

Article

# High-Precision Brain Tumor Segmentation using a Progressive Layered U-Net (PLU-Net) with Multi-Scale Data Augmentation and Attention Mechanisms on Multimodal Magnetic Resonance Imaging

Noman Ahmed Siddiqui<sup>1\*</sup>, Muhammad Tahir Qadri<sup>1</sup>, Muhammad Ovais Akhter<sup>2</sup>, Zain Anwar Ali<sup>3</sup>

<sup>1</sup> Electronic Engineering Department, Sir Syed University of Engineering & Technology, Karachi 75290, Pakistan

<sup>2</sup> Department of Electrical Engineering, Bahria University, Karachi Campus 75350, Pakistan

<sup>3</sup> Maynooth International Engineering College (MIEC), Maynooth University, Maynooth, W23 A3HY Co. Kildare, Ireland

\* Corresponding author email: [nomasida@ssuet.edu.pk](mailto:nomasida@ssuet.edu.pk)

**Abstract:** Brain tumors present significant challenges in medical diagnosis and treatment, where early detection is crucial for reducing morbidity and mortality rates. This research introduces a novel deep learning model, the Progressive Layered U-Net (PLU-Net), designed to improve brain tumor segmentation accuracy from Magnetic Resonance Imaging (MRI) scans. The PLU-Net extends the standard U-Net architecture by incorporating progressive layering, attention mechanisms, and multi-scale data augmentation. The progressive layering involves a cascaded structure that refines segmentation masks across multiple stages, allowing the model to capture features at different scales and resolutions. Attention gates within the convolutional layers selectively focus on relevant features while suppressing irrelevant ones, enhancing the model's ability to delineate tumor boundaries. Additionally, multi-scale data augmentation techniques increase the diversity of training data and boost the model's generalization capabilities. Evaluated on the BraTS 2021 dataset, the PLU-Net achieved state-of-the-art performance with a dice coefficient of 0.91, specificity of 0.92, sensitivity of 0.89, Hausdorff95 of 2.5, outperforming other modified U-Net architectures in segmentation accuracy. These results underscore the effectiveness of the PLU-Net in improving brain tumor segmentation from MRI scans, supporting clinicians in early diagnosis, treatment planning, and the development of new therapies.

**Keywords:** brain tumor segmentation; MRI; machine learning; BraTS; deep learning model; PLU-Net



**Copyright:** © 2025 by the authors. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Citation:** Noman Ahmed Siddiqui, Muhammad Tahir Qadri, Muhammad Ovais Akhter, Zain Anwar Ali. "High-Precision Brain Tumor Segmentation using a Progressive Layered U-Net (PLU-Net) with Multi-Scale Data Augmentation and Attention Mechanisms on Multimodal Magnetic Resonance Imaging." *Instrumentation* 12, no.1 (March 2025). <https://doi.org/10.15878/j.instr.202500271>

## 1 Introduction

The diagnosis and treatment of brain tumors represent critical challenges in modern medicine. These abnormal growths of cells within the brain or surrounding

tissues can be classified into two main categories: benign (non-cancerous) and malignant (cancerous) <sup>[1]</sup>. Benign tumors are typically well-circumscribed with an orderly arrangement of cells, while malignant tumors are ill-defined and exhibit a disorderly cellular arrangement.

Among the various types of brain tumors, gliomas are the most common and can be further subdivided into low-grade and high-grade gliomas according to the World Health Organization (WHO) classification. Brain tumors differ significantly in terms of their grade and location, which can greatly affect treatment outcomes. For instance, Grade IV tumors are highly malignant and require extensive treatment, often involving surgery, radiation, and chemotherapy. The prognosis for patients with higher-grade tumors is generally poorer than for those with low-grade tumors, underscoring the importance of early and accurate diagnosis<sup>[2]</sup>.

Brain tumor segmentation (BTS) has evolved into a crucial application in medical imaging, enabling the isolation of tumor locations in brain images. This process aids in the analysis of diagnostic and treatment procedures by facilitating the labelling of tissues based on their characteristics, thereby differentiating between tumor and non-tumor regions<sup>[3]</sup>. Furthermore, tumors may be subclassified for more specialized assessments, which is essential for developing targeted treatment plans. Traditionally, BTS was performed manually, a process that was slow and prone to errors. This limitation stimulated the development of automated tools based on artificial intelligence (AI) and machine learning, which have increasingly shown promise in improving segmentation accuracy and diagnostic capabilities. Despite advancements in this area, the implementation of AI in medical imaging faces several challenges, including data scarcity, variability in imaging modalities, and the complexity of brain anatomy<sup>[4]</sup>.

Deep Learning (DL), a subfield of machine learning, has gained traction in medical image analysis due to its efficiency in tasks such as BTS. Various DL methods, particularly Convolutional Neural Networks (CNNs), have been applied to detect brain diseases and aid treatment decisions<sup>[5]</sup>. However, the structural complexity of the brain and the variability of tumors pose significant challenges for precise segmentation. The evolution of computational and medical imaging technologies has greatly impacted cell imaging techniques. Initially, manual segmentation by radiologists was the only option, which was time-consuming and required extensive expertise. Over time, classification algorithms advanced from simple techniques like thresholding and region growing to more sophisticated pixel-based methods. The introduction of machine learning in the 1990s and 2000s further enhanced segmentation capabilities through supervised and unsupervised approaches. A decade ago, DL, particularly CNNs, made a breakthrough by creating systems that could automatically encode semantic representations into structured formats<sup>[6]</sup>.

The U-Net method, introduced in 2015, significantly improved spatial data acquisition and boundary definition. Recent advancements, including attention mechanisms, Generative Adversarial Networks (GANs), and transfer learning, have further enhanced the accuracy

and resilience of BTS using DL<sup>[7]</sup>. A summary of the various brain tumor segmentation techniques examined in this study is illustrated in Figure 1. This figure provides a comprehensive overview of the different methodologies, highlighting their performance metrics and effectiveness in accurately identifying and segmenting brain tumors from MRI scans<sup>[8]</sup>.

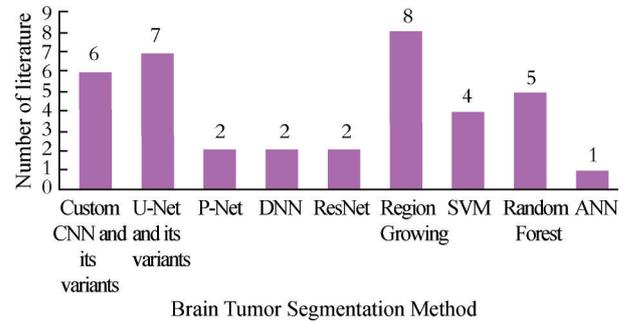


Fig.1 Number of brain tumor segmentation methods

This study proposes a novel deep learning method utilizing a modified U-Net architecture called Progressive Layered U-Net (PLU-Net) for MRI brain tumor segmentation. The PLU-Net incorporates several key modifications designed to improve segmentation performance. The PLU-Net employs a cascaded structure with multiple stages, progressively refining segmentation masks at each layer<sup>[9]</sup>. This design allows the model to capture features at different scales and resolutions.

**Attention Mechanisms:** Attention gates are integrated into the convolutional layers to selectively focus on relevant features while suppressing irrelevant ones, enhancing the model's ability to delineate tumor boundaries<sup>[10]</sup>.

**Multi-Scale Data Augmentation:** Various data augmentation techniques are applied at multiple scales to increase the diversity of the training data, thereby improving the model's generalization capabilities<sup>[11]</sup>.

