Brain Tumor Segmentation Using 3D U-Net Model

Manan Kansal, Shreya Sharma (Team No.172)
Bennett University
e22cseu1668@bennett.edu.in, e22cseu1649@bennett.edu.in

March 30, 2025

Abstract

Brain tumor segmentation is crucial for accurate diagnosis and treatment planning in medical image analysis. Here, we introduce a 3D U-Net model with spatial attention and SpatialDropout3D for enhancing the segmentation accuracy. Our model is trained on the BraTS 2020 dataset and segments brain tumors into several classes. We measure performance based on Dice coefficient, precision, sensitivity, and specificity. outcomes demonstrate the efficiancy of our method in automated brain tumor segmentation and its applicability in clinical scenarios. Our approach enhances feature extraction and minimizes overfitting, resulting in more accurate and consistent segmentation results in medical images.

Keywords: Brain Tumor, 3D U-Net, Deep Learning, Medical Image Segmentation, MRI

Contents

1 Introduction

Gliomas are the most dangerous and common brain tumors, which are among the most dangerous neurological diseases. The brain's supportive tissue is composed of glial cells, and they are responsible for gliomas. Because of their highly heterogeneous structure, gliomas are challenging to diagnose and treat. The **enhancing tumor (ET)**, **necrotic/non-enhancing tumor core (NET)**, and **peritumoral edema (ED)** are sub-regions of the glioma region, which appears in any region of the brain [1]. Tumor growth assessment and treatment planning depend on the accurate segmentation of these sub-regions in MRI images. However, radiologist manual segmentation is a time-consuming process that takes a lot of time and is subject to inter-observer variability, which may affect the accuracy of diagnosis and treatment.

The Medical Image Computing and Computer-Assisted Interventions (MICCAI)-conducted Brain Tumor Segmentation (BraTS) [2] Challenge provides a very important dataset for the automatic segmentation of brain tumors to help tackle such problems. Multi-parametric MRI scans, including **FLAIR**, **T1**, and **T1CE**, are part of the BraTS 2020 dataset. The scans capture different tumor features, and with a more comprehensive analysis, it helps identify potential problems that might have been overlooked. Automated segmentation works towards developing deep learning models that accurately detect tumor sub-regions, reducing the requirement for human annotation and improving clinical judgment.

Medical image analysis has significantly improved thanks to deep learning, especially convolutional neural networks (CNNs). In biomedical segmentation tasks, **U-Net**, a popular encoder-decoder-based architecture with skip connections, has demonstrated exceptional performance. This method is extended to volumetric data by U-Net variants like **3D U-Net** [2], which increases the segmentation accuracy of 3D medical images.

In this work, we suggest a **3D U-Net** model for brain tumor segmentation that combines **spatial attention** and **spatial dropout3D**. Using a multi-GPU **MirroredStrategy** for effective training, our model is trained on the BraTS 2020 dataset. We use a combination of **Dice Loss** and **Categorical Cross-Entropy Loss** to optimize performance, and we assess the outcomes using metrics like specificity, precision, sensitivity, and Dice coefficient.

The rest of this paper is structured as follows: Section 2 explains the dataset and preprocessing steps. Section 3 discusses the proposed model architecture and training methodology. Section 4 presents experimental results and segmentation performance.

2 Related Work

Segmentation of brain tumors has become quite remarkable with deep learning, particularly with **convolutional neural networks (CNNs)**. Handcrafted feature-based and traditional machine learning approaches were previously followed, but current trends are dominated by deep learning methods due to their high performance. **Jain and Sachdeva (2024)** conducted a systematic review of deep learning models for the identification of brain tumors, contrasting model architectures, data usage, and performance metrics [3]. Similarly, **Bouhafra and El Bahi (2024)** compared and reviewed MRI-based classification techniques between 2020 and 2024 with focus on segmentation enrichment and application of attention mechanisms for precision enhancement [4]. Such reviews provide us valuable information on research trends and potential future developments regarding

brain tumor segmentation. This section presents notable recent research studies, including 2D, 3D, and hybrid CNN architectures, and compares them to our approach.

2.1 2D U-Net and Variants

The U-Net architecture, introduced by Ronneberger et al., revolutionized medical image segmentation by employing an encoder-decoder structure with skip connections. Building upon this, Pereira et al. enhanced the U-Net by incorporating feature reorganization and recalibration modules for tumor substructure segmentation [5]. Jiang et al. proposed a cascaded U-Net framework with two stages—an initial coarse segmentation step followed by refinement using a secondary decoder [6]. Similarly, Gu et al. introduced an attention-guided U-Net to selectively enhance important tumor regions while suppressing less relevant features [7]. Colman et al. developed DR-Unet104, a deep residual U-Net with 104 convolutional layers, integrating stacked "bottleneck" residual blocks for improved feature transfer [8].

