aware loss, which combines Dice, Cross-Entropy, and edge-focused terms, improves segmentation accuracy, especially for the clinically important enhancing tumor regions.

Keywords; brain tumor segmentation, enhancing tumor segmentation, boundary delineation, TransUNet architecture

Cite this paper as : Sang Suh and Ozlem Deveci (2026) "Towards Precise Brain Tumor Segmentation in MRI with Boundary Delineation", Journal of Industrial Information Technology and Application, Vol. 10. No. 3, pp. 1388-1405

* Corresponding author : sang.suh@tamuc.edu

Received: Jul. 03. 2026 Accepted: Aug. 15. 2026 Published: Sept. 30. 2026

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright©2017. Journal of Industrial Information Technology and Application (JIITA)

1. Introduction

Brain tumor conditions are critical in the field of medicine. Evaluating the treatment and management of such conditions needs the right diagnosis and in a timely manner. Tumors are capable of damaging vital functions of the brain, and the consequences of such malignancies might lead to death. Brain metastatic tumors are responsible for about 250,000 deaths every year[1]. There is a large contribution of malignant gliomas due to their fast growth and unfavorable prognosis[2]. Patients suffering from gliomas have a median of 15 months survival, and a small rate of 5 % survives 5 years after the diagnosis[3]. This incidence reveals the critical necessity of diagnosing and treating such cases in a timely manner to enhance the rate of survival.

MRI is the gold standard for diagnosing and tracking the growth of brain tumors in patients. Moreover, it is noninvasive and safe for a patient because it does not involve dangerous radiation [4]. There are different types of MRIs that are more modern in the sense that they utilize multiple sequences including T1-weighted, contrast-enhanced T1 (T1ce), T2-weighted and are for FLAIR. These different sequences are capable of providing different types of tissues and it helps to delineate regions of a tumor with contrast such as the enhancing tumor [5]. These combinations of scans allow a summation of the review and findings of the tumor for the edematous areas of the brain, and that is a great advantage for real diagnosis and treatment.

In spite of improvements in MRI technology, tumor analysis still takes way too long. In addition to segmentation errors, aspects of volumetric MRI scans, such as tumor analyses, are incredibly nuanced requiring expert training, so they can take several hours each. Aspects of volumetric MRI scans, such as tumor analyses, are incredibly nuanced requiring expert training, so they can take several hours each. Furthermore, differences within peer observation create additional issues. It is common for different neuroradiologists to produce segmentation for the same tumor. Inter-observer differences can create problems. Neuroradiologists are not immune to errors [6]. There are documented instances where the error rates for manual segmentation reached 20%. As such, there is a demand for dependable, automated segmentation solutions that help clinicians accurately identify the regions of the tumor with great efficiency. To reduce the overall segmentation automation, segmentation, standardization, and consistency will be provided for different patients and different institutions.

Among the classic techniques used for the segmentation of tumors, support vector machines (SVM), random forests, and clustering have had their share of the spotlight. Average performance is achieved using these techniques due to their reliance on hand-crafted features from the MRI data such as texture, intensity, and form. They also capture moderately complex spatial and contextual information. Another source of under-segmentation is the class imbalance issue of the small tumors which constitutes a small fraction of the entire image and thus tend to get overshadowed and ignored. [7]. Deep learning has allowed for the automated abstraction of features and has radically changed the way in which medical images are analyzed. For the automated segmentation of brain tumors, the most commonly used architectures are the convolutional neural networks (3D) and the U-Net, as they are optimized for the processing of volumetric data. The most prevalent 3D U-Net configurations are the encoder-decoder architectures with skip-connections, as they perform adequately in maintaining high-resolution fragmented data while also preserving the most contextual layers. [8]. The skip connections are especially important in cases which require segmentation of small tumors which are easily overshadowed due to their low contrast.

This research focuses on the segmentation of brain tumors in MRIs, especially the segmenting of the enhancing tumor (ET) and its borders. Enhancing tumors are small, show low contrast, and have diverse appearances from patient to patient, which makes reliable segmentations difficult. There are many segmentation models [9] showing smooth segmentations, but failing to show the sharp and crucial borders for the segmentation to be useful in planning treatment and the monitoring of the tumor in the future. This research and its inquiries advanced deep learning architectures to segment the tumor robustly in all tumour subregions and actively preserve the details of the tumor edges. The research attempts to meet these challenges to build on the U-Net, which is a deep learning architecture useful in volumetric medical image segmentation, but has commonly exhibited challenges in small regions and boundary segmentation [10].