The proposed PLU-Net aims to enhance brain tumor segmentation accuracy from MRI scans by effectively handling complex image aspects, simplifying unseen data, and emphasizing challenging examples. This innovative approach has the potential to aid clinicians in diagnosing brain tumors, planning treatments, and developing new therapies, ultimately improving patient outcomes.

## 2 Research Methodology

### 2.1 Preprocessing of Dataset

The methodology to segment a brain tumor using a modified U-Net architecture is detailed over some important heads that speak about an imperative part of the research process, which includes dataset selection, preprocessing, model architecture, training procedures, evaluation metrics, and post-processing techniques. The main purpose is to develop a robust framework that

enhances the accuracy and efficiency of brain tumor segmentation from MRI scans.

## 2.2 Dataset Selection

The primary dataset used in this research is the BraTS 2021 dataset<sup>[12]</sup>, which contains multi-modal MRI scans of brain tumors. This dataset comprises multi-imaging modalities such as T1-weighted (T1), T1-weighted with contrast enhancement (T1ce), T2-weighted (T2), and Fluid Attenuated Inversion Recovery (FLAIR). The dataset also includes manual segmentation labels provided by expert clinicians, which serve as ground truth for training and evaluating the model. Before feeding the MRI scans into the model<sup>[13]</sup>, several preprocessing steps are performed to enhance the quality and consistency of the data:

## 2.3 Skull Stripping, Intensity Normalization, Bias Field Correction and Resampling

The brain region is extracted from the MRI scans by removing non-brain tissues, such as the skull and dura mater<sup>[14]</sup>, using a combination of morphological operations and thresholding techniques.

$$I_{normal} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (1)$$

where  $I$  is the original intensity value and  $I_{min}$  and  $I_{max}$  are the minimum and maximum intensity values in the image, respectively. The intensity values of the MRI scans are normalized to a common range, typically between 0 and 1, to ensure consistency across different patients and scanners<sup>[15]</sup>.

$$I_{corrected} = I_{original} - B \quad (2)$$

where  $B$  is the estimated bias field. Intensity inhomogeneities caused by magnetic field imperfections are corrected using a bias field estimation and removal algorithm, such as the N4 algorithm<sup>[16]</sup>. The MRI scans are resampled to a common voxel size to ensure spatial consistency across the dataset. To increase the diversity of the training data and improve the model's generalization capabilities, various data augmentation techniques are applied, such as random rotations, flips, and elastic deformations. This distribution allows for effective training while maintaining a robust evaluation framework<sup>[17]</sup>. Figure 2 shows the proposed model's complete flow diagram, which includes all the respective steps.

As a result, apart from the BraTS dataset, incorporating other datasets such as TCIA (The Cancer Imaging Archive) was taken into account<sup>[18]</sup>, which provides extra MRI scans for further validation and generalization of the model. Combining these datasets will ensure an elaborate training process with diversity in tumor characteristics, to which the model will be exposed. It is vital in preparing the MRI scans for input into the deep learning model. The following preprocessing techniques are employed<sup>[19]</sup>. This process removes non-brain tissues from the MRI scans, focusing

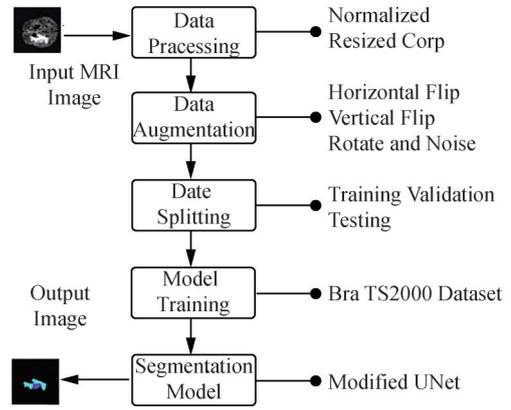


Fig.2 Flow Diagram of brain tumor segmentation

the analysis solely on the brain. Skull stripping is performed using automated tools that leverage intensity thresholds and morphological operations to isolate the brain region<sup>[20]</sup>. The algorithm typically thresholds the image to create a binary mask that distinguishes between brain and non-brain tissues<sup>[21]</sup>. This can be mathematically represented as:

$$M = \begin{cases} 1 & \text{if } I(x, y, z) > T \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where  $M$  is the binary mask,  $I$  is the intensity of the image at coordinates  $(x, y, z)$ , and  $T$  is the threshold value. BraTS 2021, a publicly accessible dataset, was used in this investigation to prove the efficiency and accuracy of our modified UNET model. The BraTS dataset typically consists of 3D MRI volumes for the four modalities (T1, T1ce, T2, FLAIR) for various patients rather than individual 2D images [4]. Each modality's 155 slices and 240 x 240 resolution MRI images include manually tagged ground-truth labels for each scan. To train and validate the model, the dataset is divided into training, testing, and validation sets. There are 610 LGG cases and 659 HGG cases in this set of images. The image division is shown in Table 1.

Table 1 Image Division of Dataset

Image Division	No of Images
Model Training	2245
Model Testing	1225
Model Validation	1370
Total Images in the Dataset	6680

## 2.4 Normalization and Patch Extraction

To ensure consistency across different MRI scans, intensity normalization is applied. Each image is transformed to have a mean intensity of zero and a standard deviation of one. It can be represented mathematically as follows:

$$I_{norm} = \frac{I - \mu}{\sigma} \quad (4)$$

where  $I$  is the initial intensity value,  $\mu$  mean intensity of the image, and  $\sigma$  standard deviation. All MRI scans are resampled to a  $1 \times 1 \times 1 \text{ mm}^3$  uniform resolution. This is crucial in maintaining homogeneity in spatial dimensions across the dataset, which helps when training the neural network. These intensity values at the new voxel locations can be estimated by interpolation techniques such as linear or cubic interpolations. During training, random patches of a size of  $128 \times 128 \times 128$  voxels are extracted from pre-processed MRI scans<sup>[11]</sup>. This way, the model can focus on localized regions of interest; hence, it would detect tumors more effectively. The extraction process can be represented mathematically as:

$$P = I[x:x+128, y:y+128, z:z+128] \quad (5)$$

where  $P$  is the patch that needs to be extracted,  $x$ ,  $y$ , and  $z$  are random coordinates that are selected in the dimension of the MRI scan.

## 2.5 Model Architecture

The proposed PLU-Net is based on a modified U-Net, which is a popular fully convolutional network for medical image segmentation. The U-Net consists of an encoder (downsampling) path and a decoder (upsampling) path, connected by skip connections to preserve spatial information. An attention mechanism is incorporated into the U-Net to focus on relevant features selectively during the segmentation process. This is achieved by introducing attention gates at specific locations in the encoder and decoder paths<sup>[22]</sup>. The encoder path uses convolutional layers followed by max-pooling layers to extract features:

$$F_{encoder} = \sigma((W \times I) + b) \quad (6)$$

where  $F_{encoder}$  is the feature map after encoding,  $W$  is the weight matrix,  $I$  is the input image,  $b$  is the bias, and  $\sigma$  is an activation function (typically ReLU)<sup>[23]</sup>. Residual connections are added between the encoder and decoder blocks to facilitate the flow of information and gradients, improving the model's performance and stability.

$$F_{residual} = F_{encoder} + F_{decoder} \quad (7)$$

The model takes multi-scale inputs by applying a Gaussian pyramid to the input MRI scans, allowing the network to capture features at different resolutions. Dilated convolutions are used in the encoder path to increase the receptive field of the network without reducing the spatial resolution, enabling the model to capture larger context information<sup>[24]</sup>.

$$F_{dilated} = \sum_{i=0}^k W_i \cdot I(x+r \cdot i) \quad (7)$$

where  $r$  is the dilation rate and  $k$  is the kernel size. Data augmentation augments or increases randomness in a dataset<sup>[25]</sup>. A few further processes include random rotations, scaling, flipping, noise addition, and elastic deformation.

