Medical Image Analysis Project

Segmentation of Gliomas in MRI

(BraTS Dataset)

Submitted By

Sparsh Sharma (DA24C028)
Shravasti Sarkar (DA24C018)
Navya Garg (DA24C011)

Indian Institute of Technology Madras

November 14, 2025

GitHub Repository:
`https://github.com/Navya-146/Medical-Image-Analysis-2025`

1 Introduction

Problem statement. Automatic segmentation of glioma sub-regions (WT, TC, ET) from multi-parametric MRI (T1, T1ce, T2, FLAIR) is essential for treatment planning and quantitative analysis. The task is formulated as a 3D semantic segmentation problem where each voxel is classified as background or tumour tissue.

2 Literature Review

Since the inaugural BraTS challenge in 2014, deep convolutional neural networks have dominated brain tumor segmentation tasks. Advances in GPU technology, the growing availability of annotated datasets, and improvements in deep learning methods have further strengthened the performance of these models. Among various architectures, the U-Net [10] has become a widely adopted baseline due to its strong segmentation capabilities. Several modifications to the standard U-Net have been proposed to enhance its performance, including the incorporation of residual connections, densely connected layers [9], and attention mechanisms [2].

All winning solutions from the past six editions of the BraTS challenge have relied on deep neural networks. In particular, Kamnitsas et al. [5], the 2017 winners, introduced an ensemble approach known as EMMA (Ensemble of Multiple Models and Architectures). This strategy combined several 3D convolutional models, including DeepMedic [7, 6], fully convolutional networks [8], and U-Net [10], which helped reduce sensitivity to individual model hyperparameters and mitigated overfitting.

For the 2020 BraTS challenge, the winning team [3, 4] adopted nnU-Net [3] as their foundation, a self-configuring U-Net variant, and introduced several task-specific enhancements. Key improvements included focusing on regional segmentation rather than individual classes, increasing batch size from 2 to 5, applying more extensive data augmentation, replacing instance normalization with batch normalization to better handle augmented data, using batch Dice loss instead of standard Dice, and employing customized post-processing techniques.

The 2021 challenge winners [9] also built upon the nnU-Net framework, reusing the prior BraTS-specific optimizations while adding new modifications. With the dataset expanding from 494 to 1470 cases, they scaled the network by doubling the encoder filters and increasing the bottleneck capacity to 512, while maintaining the decoder unchanged. Additionally, batch normalization was swapped for group normalization, allowing smaller batch sizes (2 instead of 5). An axial attention decoder was tested but did not yield performance gains in 5-fold cross-validation.

3 Dataset Description

The BraTS dataset [1] provides paired multi-parametric MRI volumes (T1, T1ce, T2, FLAIR) and voxel-wise labels for the three glioma regions: Whole Tumor (WT), Tumor Core (TC), and Enhancing Tumor (ET). Each modality offers complementary contrast—FLAIR for edema, T1ce for enhancing boundaries, and T2 for fluid abnormalities—making multi-channel input crucial for accurate segmentation.

Preprocessing

Preprocessing steps include joint non-zero cropping to remove redundant background while maintaining spatial alignment across all modalities, which reduces memory usage. Intensity normalization is performed via per-channel z-scoring on non-zero voxels to stabilize contrast between subjects. Labels are remapped from their original values (0, 1, 2, 4) to a compact set (0, 1, 2) to facilitate cleaner softmax outputs. Images are resampled using trilinear interpolation, while labels use nearest-neighbor interpolation to preserve boundaries. Finally, the preprocessed volumes are stored as compressed `.npz` files for efficient and deterministic loading.

4 Methodology

The segmentation model is a 3D U-Net implemented using MONAI, chosen due to its proven performance on volumetric medical imaging tasks and its encoder–decoder structure that preserves fine spatial detail while capturing global tumour context. The network uses increasing feature depths (e.g., 16–256) with strided convolutions for downsampling, enabling hierarchical feature extraction across scales. Skip connections are essential for tumour segmentation because glioma boundaries often contain subtle gradient; these connections ensure that shallow, high-resolution spatial features are preserved and fused with deep semantic information. A patch-based training strategy is adopted to manage memory usage and increase the effective number of training samples.