Although these 2D-based approaches have achieved competitive segmentation performance, a major limitation is the loss of volumetric information, as MRI scans are inherently 3D.

2.2 3D CNN-Based Approaches

To overcome the shortages of 2D models, scientists have worked on 3D convolutional ar-chitectures that can maintain spatial details between slices. Isensee et al. created a dropout layer and deep data augmentation improved 3D U-Net variant to maximize generalization[2]. Tian et al. extended 3D U-Net through the incorporation of stacked residual blocks, enhancing the accuracy of segmen-tation while maintaining convergent stability [9]. Mlynarski et al. introduced a 2D-3D mixed network to improve computational complexity while seg- mentation accuracy [10]. Chen et al. also presented S3D-UNet, which leverages separable 3D convolutions to minimize computational expenses while ensuring segmentation performance [11]. Ghaffari et al. used a deep densely connected 3D CNN for tumor segmentation to provide improved gradient flow with dense connections [12]. In the same manner, Weninger et al. integrated self-attention mechanisms into 3D segmentation networks for improved feature representation, enhancing tumor boundary delineation [13].

2.3 Hybrid and Attention-Based Networks

Some works have ensembled multiple methods to achieve optimal segmentation performance. Sun et al. introduced a multi-scale fusion network that combines complementary features across various scales and uses an additive attention mechanism to enhance encoder-decoder relations [14]. Choi et al. presented a dual CNN method, with one network capturing global context and another refining voxel-wise classifications [15]. Vaanathi et al. proposed a multi-U-Net ensemble that handles MRI slices along three axes (axial, sagittal, coronal), making complete use of volumetric data [16]. Integrating domain adaptation methods, Shaikh et al. experimented with unsupervised adversarial learning for ensuring robustness in segmentation on various datasets [17]. Further, Myronenko et al. combined a Variational Autoencoder (VAE) branch with a 4 3D segmentation network to regularize the encoder output, maximizing feature clustering[16]

.

Incorporating domain adaptation techniques, Shaikh et al. explored unsupervised adversarial learning, ensuring robust segmentation across different datasets [17]. Additionally, Myronenko et al. integrated a Variational Autoencoder (VAE) branch into a 3D segmentation network to regularize the encoder output, optimizing feature clustering [18].

2.4 Loss Functions and Optimization Strategies

Since class imbalance is common in medical images, segmentation performance significantly depends on the choice of the loss function. Since dice loss is sensitive to overlapping areas, it is often employed. For the purpose of boosting segmentation accuracy among tumor subregions, Tian et al. introduced a hybrid loss function that combines binary cross-entropy and dice loss[19]. To facilitate accurate segmentation around tumor edges, Rathore et al. used boundary-aware loss functions[20]. Wang et al., however, employed contrastive learning to enhance tumor core detection by enhancing feature separability [21]. Inspired by these works, our model uses categorical cross-entropy and dice loss in a weighted combination to promote balanced learning among different tumor subregions.

2.5 Our Proposed Model

Our proposed model extends 3D U-Net by integrating Spatial Attention and SpatialDropout3D, which enables the network to concentrate on important tumor regions without overfitting. Unlike the existing literature, which solely uses Dice loss, we use a hybrid loss function that combines Dice loss and Categorical Cross-Entropy loss for better class balance. We also employ multi-modal MRI scans (FLAIR, T1, T1CE) to improve tumor subregion segmentation. Our results demonstrate increased segmentation accuracy, particularly in challenging regions such as the enhancing tumor and tumor core. Our test results demonstrate exceptional segmentation accuracy, particularly in challenging areas such as the enhancing tumor and tumor core.

Our experimental results demonstrate superior segmentation accuracy, particularly in challenging areas like the enhancing tumor and tumor core.

2.6 Comparison of Methods

A comparative overview of current models and our work is presented in Table 1.

In summary, although current models have far progressed brain tumor segmentation, problems such as tumor delineation, overfitting, and computational efficiency remain. Our model resolves these problems through a 3D U-Net architecture with spatial attention, multi-modal fusion, and an optimized loss function, with performance that is better ter segmentation performance, especially for tumor core and enhancing tumor areas give for this

3 Dataset Description

The dataset employed for the brain tumor segmentation in this research is the **BraTS 2020** dataset, which comprises multi-modal MRI scans. This dataset is widely used for training deep learning models for brain tumor segmentation.