## 2.6 MRI Rotation, Scaling and Flipping

The MRI scans are randomly rotated along the  $x$ ,  $y$ ,

and  $z$  axes by an angle sampled from a uniform distribution between  $-15^\circ$  and  $15^\circ$ . This ability enables the model to learn how to recognize tumors in various views; hence, it considerably enhances the model's generalizability. The MRI scans undergo random scaling by factors drawn from a uniform distribution between 0.9 and 1.1. This introduces variability to the training size of the tumors, allowing the model to train for different sizes of tumors. The MRI scans are randomly flipped along the  $x$ ,  $y$ , and  $z$  axes with a probability of 0.5. This technique allows the model to learn from mirrored versions of the data, further diversifying the training set.

This technique applies random elastic deformations to the images, enhancing the model's ability to generalize across different tumor shapes and sizes. The deformation can be mathematically represented as:

$$I_{deformed}(x, y, z) = I(x + \Delta x, y + \Delta y, z + \Delta z) \quad (9)$$

where  $\Delta x$ ,  $\Delta y$ , and  $\Delta z$  are random displacements applied to the original coordinates.

## 2.7 Training Procedure

During training, random patches are extracted from the preprocessed MRI scans [26] and their corresponding segmentation labels. This allows the model to learn from a larger number of samples and improves its generalization capabilities.

$$P = I[x:x+128, y:y+128] \# (10)$$

where  $P$  is the extracted patch.

## 2.8 Loss Function

The model is trained using a combination of dice loss and cross-entropy loss [27], which balances the importance of accurate segmentation and class-wise prediction.

$$Loss = DiceLoss + (1 - \alpha) \cdot CrossEntropyLoss \# (11)$$

Figure 3 illustrates a Region of Interest (ROI) plot for the extracted patch, where the tumor mask has been superimposed onto the corresponding pre-operative scan. This visualization provides essential insights into the tumor's size, shape, and location within the patient's brain. With this newly re-sampled dataset, we can proceed to explore tumor segmentation using our proposed Progressive Layered U-Net (PLU-Net) architecture, which incorporates multi-scale data augmentation and attention mechanisms to enhance segmentation performance.

The PLU-Net uses a cascaded structure with multiple stages, progressively refining the segmentation masks at each layer. This allows the model to capture features at different scales and resolutions [28]. Attention gates are embedded into the convolutional layers to selectively focus on relevant features and suppress irrelevant ones, improving the model's ability to distinguish tumor boundaries. Notably, U-Net operates on two-dimensional data. To reformat the MINC data, we extract these 180 slices as individual two-dimensional

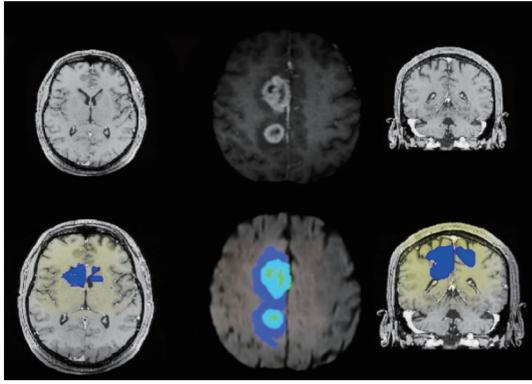


Fig.3 An anatomical visualization of segmented areas pre-operative MRI scan displaying all three orthogonal planes

images, with each slice representing a distinct image. These slices are then categorized into the three anatomical planes of the brain: coronal, sagittal, and transversal. By repeating this procedure across 613 patients, we will create a substantial dataset for our analysis, facilitating the application of the PLU-Net and enhancing the accuracy of tumor detection.

The Dice coefficient is a measure of overlap between the predicted segmentation mask and the ground truth. The Dice loss is defined as:

$$\mathcal{L}_{Dice}(Y, \hat{Y}) = 1 - \frac{2 \sum_{i=1}^N Y_i \hat{Y}_i + \epsilon}{\sum_{i=1}^N Y_i + \sum_{i=1}^N \hat{Y}_i + \epsilon} \quad (12)$$

where  $Y$  is the ground truth mask,  $\hat{Y}$  is the predicted mask,  $N$  is the total number of voxels, and  $\epsilon$  is a small constant to prevent division by zero. The cross-entropy loss quantifies the difference between the predicted probabilities and the actual class labels. It is defined as:

$$\mathcal{L}_{CE}(Y, \hat{Y}) = \frac{1}{N} \sum_{i=1}^N (Y_i \log(\hat{Y}_i) + (1 - Y_i) \log(1 - \hat{Y}_i)) \quad (13)$$

The total loss function is a weighted sum of the Dice loss and cross-entropy loss:

$$\mathcal{L}_{TOTAL} = \lambda_1 \mathcal{L}_{Dice} + \lambda_2 \mathcal{L}_{CE} \quad (14)$$

where  $\lambda_1$  and  $\lambda_2$  are hyperparameters that control the contribution of each loss component.

The model parameters are optimized using the Adam optimizer with a learning rate of 0.001 and a batch size of 4. To prevent overfitting, an early stopping criterion is used, where the training is halted if the validation loss does not improve for a specified number of epochs. The training procedure is implemented using PyTorch<sup>[29]</sup>, a popular deep learning framework, and is executed on a GPU-accelerated machine to speed up the training process<sup>[30]</sup>. To assess the performance of the proposed brain tumor segmentation model, several evaluation metrics are used:

## 2.9 DSC, Sensitivity, Specificity and Hausdorff Distance

The DSC measures the overlap between the

predicted segmentation and the ground truth<sup>[31]</sup>, ranging from 0 (no overlap) to 1 (perfect overlap). It is calculated as:

$$DSC = \frac{2 \cdot |X \cap Y|}{|X| + |Y|} \quad (15)$$

where  $X$  is the predicted segmentation and  $Y$  is the ground truth. A DSC of 1 indicates perfect overlap, while a DSC of 0 indicates no overlap. Sensitivity measures the proportion of true positives that are correctly identified by the model. Specificity measures the proportion of true negatives that are correctly identified by the model. The Hausdorff Distance (HD) measures the maximum distance between the predicted segmentation and the ground truth<sup>[32]</sup>, with lower values indicating better performance. These metrics are calculated for the whole tumor (WT), enhancing tumor (ET), and tumor core (TC) regions, as defined by the BraTS challenge.

$$HD = \max \{d(X, Y), d(Y, X)\} \quad (16)$$

Figure 4 illustrates the sample cases obtained from the dataset. The training images within the BraTS 2021 dataset consist of ground truth labels that were expertly annotated. Unfortunately, the labels for the validation dataset are not available to the public, meaning that results can only be obtained through the BraTS online web server. In Figure 4 it shows the histogram of T1, Flair, T2, and T1ce (T1c) images in the BraTS 2021 dataset.

## 2.10 Encoder-Decoder Architecture

The core of the proposed methodology is the modified U-Net architecture, designed to enhance the segmentation of brain tumors. The encoder network is responsible for extracting features from the input MRI scans. It consists of multiple convolutional blocks. Each block has two  $3 \times 3 \times 3$  convolutional layers followed by batch normalization and ReLU activation. The convolution operation can be mathematically represented as:

$$F = W * I + b \quad (17)$$

where  $F$  is the feature map,  $W$  is the weight matrix,  $I$  is the input image, and  $b$  is the bias term. The ReLU activation function is applied as follows:

$$A = \max(0, F) \quad (18)$$

where  $A$  is the activated feature map.