The methodology in question is to use a mix of TransUNet architecture. This means taking a convolutional feature extraction with a U-Net arch. and merging it with a transformer model. This is done in order to retrieve more global context. In order to increase how accurately the model is able to predict the edges of a tumor and to improve overall segmentation, a Boundary Loss function, which has been used in remote sensing and is adapted in this case, is incorporated with a hybrid Dice and Cross Entropy loss. These loss functions are chosen to emphasize accurate boundary delineation and balance the segmentation of small tumor regions with larger areas. By comparing the performance of

the conventional U-Net baseline, the TransUNet baseline, and the proposed Boundary-Supervised TransUNet (BS-TransUNet) on the BraTS dataset, the study confirms achieving more precise segmentation of the enhancing tumor and sharper tumor boundaries. The purpose of this research is to provide a method that can improve clinical assessment, treatment planning, and monitoring of tumor progression.

2. Related Work

Medical Image Segmentation: Most medical image segmentation approaches are based on neural networks, mainly convolutional neural networks (CNNs). These networks include fully convolutional networks (FCN) [11] and autoencoders derived from the U-Net architecture. Variants of these networks have been applied for semantic segmentation tasks in both medical and non-medical images, where semantic segmentation assigns a class or label to every pixel or voxel in an image. In medical imaging, the Attention U-Net enhances the original U-Net by adding attention gates. These gates emphasize relevant information while suppressing irrelevant information in the skip connections of the U-Net. The Attention U-Net showed pancreas segmentation, where it proved a strong performance highlighting benefits of soft attention mechanisms [12]. Another variation, UNet++ proposed by Z. Zhou, et. al., introduced skip connections in U-Net's connections, adding denser convolutional blocks in order to replace the original skip pathways. The output of these blocks is then concatenated with that of the corresponding encoder node and the preceding decoded node. This step helps to improve segmentation accuracy but adds new parameters and computational complexity [13]. Besides U-Net models, cascaded convolutional neural networks also present promising results. H. Roth, et al. used a two-stage cascaded network for organ and major vessel segmentation in the abdomen. In stage one, it creates a human body mask, which in turn is utilized in a second stage for the final segmentation. This technique greatly boosted the range of mean Dice score from 68.5% to 82.2% [14].

Brain Tumor Segmentation: The aim of brain tumor segmentation is to identify various areas in the brain, including normal, peritumoral edematous, and enhancing tumor brain structures, and necrotic, non-enhancing tumor regions [15]. This work uses the BraTS 2023 datasets, which are multi-modal MRI data with various slices. Brain tumor segmentation is a particularly difficult task, given that it involves data of a high dimensionality and marked variations from scan to scan, including differences in tumor shape, size, and location [16]. Furthermore, the poor quality of some of the MRI images

taken contributes to this difficulty. Prior studies have examined the extent to which the fusion of 2D slices can lead to a better segmentation than using full 3D MRI volumes. Avesta et al. compared 2D slices, 2.5D (five consecutive slices), and complete 3D MRI data for anatomical structure segmentation, such as the 3rd ventricle, thalamus, and hippocampus. Their findings showed that 3D MRI inputs were evidently much superior to 2D and 2.5D models, even though 3D models needed additional memory for training than 2.5 [17].

Attention Mechanism: Attention mechanisms are extensively used in deep learning to enhance model performance. These mechanisms facilitate the network identifying important input features while discounting any information that wasn't relevant [18]. Attention was first popularized in the transformer model, originally developed for natural language processing, where it enables the model to capture relationships and the importance of different words, leading to significant performance gains. Attention mechanisms have since been adapted for applications in computer vision and geometric deep learning [19]. There are two main types of attention: soft and hard attention. Soft attention assigns weights to different inputs, making the mechanism fully differentiable and trainable via gradient descent, as in the transformer. Hard attention also maps relationships between inputs, but it is non-differentiable, making gradient-based learning impossible [20]. Another approach that extends TransUNet by implementing self-attention into the U-Net architecture by using a Transformer encoder to convert CNN feature maps into patch-tokens for capturing global context, and a cross-attention decoder to refine regions by fusing Transformer representations with U-Net features. This framework supports multiple configurations (encoder, decoder, or both) and has flexible 2D/3D implementations that can be adjusted for various medical segmentation tasks [21].