Model	Advantage	Disadvantage
Multi U-Net [16]	Strong feature learning with three U-Nets working together	Overfitting and incorrect tumor core predictions
2D–3D Hybrid [10]	Combines short-distance 3D and long-distance 2D contexts	Network complexity, underutilized tumor features
DR-Unet104 [8]	Enhanced feature transfer via residual blocks	High variance in DSC between validation and test sets
Improved U-Net [2]	Deep supervision and extensive data augmentation	Long training time, limiting architecture variations
VAE-encoder [18]	Regularization via VAE branch for better feature clustering	Smaller input size reduces whole tumor segmentation accuracy
Ours	Spatial attention, multi-modal inputs, and hybrid loss function	Some spatial information loss between blocks

Table 1: Comparison of different models for brain tumor segmentation

3.1 Dataset Overview

The dataset used for brain tumor segmentation in this study is the BraTS 2020 dataset, which consists of multi-modal MRI scans. The dataset is widely used for deep learning models for brain tumor segmentation.

3.2 Dataset Overview

- **Modality:**Multi-modal MRI (T1, T1ce, T2, FLAIR)
- **Tumor Regions:** 3 tumor subregions
- **Format:** NIfTI (.nii.gz) images
- **Task:** 3D segmentation of brain tumors

3.3 MRI Modalities (4 Channels per Scan)

Each patient scan includes four MRI modalities, capturing different tissue characteristics:

- **T1-weighted (T1):** High contrast for anatomical details.
- **T1-contrast (T1ce):** Salient in enhancing tumor areas.
- **T2-weighted (T2):** Aids in viewing edema (swelling)..
- **FLAIR:** Highlights fluid and non-enhancing tumor parts.

All modalities are stored as a separate `.nii.gz` file for each patient.

3.4 Tumor Labels (Ground Truth)

The data include segmentation masks that classify brain tumors into 3 subregions:

- **Label 1 - Enhancing Tumor (ET):** Area of actively proliferating tumor.
- **Label 2 - Peritumoral Edema (ED):** Swelling of the tumor.
- **Label 3 - Necrotic & Non-enhancing Tumor Core (NCR/NET):** necrotic tumor tissue.

These labels are provided in `.nii.gz` format, with each voxel given a value of 0 (background) or 1-3 (tumor class)

3.5 Dataset Statistics

- **Total Training Patients:** 371
- **Total Validation Patients:** 127
- **Total Samples (Train + Val):** 498
- **MRI Modalities per Patient:** 4 (FLAIR, T1, T1ce, T2)
- **Example Image Shape:** (240, 240, 155) *(Width, Height, Depth)*

3.6 Folder Structure

The dataset is typically organized as follows:

```
BraTS2020/
 Training/
    BraTS20_0001/
       BraTS20_0001_flair.nii.gz
       BraTS20_0001_t1.nii.gz
       BraTS20_0001_t1ce.nii.gz
       BraTS20_0001_t2.nii.gz
       BraTS20_0001_seg.nii.gz  <-- Ground truth mask
 Validation/
 Test/
```

- **Training set:** MRI scans and segmentation labels are included in the training set
- **Validation/Test sets:** Used for testing (ground truth may not be provided for test).

3.7 Preprocessing Steps

Data Normalization: The MRI pixel values are normalized by dividing by the maximum intensity:

$$X = \frac{X}{\max(X)} \tag{1}$$

Resizing: Each slice of FLAIR, T1, and T1ce scans is resized to (`IMG_SIZE`, `IMG_SIZE`) using OpenCV (`cv2.resize`).

Channel Stacking: The MRI modalities (FLAIR, T1, T1ce) are stacked together to form a 3-channel input.

One-Hot Encoding: The segmentation labels are converted to a categorical format inside the DataGenerator, making them suitable for multi-class segmentation.

4 Results

4.1 Training Performance

The **3D U-Net** model for brain tumor segmentation was trained over five epochs on a **multi-GPU** setup using the **BraTS 2020** dataset. The training process showed steady improvements in segmentation accuracy, **Dice coefficients**, and loss minimization over epochs.

- **Epoch 1:** The original **Dice coefficient** of 0.7539 and loss of 0.6649 indicate a moderate level of segmentation ability.
- **Epoch 5:** The **Dice coefficient** improved to 0.9999, and the loss decreased to 0.0298, indicating excellent segmentation performance.

4.2 Validation Performance

The validation performance demonstrated steady improvements over epochs, with **Dice coefficients** for tumor subregions approaching 1.0000, indicating nearly perfect segmentation.