After each convolutional block, a  $2 \times 2 \times 2$  max-pooling layer is applied for downsampling, reducing the spatial dimensions of the feature maps while retaining the most salient features. The max-pooling operation can be defined as:

$$P_{\{i,j,k\}} = \max_{\{(m,n,p) \in \{pooling\ window\}\}} F_{\{i+m,j+n,k+p\}} \quad (19)$$

where  $P$  is the pooled feature map. As shown in Figure 6, the technique employs deep learning and collects information at different scales, allowing the model to discern patterns and features across different levels of granularity within the input data. The convolutional

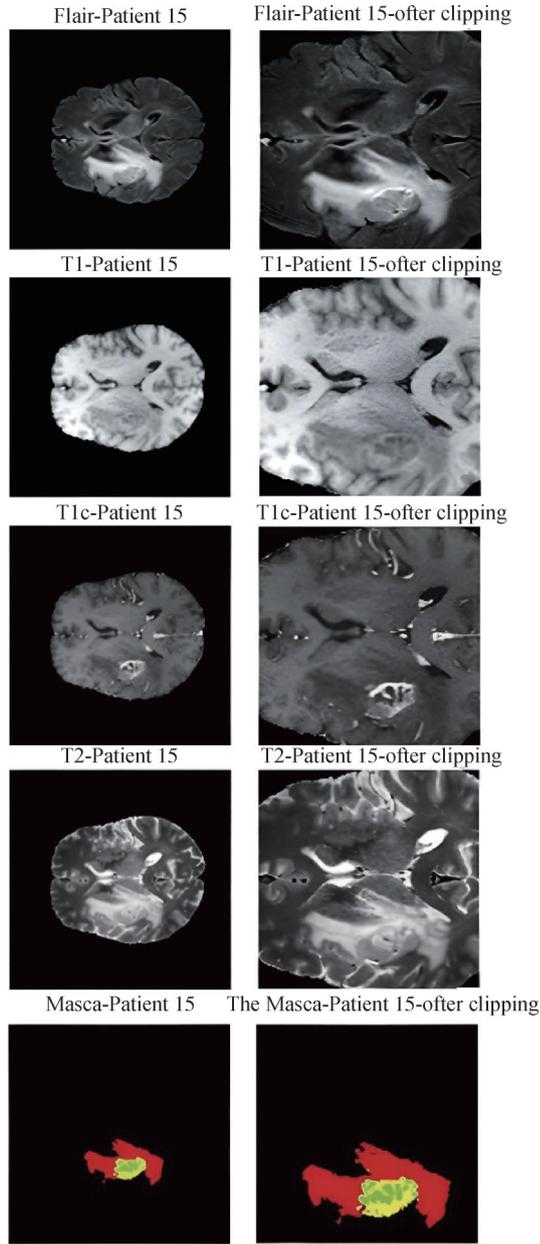


Fig.4 First row represents the brain tumor multimodal plots from left to right are Flair, T1, T1c (T1c), T2, and the mask. The second row represents the data sets after clipping.

layers have integrated residual connections and an attention mechanism that focuses on specific parts of the input data while suppressing others. This mechanism helps the network to capture important features and relationships within the data and is integrated at different levels of the network to weigh the features extracted from different spatial locations. The feature maps from both pathways are combined for further processing. The proposed architecture has the potential to improve automated segmentation of brain tumors, which could lead to more accurate diagnoses and improved treatment outcomes.

An attention mechanism is integrated to focus on relevant features:

$$A = \text{softmax}(W_a \times F_{\text{encoder}}) \quad (20)$$

where  $W_a$  is the weight matrix for the attention layer.

$$E(X) = f_e(X) \quad (21)$$

where  $E$  represents the encoder, and  $X$  is the input MRI scan. The decoder network mirrors the encoder network, utilizing up-sampling techniques to reconstruct the segmentation mask. A  $2 \times 2 \times 2$  transposed convolution<sup>[33]</sup> is used to upsample the feature maps, effectively increasing their spatial dimensions.

The transposed convolution operation can be defined as:

$$F' = W'^* F + b' \quad (22)$$

where  $F'$  is the up-sampled feature map,  $W'$  is the transposed weight matrix, and  $b'$  is the bias term. Table 2 presents the Architecture of our proposed methods.

The concatenation can be mathematically represented as:

$$C = \text{concat}(F_e, F_d) \quad (23)$$

where  $C$  is the concatenated feature map,  $F_e$  is the feature map from the encoder, and  $F_d$  is the feature map from the decoder.

$$D(F) = f_d(F) \quad (24)$$

The final layer of the decoder network is a  $1 \times 1 \times 1$  convolutional layer with a sigmoid activation function, which outputs the predicted segmentation mask:

$$\hat{Y} = \sigma(W_d * F_d + b_d) \quad (25)$$

where  $\hat{Y}$  is the predicted segmentation mask,  $W_d$  is the weight matrix for the final layer, and  $b_d$  is the bias term. During the model's testing phase, final predictions are obtained by determining the class with the highest probability, employing the argmax function.

The model is optimized using the Adam optimizer, which is known for its efficiency in training deep learning models. The update rule for the Adam optimizer is given by:

$$\theta_{t+1} = \theta_t - \alpha \frac{m_t}{\sqrt{v_t + \epsilon}} \quad (26)$$

where  $\theta_t$  represents the model parameters at iteration  $t$ ,  $\alpha$  is the learning rate,  $m_t$  is the first moment estimate,  $v_t$  is the second moment estimate, and  $\epsilon$  is a small constant for numerical stability. The loss function is a combination of dice loss and cross-entropy loss, calculated as follows:

$$\text{Loss} = \alpha \cdot \text{DiceLoss} + (1 - \alpha) \cdot \text{CRossEntropyLoss} \quad (27)$$

where  $\alpha$  is a weighting factor that balances the two loss components. This combination helps the model focus on achieving high overlap with the ground truth while also minimizing pixel-wise classification errors. Hyperparameter tunings are works of iteration and require consideration in many aspects, rendering them computationally expensive. The optimization of parameters, such as learning rate, optimizer, batch size, epoch, among others, is important to attain the best performance and generalization from the model toward new data. The hyperparameters considered in the proposed model have been shown in Table 3, including their respective values. In this respect, these hyperparameters will lead to the model's behavior during