Uncertainty Estimations: Uncertainty estimates for neural networks are a must, especially for their application in medicine. Deterministic models of neural networks, Bayesian models, ensemble methods, and test-time augmentation. Deterministic models estimate uncertainty either using an external model or through predictions from the network itself. Activation functions such as softmax, sigmoid, etc, which are usually used as proxies for uncertainty, though their estimates are usually overconfident and unreliable in result [23]. To tackle this, alternative methods like that of the Dirichlet distribution are used. Another strategy involves aggregating multiple predictions, such as combining outputs from different model weights in Bayesian methods, using ensembles, or aggregating predictions from multiple augmentations of the same input image [24].

While existing methods as above significantly improved medical image segmentation, most approaches rely solely on convolutional architectures which do not fully capture global context crucial for precise brain tumor boundary delineation [10]. Transformer based models have demonstrated superior capability in modeling global contextual relationships [13]. Unlike prior works that typically optimize segmentation using standard losses, the proposed method utilizes a hybrid TransUNet architecture augmented with boundary aware and hybrid Dice Cross Entropy loss functions to enhance edge sensitivity and improve segmentation of both small and large tumor subregions.

3. Methodology

3.1 Overview of the Approach

To address this limitation, we use the TransUNet architecture with our composite loss function, now termed the Boundary-Supervised TransUNet (BS-TransUNet), using Dice, Cross-Entropy, and the Boundary Loss adopted from [28], each weighted to optimize both region overlap and edge definition. This improvement ensures that the model concentrates more on the edge information and Boundary pixels, which enhance segmentation performance in small and low-contrast areas.

3.2 Dataset Description

For this research, we employed the BraTS 2023 dataset, which has been found to be the most trustworthy among other datasets available benchmarks for brain tumor segmentation [25]. This set of data involves multi-modal MRI images from patients with gliomas. Every case has four MRI modalities:

- T1-weighted (T1)
- T1 contrast-enhanced (T1ce)
- T2-weighted (T2)
- FLAIR images

Together with the MRI images, there are also ground truth masks available in the dataset. segment the tumor into three subregions:

- Enhancing Tumor (ET)

- Tumor Core (TC)
- Whole Tumor (WT)

In our experiment, we split our data into 70% for training, 15% for validation, and 15% for testing.

3.3 Data Preprocessing

Prior to training, preprocessing techniques are used for data preparation of the MRI images which includes:

- Because of computational constraints, each volume was resized to $128 \times 128 \times 100$. Every 3D scan consisted of 100 slices being extracted and stored as separate PNG images. This analysis concentrates on T1CE and FLAIR modalities, as these offer supplementary information about brain structures and tissue contrast. These two were layered along the channel dimension, yielding an input with a shape of $2 \times 128 \times 128$.
- Normalization: To address differences in intensity values among various scanners, patients, every MRI modality has been normalized using z-score normalization, which involves subtracting subtracted, and then normalised by taking the standard deviation of voxel intensities.
- Mask Encoding: The labels of segmentation are turned into one-hot encoded masks, where each channel represents a different tumor sub-region (ET, TC, WT). This representation enables it to make independent predictions for every area.

This preprocessing is used for standardizing input data, which in turn affects normalizing training processes, which results in a consistent model performance.

3.4 Baseline Models

3.4.1 U-Net Architecture

The U-Net architecture consists of two main parts: a contracting path, shown in Figure 1 on the left and a wide path on the right.

The contracting path acts like a regular convolutional encoder. It has two successive 3×3 convolutional layers (without padding), each followed by a Rectified Linear followed by a unit ReLU activation and then a 2×2 max-pooling operation with a stride of 2 to perform

downsampling. With each downsampling, the number of feature channels is increased two fold. This allows the network to capture higher-level, more abstract, and complex representations of the input in higher layers.

The expansive path performs the decoding operation to recover spatial resolution. Each step in this path begins with an upsampling of the feature map, followed by a 2×2 transposed convolution (“up-convolution”) that reduces the number of feature channels by half. The resulting feature map is concatenated with the corresponding cropped feature map from the contracting path through skip connections. Two 3×3 convolutions with ReLU activations are then applied to refine the fused features. Cropping of the encoder features is required because border pixels are lost during unpadded convolutions.