Training uses a hybrid Dice + CrossEntropy loss, a standard but empirically validated choice for medical segmentation. Dice loss directly optimizes region overlap, which is critical for handling class imbalance as tumour voxels often occupy less than 3% of the volume. CrossEntropy stabilizes early training and improves voxel-wise discrimination. AdamW is selected as the optimizer due to its decoupled weight decay, which allows better generalization under strong data augmentations. Automatic Mixed Precision (AMP) is used to reduce memory consumption and accelerate training while keeping numerical precision in sensitive operations via dynamic loss scaling. Data augmentation is implemented through TorchIO to counter overfitting and simulate realistic anatomical variance. Random affine transforms and elastic deformations approximate natural patient-to-patient anatomical variability, while noise, blur, and intensity shifts mimic scanner artefacts and acquisition differences. These augmentations are crucial because BraTS includes images from multiple hospitals, and a model trained without augmentation tends to overfit to institution-specific intensity profiles. Evaluation uses per-class Dice, HD95, and BraTS regional metrics (WT/TC/ET), ensuring consistency with challenge standards and enabling easy comparison to reported baselines. Deterministic data splits, fixed seeds, and explicit checkpointing ensure full experiment reproducibility.

5 Results & Evaluation

5.1 Training Curve Analysis

This section summarizes the behavior of the model across 50 epochs using the training log. The curves provide a picture of model convergence and evolution of segmentation performance during optimization.

5.1.1 Train Loss vs. Epoch

The training loss (Fig. 1) decreases smoothly from 0.53 to 0.20 over 50 epochs, with no spikes or instability. The monotonic trend indicates stable gradients and consistent optimization. The mean loss is 0.27 (SD 0.08), reflecting steady improvement throughout training.

5.1.2 Validation Dice vs. Epoch

The validation Dice score (Fig. 2) improves from 0.61 to 0.77, reaching its peak value of 0.77 at epoch 46. The mean Dice is 0.73 (SD 0.04), with an early rise to 90% of the final score by epoch 14. Fluctuations are minor, and the curve shows stable, gradually improving generalization with no indication of overfitting.

5.1.3 Validation Hausdorff95 vs. Epoch

The Hausdorff95 distance (Fig. 3) decreases from 14.33 to approximately 5.3–6.0 in later epochs, indicating improved boundary accuracy. The mean value is 7.35 (SD 2.29). The total improvement across training is $\Delta H \approx 8.93$, showing a substantial reduction in contouring errors over time.

Based on all curves and statistics, the training process is stable, monotonically improving, and free of pathological behavior.

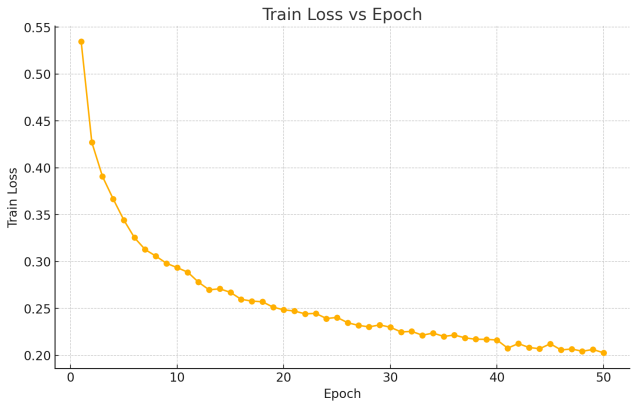


Figure 1: Training loss across 50 epochs.

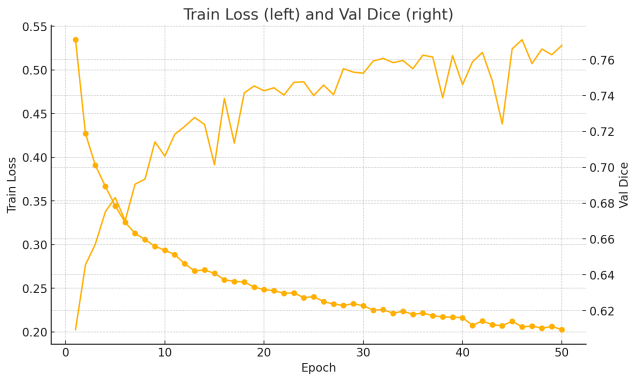


Figure 2: Train loss and validation Dice using dual axes.