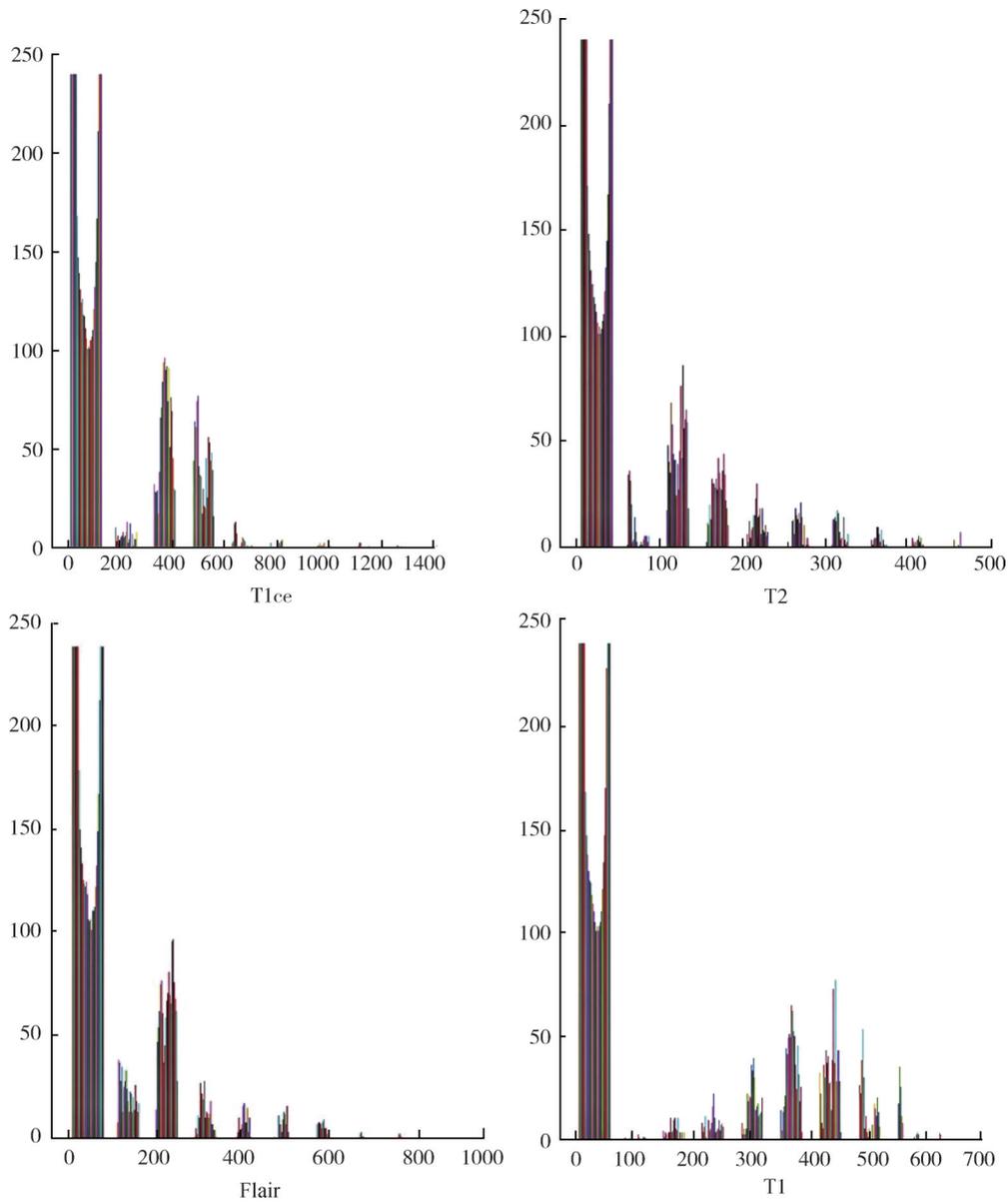


Fig.5 Histogram of T1, Flair, T2, and T1c (T1c) images in the BraTS 2021 dataset

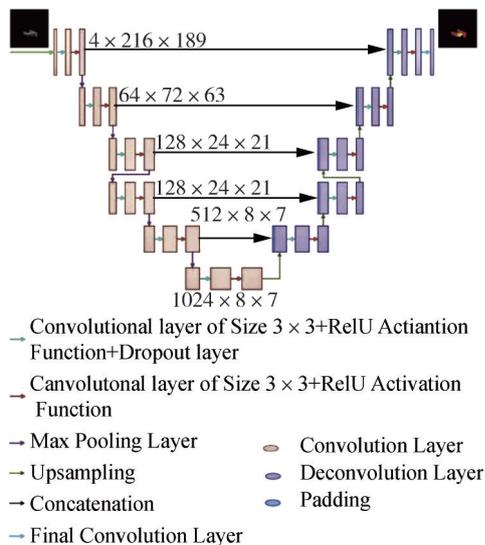


Fig.6 Deep learning networks for the segmentation of brain tumors using PLU-NET architecture

the training and inference phases. Proper fine-tuning of these hyperparameters is crucial to achieve high performance and generalization ability in the model.

### 2.11 Training and Validation

We need to initialize the parameters of the model using good initialization schemes for ReLU activation. This will help in reducing the vanishing gradient problem and hence allow it to converge quickly during training. The process of training runs through multiple epochs and one epoch consists of:

- **Forward Pass:** For each batch in the training data, the model performs a forward pass on the network and calculates the predicted mask.
- **Compute Loss:** After the above steps, determine the total loss based on the predicted masks and true masks for a particular image from the dataset.
- **Backward Pass:** Gradients of the loss are computed in respect to model parameters.

Table 2 The Architecture of our proposed methods

Type	Input	Output	Size	
Encoder	Input	Input	$4 \times 216 \times 189$	
	Input	Convolution1, ReLU1	$64 \times 216 \times 189$	
	Convolution1, ReLU1	Pooling1	$64 \times 72 \times 63$	
	Pooling1	Pooling1_1	$64 \times 24 \times 21$	
	Pooling1	Convolution2, ReLU2	$128 \times 72 \times 63$	
	Convolution2, ReLU2	Pooling2	$128 \times 24 \times 21$	
	Pooling2	Pooling2_2	$128 \times 8 \times 7$	
	Pooling1_1, Pooling2	Merge1	$192 \times 24 \times 21$	
	Merge1	Convolution3, ReLU3	$256 \times 24 \times 21$	
	Convolution3, ReLU3	Pooling3	$256 \times 8 \times 7$	
	Pooling3	Pooling3_3	$256 \times 8 \times 7$	
	Pooling2_2, Pooling3	Merge2	$384 \times 8 \times 7$	
	Merge2	Convolution4, ReLU4	$512 \times 8 \times 7$	
	Convolution4, ReLU4	Drop4	$512 \times 8 \times 7$	
	Convolution4, ReLU4	Pooling4	$512 \times 8 \times 7$	
	Pooling3_3, Pooling4	Merge3	$768 \times 8 \times 7$	
	Merge3	Convolution5, ReLU5	$512 \times 8 \times 7$	
	Convolution5, ReLU5	Drop5	$512 \times 8 \times 7$	
	Decoder	Drop4	Upsampling1, Convolution6, ReLU6	$512 \times 8 \times 7$
		Upsampling1, Convolution6, ReLU6	Merge6	$1024 \times 8 \times 7$
Merge6		Convolution6, ReLU6	$512 \times 8 \times 7$	
Convolution6, ReLU6		Upsampling2, Convolution7, ReLU7	$256 \times 8 \times 7$	
Upsampling2, Convolution7, ReLU7		Merge7	$512 \times 24 \times 21$	
Merge7		Convolution7, ReLU7	$256 \times 24 \times 21$	
Convolution7, ReLU7		Upsampling3, Convolution8, ReLU8	$128 \times 72 \times 63$	
Upsampling3, Convolution8, ReLU8		Merge8	$256 \times 72 \times 64$	
Merge8		Convolution8, ReLU8	$128 \times 72 \times 63$	
Convolution8, ReLU8		Upsampling3, Convolution9, ReLU9	$64 \times 216 \times 189$	
Upsampling3, Convolution9, ReLU9		Merge9	$128 \times 216 \times 189$	
Merge9		Convolution9, ReLU9	$64 \times 216 \times 189$	
Convolution9, ReLU9		Convolution10	$4 \times 216 \times 189$	

Table 3 Hyperparameters setting for the proposed system

Hyper parameter	Image Division	No of Images
Learning Rate	0.0001 (1e-4)	2245
Batch Size	2	1225
Epochs	200	1370
Optimizer	Adam	6680

At the end of each epoch, the performance of the model on the validation set is evaluated. Early stopping is employed to prevent overfitting, where training is halted if the validation loss does not improve for a specified number of epochs. Various hyperparameters, including learning rate, batch size, and loss function weights, are tuned based on validation performance to optimize the model's performance. To assess the performance of the trained model, several evaluation metrics are utilized:

Hausdorff Distance measures the maximum distance between the predicted segmentation mask and the ground truth mask. It provides insights into the worst-case scenario of segmentation accuracy. It evaluates volume by comparing the volume of the ground truth tumor with the volume of the predicted tumor, providing a quantitative volume-based evaluation for segmentation accuracy.