Table 2: Validation Performance Metrics Across Epochs

Epoch	Dice	Edema	Enhancing	Necrotic	Loss	Mean IoU	Precision	Sensitivity	Specificity
1	0.9985	0.9989	0.9991	0.9985	0.0651	0.7073	0.9905	0.9414	0.9969
2	0.9998	0.9999	0.9999	0.9998	0.0318	0.7842	0.9867	0.9867	0.9956
3	1.0000	1.0000	1.0000	1.0000	0.0290	0.8188	0.9871	0.9871	0.9957
4	0.9999	1.0000	1.0000	0.9999	0.0279	0.7580	0.9859	0.9859	0.9953
5	1.0000	1.0000	1.0000	1.0000	0.0286	0.8258	0.9868	0.9868	0.9956

4.3 Performance Metrics Analysis

- **Mean IoU:** Improved from **0.7073** (Epoch 1) to **0.8258** (Epoch 5), demonstrating enhanced segmentation accuracy.

- **Precision & Sensitivity:** Remained consistently high, ensuring the model effectively identifies tumor regions while minimizing false positives.
- **Specificity:** Stayed above **0.995**, confirming minimal misclassification of non-tumor regions.

4.4 Model Predictions (Segmentations)

The model predicts brain tumor areas and visualizes them well, emphasizing:

- **Necrotic Core (Dead Tumor Cells)**
- **Edema (Swelling Around Tumor)**
- **Enhancing Tumor (Active Tumor Region)**

How Results Are Displayed:

1. **Original FLAIR MRI Scan** – Raw input image.
2. **Ground Truth Segmentation** – Expert-annotated labels.
3. **Predicted Segmentation** – Model output (overlaid on FLAIR).
4. **Heatmaps & Binary Masks** – Aid in analyzing prediction confidence.

Visualization Details:

- **Heatmaps:** Show tumor region probability (red = high probability).
- **Binary Masks:** Provide thresholded segmentation with sharper boundaries.

4.5 Observations & Analysis

- The model converged quickly, reaching near-perfect segmentation by **Epoch 3**.
- Validation loss consistently dropped, affirming strong generalization to unseen data.
- **Dice coefficients** close to 1.0000 signify that the model effectively separates tumor subregions.

The above findings verify the efficacy of the **3D U-Net** model with **Spatial Attention** and **SpatialDropout3D** in multi-class brain tumor segmentation.

5 Conclusion

In this paper, we employed a **3D U-Net** network with **Spatial Attention** and **SpatialDropout3D** for segmenting brain tumors in the **BraTS 2020** dataset. Our model effectively separated different sub-regions of tumors, i.e., **necrotic/core**, **edema**, and **enhancing tumor**, from multi-modal MRI scans (**FLAIR**, **T1**, and **T1CE**). Training was enhanced through **multi-GPU MirroredStrategy**, and testing was carried out using a combination of **Dice Loss** and **Categorical Cross-Entropy Loss** for optimizing learning.

The segmentation performance was tested with visual summaries of ground truth and predicted masks, as well as quantitative evaluation using **Dice similarity scores**, **precision**, and **recall** metrics. Our results indicated good accuracy in segmentation, particularly for the **edema** and **enhancing tumor** regions, though some challenges were encountered in segmenting the **necrotic core** due to class imbalance and intensity variations.

Using fine-grained visualization techniques like **probability heatmaps** and **binary masks**, we were able to showcase the ability of the model to accurately localize tumors.

Further improvement can be made, however, with the addition of **post-processing techniques**, **ensemble learning**, and **data augmentation** methods for improved generalizability. Future work could also explore integrating **self-supervised learning** or **transformer-based architectures** to further enhance segmentation performance.

This work contributes to **deep learning**-based medical image segmentation, offering a potential aid for **computer-aided treatment planning** and **diagnosis** in **neuro-oncology**.

References

[1] D. N. Louis, P. Wesseling, K. Aldape, D. J. Brat, D. Capper, I. A. Cree, C. Eberhart, D. Figarella-Branger, M. Fouladi, G. N. Fuller *et al.*, "cimpact-now update 6: new entity and diagnostic principle recommendations of the cimpact-utrecht meeting on future cns tumor classification and grading," 2020.

[2] Z. Li, X. Wu, and X. Yang, "A multi brain tumor region segmentation model based on 3d u-net," *Applied Sciences*, vol. 13, no. 16, p. 9282, 2023.

[3] S. Jain and B. Sachdeva, "A systematic review on brain tumor detection using deep learning," in *AIP Conference Proceedings*, vol. 3168, no. 1. AIP Publishing, 2024.