At the output, a 1×1 convolution maps each 64-channel feature vector to the desired number of segmentation classes. Overall, the architecture comprises 23 convolutional layers. To ensure a seamless tiling of the predicted segmentation maps, the input image size must be chosen such that all 2×2 max-pooling operations act on feature maps with even spatial dimensions.

U-Net has proven highly effective for medical image segmentation because of its ability to combine local and global features efficiently. It serves as our first baseline model trained with Dice loss and cross-entropy loss.

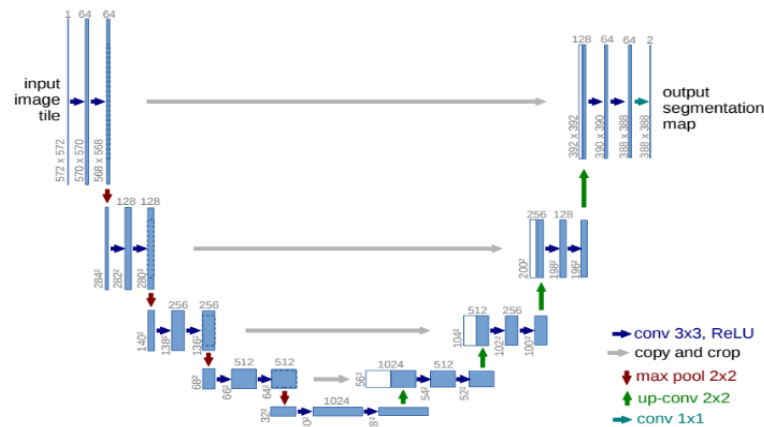


Figure 1. U-Net architecture(example for 32×32 pixels in the lowest resolution).

3.4.2 TransUNet Architecture

The TransUNet architecture extends the conventional U-Net by integrating Transformer modules to model global dependencies in medical images. As shown in Figure 2 (a), the Transformer encoder operates on a sequence of image patches extracted from CNN feature maps. Each patch is linearly embedded into a vector representation, combined with positional embeddings, and processed through multiple layers of multi-head self-attention and feed-forward networks. This mechanism enables the model to capture long-range contextual relationships that traditional convolutional layers may miss.

In the complete architecture shown in Figure 2 (b), the encoder first applies convolutional layers to extract low-level spatial features.

This hybrid CNN Transformer design allows for the local texture representation of CNNs, along with leverage global semantic understanding from Transformers.

3.5 Boundary-Aware Training Strategy

Although the TransUNet architecture combines both convolutional networks and transformers in their model structure, features for effective medical image segmentation, it suffers from some of its own features, in that it can produce blurred or inaccurate boundaries of tumors. These boundaries of tumors are particularly apparent when it comes to small and low-contrast enhancing tumors. To address this problem, we use a boundary-aware training strategy that can improve the attention mechanism for fine details of boundaries in TransUNet without changing architecture.

The key part of this approach is to combine Dice, Cross-Entropy, and Boundary Aware losses. These losses provide different strengths. Dice Loss aims to optimize the region, reduce overlap and enhance global segmentation consistency. Cross-Entropy Loss offers stable pixel-wise classification. Boundary-Aware Loss particularly focuses on precision in tumor using penalties for discrepancies between predicted boundaries and actual boundaries.

The combined objective promotes region-level accuracy as well as improvement of local boundary delineation. The weighting coefficients in each of these losses were tuned balancing a trade-off between smoothness and detail.

All architectural parameters, hyperparameters, as well as preprocessing parameters, were set. aligned with that of the baseline TransUNet to provide a fair and controlled evaluation. This ensures a direct evaluation of the influence of hybrid loss functions on segmentation quality, especially in difficult-to-enhance areas of tumors.

3.5.1 Motivation

A standard segmentation loss regards all pixels as identical, which often ignores the diagnostic role of boundary pixels. In brain tumor segmentation, small discrepancies around a brain tumor area margin can cause substantial inaccuracies in patient care-related measurements, such as calculations of tumor size and treatment planning. The reasoning for using a boundary-aware loss function is to help the model's learning mechanism focus on delineating a precise tumor boundary, which plays a key role in precise diagnostics and radiotherapy planning. Because of introducing sensitivity over boundaries using the loss function, the model would learn to maintain structure-level information which usually gets blurred in training with conventional losses.