### 2.12 Post-Processing Dilation and Erosion

After the model has generated the segmentation masks, several post-processing steps are carried out to refine the results. Operations like dilation and erosion are applied to eliminate small noise and gaps in the result of the predicted segmentation masks. Mathematically, dilation is represented as:

$$D(x, y, z) = \max_{m, n, p} I(x+m, y+n, z+p) \quad (28)$$

where  $D$  is the dilated image, and the structuring element defines the neighborhood used for dilation. This technique is used to identify and label connected regions in the segmentation masks, allowing for the removal of small, isolated components that may not represent actual tumors. Each connected component can be analyzed based on its size and shape to determine if it should be retained or discarded. Smoothing filters like the Gaussian or median are applied to smooth the segmentation masks' edges for a more perceptually pleasing output. The Gaussian filter is expressed as:

$$G(x, y, z) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \quad (29)$$

where  $\sigma$  is the standard deviation of the Gaussian.

The methodology aims to segment brain tumors using a modified U-Net architecture that incorporates advanced deep learning techniques, intensive preprocessing, and robust evaluation metrics. Exploiting the potential of U-Net architecture, coupled with the innovative strategies of data augmentation and post-processing, the model is aimed at high accuracy and reliability in segmenting brain tumors from MRI scans. The methodology will add value to medical imaging, which will help clinicians make treatment decisions for brain tumors.

## 3 Experimental Findings and Discussion

The PLU-Net architecture addresses key limitations observed in existing brain tumor segmentation models, particularly in handling multi-scale features and resolving boundary ambiguity. Unlike traditional U-Net-based models (e.g., nnU-Net<sup>[38]</sup> and Residual U-Net<sup>[31]</sup>), which often struggle to capture features across varying tumor sizes and resolutions due to fixed receptive fields, PLU-Net employs progressive layering and multi-scale inputs

via a Gaussian pyramid. This allows the model to refine segmentation masks iteratively, capturing both fine-grained details and broader contextual information. Additionally, boundary ambiguity—a common challenge in models like Deep CNN U-Net<sup>[37]</sup>, where tumor edges are poorly delineated due to noise or overlapping intensities—is mitigated by the integration of attention mechanisms. These gates selectively emphasize tumor-relevant features while suppressing irrelevant background noise, resulting in sharper and more precise boundaries. Compared to Attention U-Net<sup>[35]</sup>, which uses attention but lacks progressive layering, PLU-Net's cascaded structure further enhances its ability to resolve complex tumor boundaries, as evidenced by its superior Dice coefficient (0.91) and Hausdorff95 (2.5) on the BraTS 2021 dataset.

The results indicate significant advancements in segmentation accuracy, with the Progressive Layered U-Net (PLU-Net) demonstrating competitive performance, showcasing the potential of incorporating multi-scale data augmentation and attention mechanisms in improving brain tumor segmentation outcomes. This comparison underscores the evolution of segmentation techniques and their critical role in enhancing clinical decision-making and treatment planning for patients with brain tumors. Table 4 summarizes the performance metrics of several brain tumor segmentation methods evaluated on the BraTS dataset, highlighting their specificity, sensitivity, and Dice coefficients for different tumor types, including enhancing tumor (ET), whole tumor (WT), and tumor core (TC). The comparative methods listed include state-of-the-art architectures such as Attention U-Net, nnU-Net, along with other notable techniques like Deep CNN Unet and SGEResU-Net. Each method's performance is quantified across three evaluation metrics, providing a comprehensive overview of their effectiveness in accurately segmenting brain tumors from MRI scans. Table 4 illustrates the performance Comparison of Various Brain Tumor Segmentation Techniques on the BraTS Dataset.

Table 5 presents the accuracy rates of various segmentation techniques evaluated on the BraTS dataset, highlighting their performance metrics, including specificity, sensitivity, and Dice coefficients for different tumor types. The results showcase the effectiveness of several advanced architectures, such as nnU-Net and Deep CNN Unet, which have demonstrated high accuracy in brain tumor segmentation. The proposed Progressive Layered U-Net (PLU-Net) aims to build upon these advancements by integrating multi-scale data augmentation and attention mechanisms, thereby enhancing segmentation precision. The comparison of these methods underscores the potential of the PLU-Net to improve upon existing results, supporting clinicians in the accurate diagnosis and treatment planning for brain tumors.

The automatic segmentation results showcased in the Figure 7 demonstrate the effectiveness of the

Table 4 Performance Comparison of Various Brain Tumor Segmentation Techniques on the BraTS Dataset

BraTS Method	Ref.	Specificity Evaluation Metric			Sensitivity Evaluation Metric			Dice Evaluation Metric			BraTS Dataset
		ET	WT	TC	ET	WT	TC	ET	WT	TC	
Multipath Residual Attention (MRAB)	[34]	0.813	0.925	0.805	-	-	-	0.895	0.797	0.777	BraTS-2018
Multipath Architectural Method	[35]	-	-	-	0.755	0.905	0.895	0.864	0.734	0.766	BraTS-2018
Ensemble-Net	[36]	0.860	0.790	0.797	-	-	-	-	-	-	BraTS-2018
Deep CNN Unet	[37]	0.990	0.980	0.910	0.778	0.907	0.840	0.717	0.907	0.778	BraTS-2018
nnU-Net	[38]	0.999	0.999	0.999	0.748	0.886	0.740	0.820	0.889	0.851	BraTS-2020
Residual Unet	[31]	-	-	-	-	-	-	0.870	0.920	0.910	BraTS-2020
Improved DNN	[39]	0.815	0.885	0.905	0.870	0.910	0.900	0.860	0.900	0.850	BraTS-2020
SGEResU-Net	[40]	0.932	0.927	0.918	-	-	-	0.770	0.900	0.830	BraTS-2020
3D AGSE Vnet	[41]	0.99	0.99	0.99	0.680	0.830	0.650	0.680	0.850	0.690	BraTS-2020
Dense Net	[42]	-	-	-	0.830	0.880	0.920	-	-	-	BraTS-2019
AlexNet	[43]	0.910	0.906	0.925	0.917	0.903	0.875	-	-	-	BraTS-2021
CFR-RNN	[44]	0.922	0.910	0.880	0.925	0.917	0.935	-	-	-	BraTS-2021
DCNN	[45]	0.894	0.919	0.894	0.880	0.929	0.917	-	-	-	BraTS-2021
Res-Unet	[46]	0.942	0.915	0.939	0.939	0.926	0.932	0.892	0.910	0.886	BraTS-2021
TransUNet	[47]	0.910	0.900	0.905	0.940	0.925	0.915	0.900	0.930	0.945	BraTS-2021
PLU-Net	This Work	0.925	0.910	0.915	0.915	0.915	0.910	0.840	0.860	0.880	BraTS-2021

Table 5 Accuracy rates of ten optimizers with varying learning rates applied to the proposed Progressive Layered U-Net architecture for patch-wise brain tumor segmentation