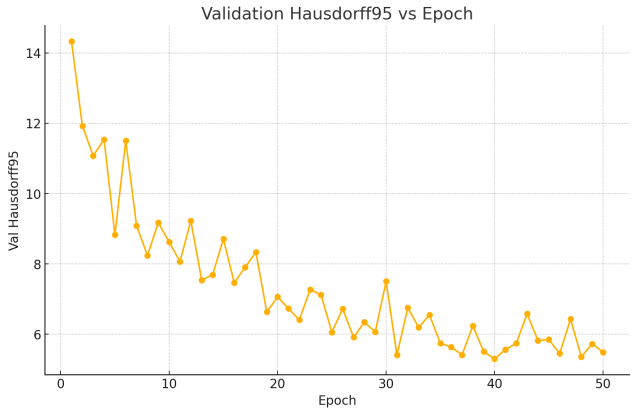


Figure 3: Validation Hausdorff95 distance per epoch. Lower values indicate better boundary accuracy.

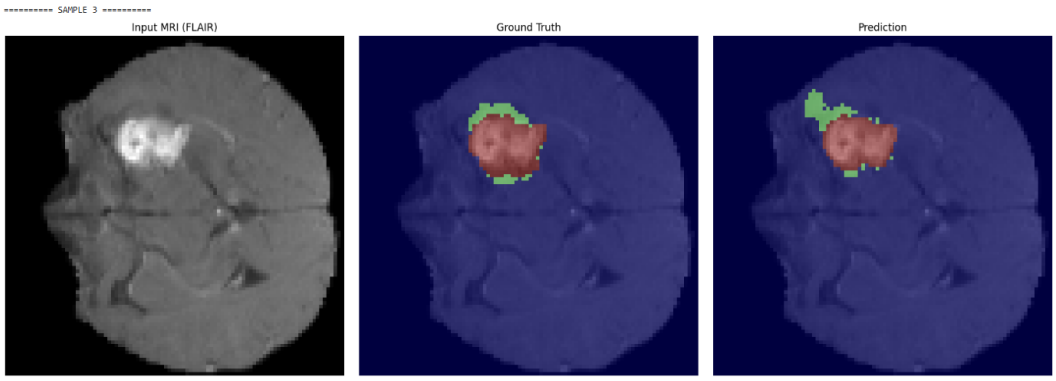
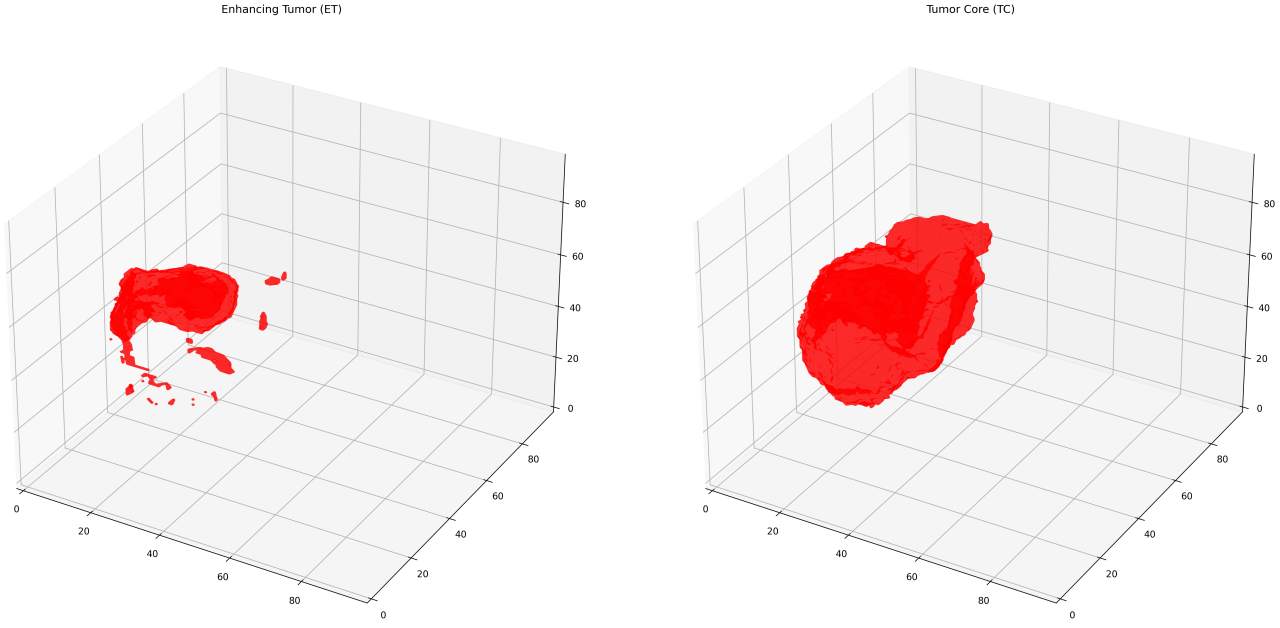