3.5.2 Boundary-Aware Loss Function

We augment the original hybrid loss function used in the baseline models with Boundary Loss [Shown in Figure 2 (b)], adopted from [28], which aims to provide a precise definition of object boundaries. The new total loss is given as:

$$Total\ Loss = \alpha L(Dice) + \beta L(Cross\ Entropy) + \gamma L(Boundary)$$

Boundary Loss: The boundary term, $L(Boundary)$, is a differentiable approximation of BF_1 metric (Boundary F_1 score), which aims at penalizing misaligned boundaries.. It is defined as:

$$L(Boundary) = 1 - BF_1$$

Where BF_1 is computed based on the precision and recall of the predicted boundary against an expanded ground truth boundary using pooling operations. This term forces the model to minimize errors specifically along the tumor edges.

Where BF_1 is computed based on the precision and recall of the predicted boundary against an expanded ground truth boundary using pooling operations. This term forces the model to minimize errors specifically along the tumor edges.

In this formulation, α , β , and γ are weighting coefficients that balance the contribution of each loss component. In our experiments, we empirically set $\alpha=0.5$, $\beta=0.5$, and $\gamma=0.1$,

which helps the network learn not only to segment regions correctly but also to produce sharp and accurate edges, particularly for the enhancing tumor (ET) area.

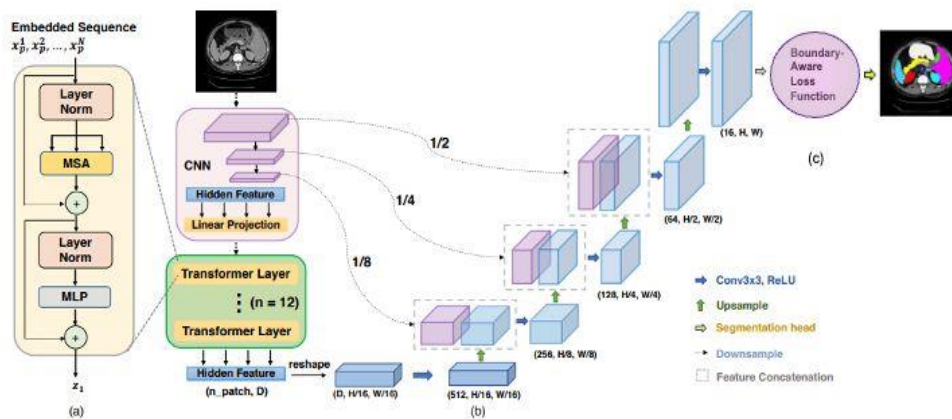


Figure 2. Overview of the TransUNet

3.6 Training Configuration

All models U-Net, TransUNet, and Boundary-Supervised TransUNet are trained under identical experimental settings for fair comparison. Framework: PyTorch Optimizer: Adam with initial learning rate of 0.0003 (base learning) along with ReduceLROnPlateau mode set to 'min' with factor 0.2, patience 2 and minimum learning rate \$0.000001 Batch Size: 64 Epochs: 30 Loss Functions: As defined in previous sections Hardware: NVIDIA TESLA P100 GPU with 16 GB of VRAM Early Stopping: Applied if validation loss does not improve for 5 consecutive epochs Model checkpoints are saved based on the highest validation Dice score. During training, learning rate scheduling is used to gradually reduce the learning rate once the validation loss plateaus.

3.7 Evaluation Metrics

To evaluate segmentation performance, we use commonly accepted metrics in medical image segmentation task which includes [25]:

- Dice Similarity Coefficient (DSC) – Measures overlap between prediction and ground truth.
- Dice Similarity for each of the predicted labels.
- Train and Validation Loss Dice provide a balanced view of both volumetric and boundary performance.

4. Implementation and Experiments

The implementation of our experiments was performed using PyTorch on NVIDIA TESLA P100. All models were trained under consistent conditions. This included the same optimizer (Adam, lr=0.0003), batch size (64), number of epochs (30). We compared two models:

We compared two models:

- Baseline TransUNet trained with Dice + Cross-Entropy loss.
- TransUNet with Boundary-Aware Loss: Dice (weighted 0.5) + Cross-Entropy (weighted 0.5) + Boundary loss (weighted 0.1).

During training, loss values were monitored on both training and validation sets, and Dice similarity coefficients were computed for each tumor subregion: Enhancing Tumor, Tumor Core, and Whole Tumor. The boundary-aware loss was designed to guide the network toward sharper tumor boundaries.