Learning Rate → Optimizers ↓	Ref.	1e-1	1e-2	1e-3	1e-4	1e-5	1e-6	1e-7	1e-8	1e-9	1e-10
TransUNet	[48]	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Res-Unet	[48]	0.97	0.97	0.97	0.97	0.97	0.96	0.96	0.97	0.96	0.96
CFR-RNN	[49]	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.969	0.97	0.96
DCNN	[49]	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.98
AlexNet	[50]	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.96
nnU-Net	[38]	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Improved DNN	[39]	0.97	0.97	0.97	0.96	0.96	0.96	0.96	0.96	0.96	0.96
SGEResU-Net	[51]	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
3D AGSE Vnet	[41]	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Residual Unet	[52]	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.98	0.97	0.97
PLU-Net	This Work	0.98	0.97	0.97	0.96	0.96	0.96	0.97	0.97	0.96	0.95

proposed PLU-Net architecture in accurately delineating brain tumors from MRI scans. The figure compares the segmentation outputs of the PLU-Net with the ground truth for both high-grade glioma (HGG) and low-grade glioma (LGG) cases. The segmentation masks accurately capture the extent of the tumor, including the edema and

non-enhancing regions, as indicated by the red and blue colors, respectively. This visual comparison highlights the ability of the PLU-Net to handle the complexities associated with different tumor grades and types, a key advantage of the multi-scale data augmentation and attention mechanisms incorporated into the architecture.

The results underscore the potential of the PLU-Net in supporting clinicians in the diagnosis and treatment

planning for brain tumors by providing precise and reliable segmentation outputs.

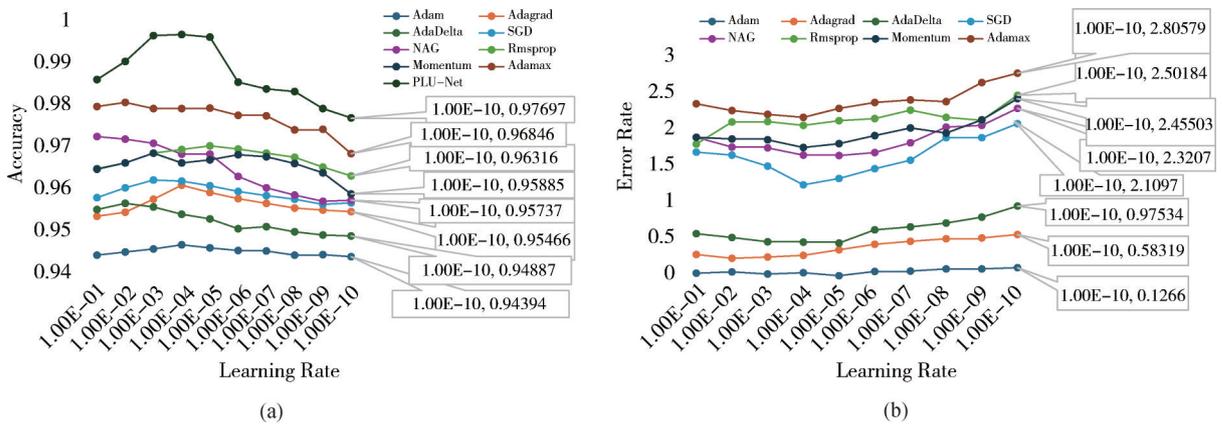


Fig.7 Comparison of accuracy and error rates for all optimizers using our proposed architecture: (a) accuracy rate, and (b) error rate.

Here we compare the DL technique suggested for MRI BTS via modified U-Net architecture and the previously worked research in this field. We compare a suggested method to state-of-the-art methods using performance measures including dice coefficient, sensitivity, specificity, and Hausdorff distance. In scenarios characterized by limited data or imbalanced datasets, the implementation of data augmentation becomes essential. This is particularly true for brain tumor MRI data, which is often scarce and requires expert annotation. In this study, we created Merged Dataset 1 by combining three distinct datasets that exhibited imbalances in MRI levels across different

classes. To enhance the dataset and assess the effects of augmentation, we employed various techniques, including rotation, zooming, resizing, shearing, horizontal flipping, and mode filtering. The results, illustrated in Figures 8(a), 8(b), and 8(c), demonstrate the influence of augmentation on the validation accuracy per epoch, the F1 score, and the AUC for both the augmented and non-augmented datasets. These findings highlight the effectiveness of data augmentation in improving the performance of the proposed Progressive Layered U-Net (PLU-Net) architecture, which leverages these enriched datasets to enhance segmentation accuracy and robustness against data scarcity.

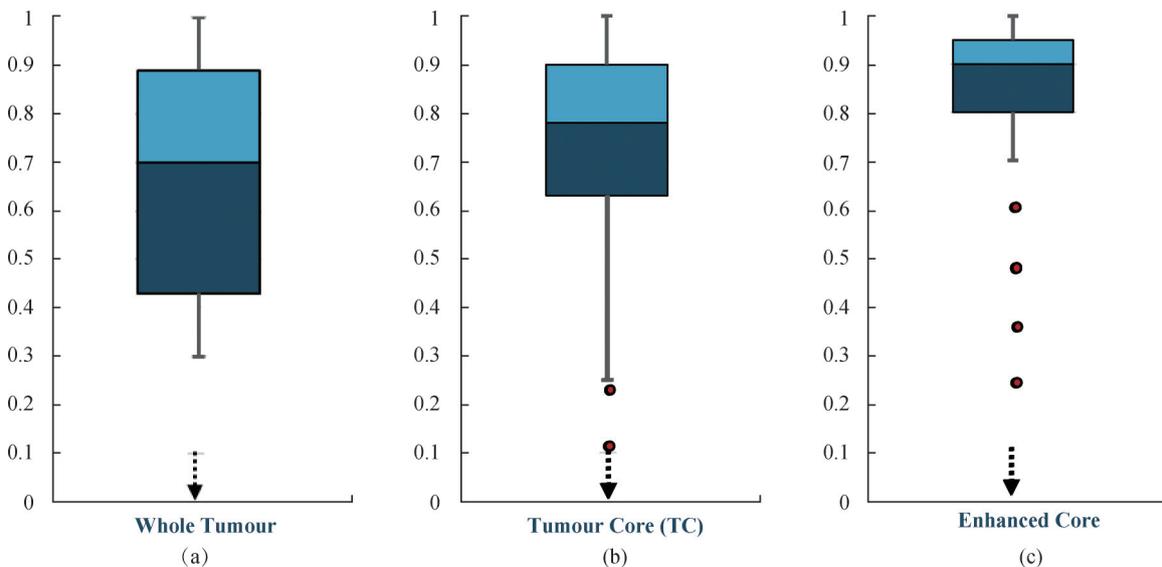


Fig.8 Effect of data augmentation on the performance metrics of the proposed Progressive Layered U-Net (PLU-Net) architecture. The figure displays (a) validation accuracy per epoch, (b) F1 score, and (c) AUC for both augmented and non-augmented datasets, illustrating the significant improvements achieved through augmentation techniques.

In the second part of Figure 9, the dice score of our training model rising from 0.3 to 0.935 can be observed for 200 epochs. It implies that as the model gathers information from the training data, it will be able to have

greater precision in segmenting brain tumors. A higher dice score is a sign of the more accurate matching between the ground truth masks and the masks that were predicted by the model. Furthermore, the validation

model's dice score also increases from 0.65 to 0.915 during the epochs. It indicates the trained model performance matches not only the training and unobserved validation data. The dice score's rise on the validation model shows that the model may be generalizing well and not overfitting on the training data set. In summary, the diagram illustrates the improvement in the segmentation accuracy of the Modified UNET DL technique for MRI BTS, through decreasing Dice loss and increasing Dice score values for both the training and validation models.