Figure 4: Input Image, Ground Truth and Prediction Visualization for a Brain Tumor sample.



(a) Enhancing Tumor (ET) (b) Tumor Core (TC)

Figure 5: 3D visualizations of tumor subregions: (a) Enhancing Tumor and (b) Tumor Core.

5.2 Final Evaluation Report

Table 1: Final evaluation metrics for per-class and BraTS regions.

Metric	Class 0 (NCR/NET)	Class 1 (ED)	Class 2 (ET)
Dice	0.7907	0.7266	0.7266
Hausdorff95	5.98	5.03	5.42
BraTS Region Dice:			
Whole Tumor (WT)	0.999999999999435		
Tumor Core (TC)	0.9965942204600237		
Enhancing Tumor (ET)	0.7265807559478199		
Overall Mean Dice	0.7586402893066406		

Table 2: Summary of training and evaluation statistics.

Metric	Value
Final Dice	0.7680
Best Dice (epoch 46)	0.7712
Initial $\rightarrow$ Final Hausdorff95	$14.33 \rightarrow 5.49$
Dice (mean / std)	0.7291 / 0.0368
Hausdorff95 (mean / std)	7.35 / 2.29

These results demonstrate consistent convergence and strong segmentation performance across training.

References

[1] Ujjwal Baid, Satyam Ghodasara, Suyash Mohan, Michel Bilello, Evan Calabrese, Errol Colak, Keyvan Farahani, Jayashree Kalpathy-Cramer, Felipe C Kitamura, Sarthak Pati, et al. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*, 2021.

[2] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H.R. Roth, and D. Xu. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI Brainlesion Workshop*, pages 272–284. Springer, 2021.

[3] F. Isensee, P.F. Jaeger, S.A. Kohl, J. Petersen, and K.H. Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, 2021.

[4] F. Isensee, P.F. Jäger, P.M. Full, P. Vollmuth, and K.H. Maier-Hein. nnu-net for brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 118–132. Springer, 2021.

[5] K. Kamnitsas, W. Bai, E. Ferrante, S. McDonagh, M. Sinclair, N. Pawlowski, M. Rajchl, M. Lee, B. Kainz, D. Rueckert, et al. Ensembles of multiple models and architectures for robust brain tumour segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 450–462. Springer, 2018.

[6] K. Kamnitsas, L. Chen, C. Ledig, D. Rueckert, and B. Glocker. Multi-scale 3d convolutional neural networks for lesion segmentation in brain mri. *Ischemic Stroke Lesion Segmentation*, 13:46, 2015.

[7] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, and B. Glocker. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis*, 36:61–78, 2017.

[8] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[9] H.M. Luu and S.H. Park. Extending nn-unet for brain tumor segmentation. In *International MICCAI Brainlesion Workshop*, pages 173–186. Springer, 2021.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pages 234–241. Springer, 2015.