Training and evaluation pipelines were fully automated. Model checkpoints were saved based on the highest validation Dice score, and early stopping was applied if validation loss did not improve for 5 consecutive epochs due to computation resources.

5. Results

The performance of the baseline TransUNet and BS-TransUNet was evaluated on the BraTS dataset using training/validation loss curves and Dice-based segmentation metrics. The learning dynamics for both models are presented in Figure 3. Although both training runs converge quickly, the two-loss setup typically shows lower overall loss values

compared to the three-loss setup. This difference occurs because adding Boundary Loss introduces an additional optimization objective, increasing the total loss magnitude even when the model is improving. This behaviour indicates improved generalization and a reduced tendency to overfit, particularly in regions where tumor boundaries are ambiguous or weakly contrasted. Table 1 presents the final training and validation losses for the baseline TransUNet and the proposed BS-TransUNet. Contrary to the expectation that boundary-aware loss would reduce the overall losses, the baseline TransUNet achieved lower training (0.02) and validation loss (0.03) compared to BS-TransUNet, which obtained 0.08 and 0.09, respectively because adding boundary loss introduces an additional optimization objective, increasing the total loss magnitude. This indicates that, in terms of global loss optimization, the baseline model converged more efficiently. However, loss values alone do not fully capture segmentation quality, especially for tumor regions with complex boundaries.

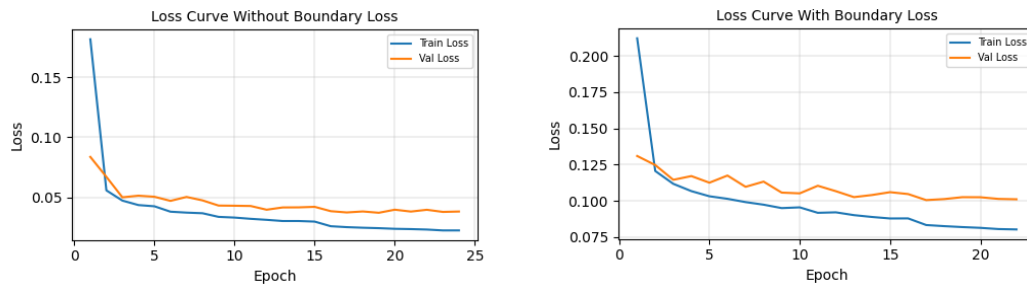


Figure 3. Training and validation loss curves for models trained with and without the boundary-aware loss.

Table 1. Training and validation loss for baseline and BS-TransUNet.

Model	Train Loss ↓	Validation Loss ↓
TransUNet	0.02	0.03
BS-TransUNet(Ours)	0.08	0.09

A more detailed evaluation using Dice similarity coefficients is given in Table 2. Here, The BS-TransUNet yields evident performance improvements for multiple tumor sub-

regions despite having higher training and validation losses. The overall Dice coefficient improves from 0.862 (baseline) to 0.879 under boundary-aware supervision. For the core of the Necrotic tumor, where the Dice score increases significantly from 0.903 to 0.913, which makes more reliable delineation of the internal tumour structures. Improvement in enhancing tumour performance does also increase slightly ($0.927 \rightarrow 0.928$), retaining high precision on the most clinically important subregion.

Table 2. Dice scores for each tumor subregion comparing baseline TransUNet and BS-TransUNet.

Model	Dice Coefficient $\uparrow$	Dice Necrotic $\uparrow$	Dice Edema $\downarrow$	Dice Enhancing $\uparrow$
TransUNet Model	0.862	0.903	0.911	0.927
BS-TransUNet(Ours)	0.879	0.913	0.909	0.928

However, the Dice value for the sub-region of edema slightly decreases ($0.911 \rightarrow 0.909$). Hence, despite matching with the conclusion in regard to improving the performance as well as enhancing tumor segmentation, it does not match with improving the Dice performance for the sub-region of edema, as no improvement is observed in the numerical Dice value for edema subregion.

Overall, the gain in Dice suggests that focusing on boundary information better assists the model in capturing fine structural details even when the total loss magnitude does not decrease. Qualitative observations further strengthen the quantitative results: BS-TransUNet produces smoother contours with less fragmentation around problematic regions like the Refining Tumor. These improvements indicate the advantage of incorporating the boundary-guided Supervision in segmentation models based on transformers.