[4] S. Bouhafra and H. El Bahi, "Deep learning approaches for brain tumor detection and classification using mri images (2020 to 2024): A systematic review," *Journal of Imaging Informatics in Medicine*, pp. 1–31, 2024.

[5] X. Dong, X. Mao, and J. Yao, "A novel cardiac image segmentation method using an optimized 3d u-net model," *Journal of Mechanics in Medicine and Biology*, vol. 23, no. 09, p. 2340102, 2023.

[6] S. Montaha, S. Azam, A. Rakibul Haque Rafid, M. Z. Hasan, and A. Karim, "Brain tumor segmentation from 3d mri scans using u-net," *SN Computer Science*, vol. 4, no. 4, p. 386, 2023.

[7] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using mri multi-modalities brain images," *Scientific reports*, vol. 11, no. 1, p. 10930, 2021.

[8] J. Colman, L. Zhang, W. Duan, and X. Ye, "Dr-unet104 for multimodal mri brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II 6*. Springer, 2021, pp. 410–419.

[9] S. Vijay, T. Guhan, K. Srinivasan, P. D. R. Vincent, and C.-Y. Chang, "Mri brain tumor segmentation using residual spatial pyramid pooling-powered 3d u-net," *Frontiers in public health*, vol. 11, p. 1091850, 2023.

[10] S. Sangui, T. Iqbal, P. C. Chandra, S. K. Ghosh, and A. Ghosh, "3d mri segmentation using u-net architecture for the detection of brain tumor," *Procedia Computer Science*, vol. 218, pp. 542–553, 2023.

[11] Z. Liu, L. Tong, L. Chen, Z. Jiang, F. Zhou, Q. Zhang, X. Zhang, Y. Jin, and H. Zhou, "Deep learning based brain tumor segmentation: a survey," *Complex & intelligent systems*, vol. 9, no. 1, pp. 1001–1026, 2023.

[12] C. M. Umarani, S. Gollagi, S. Allagi, K. Sambrekar, and S. B. Ankali, "Advancements in deep learning techniques for brain tumor segmentation: a survey," *Informatics in Medicine Unlocked*, p. 101576, 2024.

[13] H. Byeon, M. Al-Kubaisi, A. K. Dutta, F. Alghayadh, M. Soni, M. Bhende, V. Chunduri, K. Suresh Babu, and R. Jeet, "Brain tumor segmentation using neurotechnology enabled intelligence-cascaded u-net model," *Frontiers in Computational Neuroscience*, vol. 18, p. 1391025, 2024.

[14] J. Sun, R. Zhang, L. Guo, and J. Wang, "Multi-scale feature fusion and additive attention guide brain tumor mr image segmentation, j," *Image Graphics*, vol. 28, pp. 1157–1172, 2023.

[15] Z. Schwehr and S. Achanta, "Brain tumor segmentation based on deep learning, attention mechanisms, and energy-based uncertainty predictions," *Multimedia Tools and Applications*, pp. 1–20, 2025.

[16] R. A. Zeineldin, M. E. Karar, O. Burgert, and F. Mathis-Ullrich, "Multimodal cnn networks for brain tumor segmentation in mri: a brats 2022 challenge solution," in *International MICCAI Brainlesion Workshop*. Springer, 2022, pp. 127–137.

[17] E. K. Rutoh, Q. Z. Guang, N. Bahadar, R. Raza, and M. S. Hanif, "Gair-u-net: 3d guided attention inception residual u-net for brain tumor segmentation using multimodal mri images," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 6, p. 102086, 2024.

[18] Z. Mansur, J. Talukdar, T. P. Singh, and C. J. Kumar, "Deep learning-based brain tumor image analysis for segmentation," *SN Computer Science*, vol. 6, no. 1, pp. 1–18, 2025.

[19] B. AG, S. Srinivasan, M. P, S. K. Mathivanan, and M. A. Shah, "Robust brain tumor classification by fusion of deep learning and channel-wise attention mode approach," *BMC Medical Imaging*, vol. 24, no. 1, p. 147, 2024.

[20] Z. U. Abidin, R. A. Naqvi, A. Haider, H. S. Kim, D. Jeong, and S. W. Lee, "Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: A prospective survey," *Frontiers in Bioengineering and Biotechnology*, vol. 12, p. 1392807, 2024.

[21] A. B. Subba and A. K. Sunaniya, "Computationally optimized brain tumor classification using attention based googlenet-style cnn," *Expert Systems with Applications*, vol. 260, p. 125443, 2025.