Overall, the results show that while BS-TransUNet is not considering minimizing the global loss, as effective as the baseline, it provides stronger and more anatomically coherent

segmentation performance, especially for tumour subregions where boundary precision is most important clinically.

Figure 5 shows the segmentation results of the TransUNet model trained with the boundary-aware loss. As with the baseline, the top row shows the original slice, ground truth mask, and individual subregions, while the bottom row shows the predicted masks. Compared to the baseline, the boundary-aware model produces sharper and more accurate boundaries, particularly for the Enhancing Tumor region, while maintaining accurate segmentation of other tumor subregions. This shows that if the boundary is included, it effectively enhances the edge delineation without affecting volumetric segmentation performance.

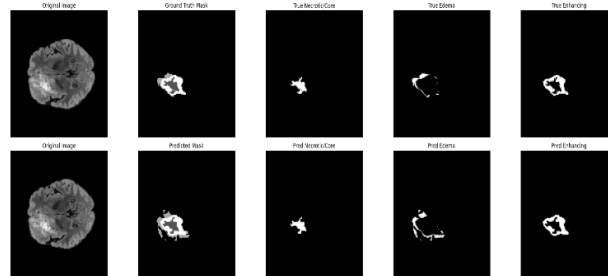


Figure 4. Visualization of brain tumor segmentation results on the 70th slice with Base TransUNet.

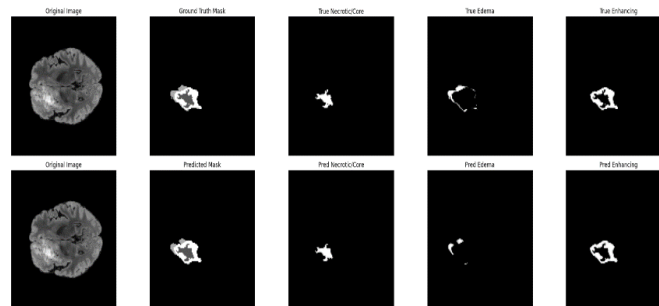


Figure 5. Visualization of brain tumor segmentation results on the 70th slice with TransUNet and Boundary Loss.

6. Conclusion and Future Work

In this research, we evaluated the performance of the TransUNet model on the BraTS dataset for brain tumor segmentation and used an extension with a boundary-aware loss to improve edge delineation.

Our experiments have found that the initial TransUNet model works well in identifying prominent tumorous areas but tends to find it difficult in small and low-contrast ones especially in the Enhancing Tumor area. With the incorporation of our boundary-aware loss, we found a more stable training procedure, a decreased overfitting rate, and sharper and more precise segmentation boundaries for clinically significant tumor subregions. Quantitative analysis results showed that improvements in Scores for Dice are shown for Necrotic and Enhancing Tumor, with qualitative visualization cleaner contours, fewer isolated false positives.

In our study we used a dataset including the data of the 600 patients. For future work; experiment can be expanded with a larger dataset to check the Dice scores especially for the Enema Tumor subregion.

References

- [1] L.-K. Ge, P. Gao, D. Chang, J.-J. Nie, Y.-S. Wang, X.-N. Zuo, and G.-X. Wei, "An open data for imaging acute aerobic exercise effects on brain and mind in emerging adulthood," *Scientific Data*, vol. 11, no. 1, p. 1422, 2024.
- [2] Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249.
- [3] U. Baid, S. U. Rane, S. Talbar, S. Gupta, M. H. Thakur, A. Moiyadi, A. Mahajan, Overall survival prediction in glioblastoma with radiomic features using machine learning, *Frontiers in Computational Neuroscience* 14 2020.
- [4] M. Usman Akbar, M. Larsson, I. Blystad, and A. Eklund, "Brain tumor segmentation using synthetic MR images: a comparison of gas and diffusion models," *Scientific Data*, vol. 11, no. 1, p. 259, 2024.
- [5] Ali, A., Syed, S. M., Syed, A. H., et al.: Brain tumor segmentation with deep learning: BraTS challenges from 2012 to 2020. *IEEE Trans Med Imaging* 40(10), 2557–2567 2022.
- [6] L. Dixon, G. K. Jandu, J. Sidpra, K. Mankad, Diagnostic accuracy of qualitative mri in 550 paediatric brain tumours: evaluating current practice in the computational era, *Quantitative Imaging in Medicine and Surgery* 12, 2022, 131–143.

- [7] S. Dorosti, T. Landry, K. Brewer, A. Forbes, C. Davis, and J. Brown, “High-resolution ultrasound data for AI based segmentation in mouse brain tumor,” *Scientific Data*, vol. 12, no. 1, p. 1322, 2025.
- [8] Myronenko, A.: 3D MRI brain tumor segmentation using autoencoder regularization. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 311–320. Springer 2019.
- [9] M. Shao, Z. Wang, H. Duan, Y. Huang, B. Zhai, S. Wang, Y. Long, and Y. Zheng, “Rethinking brain tumor segmentation from the frequency domain perspective,” *arXiv preprint arXiv:2506.10142**, Jun. 2025.
- [10] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2015*, Springer International Publishing, Cham, 2015, pp. 234–241.
- [11] E. Shelhamer, J. Long, T. Darrell, Fully convolutional networks for semantic segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 2017, 640–651.
- [12] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention u-net: Learning where to look for the pancreas, 2018.
- [13] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, 2018.
- [14] H. R. Roth, H. Oda, X. Zhou, N. Shimizu, Y. Yang, Y. Hayashi, M. Oda, M. Fujiwara, K. Misawa, K. Mori, An application of cascaded 3d fully convolutional networks for medical image segmentation, *Computerized Medical Imaging and Graphics* 66, 2018, 90–99.
- [15] Z. Li, H. Zhu, X. Zhao, Y. Gao, and Y. Han, “A multi brain tumor region segmentation model based on enhanced 3D U-Net,” *Applied Sciences*, vol. 13, no. 16, p. 9282, Aug. 2023.
- [16] T. T. M. Thao, P. V. Truong, and N. H. Thang, “Developing a loss function with TransUNet for brain tumor segmentation from MRI images,” *J. Mech. Sci. Technol.*, vol. 36, no. 8, pp. 3993–4003, 2022.
- [17] A. Avesta, S. Hossain, M. Lin, M. Aboian, H. M. Krumholz, S. Aneja, Comparing 3d, 2.5d, and 2d approaches to brain image segmentation, *medRxiv* (2022).
- [18] S. Khan, M. Naseer, M. Hayat, F. Shahbaz Khan, M. Yang, and F. Porikli, “Transformers in vision: A survey,” *ACM Computing Surveys*, vol. 55, no. 4, pp. 1–41, 2022.
- [19] T. Gonçalves, I. Rio-Torto, L. F. Teixeira and J. S. Cardoso, “A survey on attention mechanisms for medical applications: are we moving towards better algorithms?,” 2022.
- [20] G. Brauwiers, F. Frasincar, A general survey on attention mechanisms in deep learning, *IEEE Transactions on Knowledge and Data Engineering* 35, 2023, 3279–3298.

- [21] J. Chen, J. Mei, X. Li, Y. Lu, Q. Yu, Q. Wei, X. Luo, Y. Xie, E. Adeli, Y. Wang, M. P. Lungren, S. Zhang, L. Xing, L. Lu, A. Yuille, and Y. Zhou, "TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers," *Medical Image Analysis*, vol. 97, 2024, Art. no. 103280.
- [22] K. Zou, Z. Chen, X. Yuan, X. Shen, and M. Wang, "A review of uncertainty estimation and its application in medical imaging," Feb. 2023.
- [23] M. Teye, H. Azizpour, and K. Smith, "Bayesian uncertainty estimation for batch normalized deep networks," *Pattern Recognition*, vol. 142, p. 109741, 2023.
- [24] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, M. Shahzad, W. Yang, R. Bamler, X. X. Zhu, A survey of uncertainty in deep neural networks, 2022.
- [25] M. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [26] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "TransUNet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [27] K. Avazov, S. Mirzakhililov, S. Umirzakova, A. Abdusalomov, and Y. I. Cho, "Dynamic Focus on Tumor Boundaries: A Lightweight U-Net for MRI Brain Tumor Segmentation," **Bioengineering (Basel)**, vol. 11, no. 12, p. 1302, Dec. 2024.
- [28] A. Bokhovkin and E. Burnaev, "Boundary Loss for Remote Sensing Imagery Semantic Segmentation," *Proc. 16th Int. Symp. Neural Netw.*, 2019.