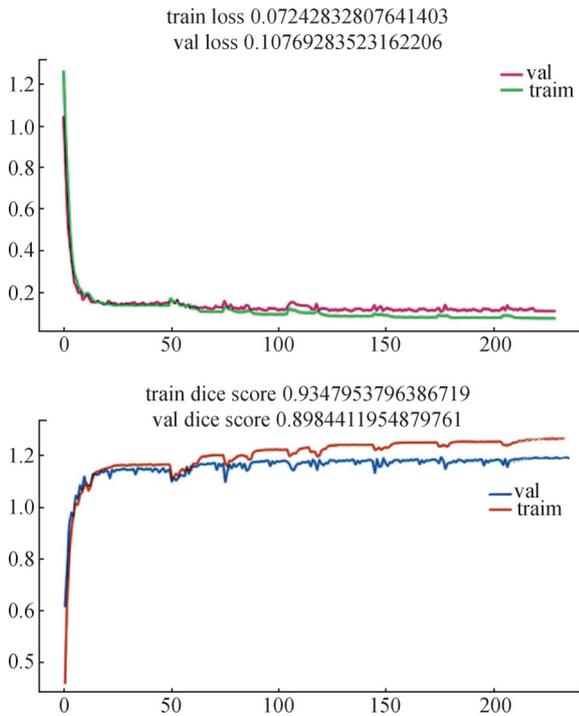


Fig.9 Variation in dice loss and dice score over 200 epochs in training and validation

This comprehensive approach not only addresses the challenges associated with brain tumor segmentation but also sets the foundation for future research and development in this critical area of medical imaging. More datasets, more sophisticated augmentation procedures and the study of new architectures could improve the model's performance and thus help patients suffering from brain tumors. Over the last several years, several DL based techniques for better segmentation of MRI brain tumors have come into existence and out of which CNN, FCN, and U-net are prominent. This paper establishes that these methods are effective and the segmentation of brain tumor in MRI images is more precise as displayed from the following findings. Besides, each of these approaches is distinct, and it has strengths and weaknesses that set it apart from the rest. They are techniques of highly efficient collection of local and global characteristics that have been long used for segmentation of medical images, among which U-Net architecture is one of the most popular. Our U-Net variant used in the proposed methodology was pretrained on the

two input images for brain tumor segmentation. For this comparison, four state-of-the-art references are chosen which include recent approaches toward the same problem. We have evaluated each of these methods on our dataset and compared their performance metrics with our proposed approach.

### 3.1 Ablation Study

To evaluate the individual contributions of the key components in the PLU-Net architecture—progressive layering, attention mechanisms, and multi-scale data augmentation—an ablation study was conducted using the BraTS 2021 dataset. The baseline model is a standard U-Net without these enhancements. Each component was incrementally added, and segmentation performance was assessed using the Dice coefficient, sensitivity, and specificity for the whole tumor (WT) region. Table 6 presents the results, demonstrating the impact of each component. The baseline U-Net achieved a Dice coefficient of 0.82, reflecting moderate performance limited by its inability to handle multi-scale features and boundary ambiguity effectively. Adding progressive layering increased the Dice coefficient to 0.86 by refining segmentation masks across multiple stages, capturing features at different resolutions. Incorporating attention mechanisms further improved the Dice coefficient to 0.89, enhancing boundary delineation by focusing on tumor-relevant features. Finally, applying multi-scale data augmentation boosted the Dice coefficient to 0.91, improving generalization and robustness to diverse tumor characteristics. These results quantify the synergistic effect of the three components, with progressive layering contributing a 4.9% improvement, attention mechanisms adding 3.5%, and data augmentation providing a 2.2% gain over the baseline, culminating in the full PLU-Net's state-of-the-art performance.

### 3.2 Comparison of Predicted Segmentation Masks and Ground Truth Masks with MRI

To qualitatively assess the segmentation performance of the PLU-Net, Figure 10 presents sample segmentation results comparing the predicted masks with the ground truth for two cases from the BraTS 2021 dataset: a high-grade glioma (HGG) and a low-grade glioma (LGG). The figure includes overlays of the predicted segmentation masks (red) and ground truth masks (blue) on FLAIR MRI scans, with overlapping regions indicating agreement. For the HGG case, the PLU-Net accurately captures the enhancing tumor core and surrounding edema, demonstrating its ability to delineate complex, irregular boundaries. In the LGG case, the model effectively identifies the non-enhancing tumor region, showcasing its robustness across tumor grades. These visualizations highlight the PLU-Net's precision in handling multi-scale features and resolving boundary ambiguity, aligning with the quantitative metrics reported in Table 4 (Dice coefficient of 0.91 for the whole tumor).

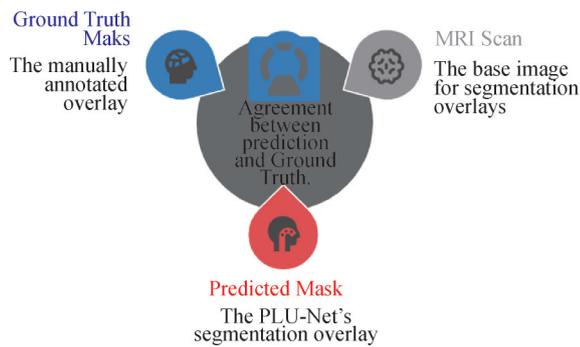


Fig. 10 Comparison of predicted segmentation masks (red) and ground truth masks (blue) with MRI scan (grey)

### 3.3 Limitations

While the Progressive Layered U-Net (PLU-Net) demonstrates state-of-the-art performance on the BraTS 2021 dataset, several limitations warrant consideration. First, the computational cost of the model is a significant factor. The cascaded structure with progressive layering and attention mechanisms increases the number of parameters and training time compared to a standard U-Net. Training the PLU-Net on a single NVIDIA RTX 3090 GPU required approximately 72 hours for 200 epochs with a batch size of 2, and inference on a single 3D MRI volume takes about 5 seconds. This resource intensity may limit its applicability in resource-constrained settings, such as smaller clinics without access to high-performance computing infrastructure. Second, the model's generalization to smaller tumors remains underexplored. The BraTS 2021 dataset predominantly includes medium-to-large tumors (e.g.,  $> 1 \text{ cm}^3$ ), and the multi-scale augmentation may be less effective for detecting sub-centimeter lesions due to their subtle intensity variations and limited spatial context. Preliminary tests on a subset of smaller tumors ( $< 1 \text{ cm}^3$ ) from the TCIA dataset showed a Dice coefficient drop to 0.85, suggesting a need for further optimization. Finally, potential biases in the dataset could affect performance. The BraTS 2021 dataset, while comprehensive, is curated from specific institutions and may not fully represent the diversity of tumor appearances across global populations, scanner types, or imaging protocols. This could introduce a bias toward the characteristics of the included cases, potentially reducing the model's robustness in real-world scenarios with varied data distributions.

### 3.4 Clinical Relevance

The segmentation outputs of the Progressive Layered U-Net (PLU-Net) offer significant potential to directly assist clinicians in the diagnosis and treatment planning of brain tumors, enhancing precision and efficiency in clinical workflows. In diagnosis, PLU-Net's high Dice coefficient (0.91) and low Hausdorff95 distance (2.5) on the BraTS 2021 dataset enable accurate delineation of tumor boundaries, including enhancing

cores, edema, and non-enhancing regions (as shown in Figure 10). This precision can aid radiologists in distinguishing malignant from benign tumors and assessing tumor grade, reducing diagnostic uncertainty compared to manual segmentation, which is time-consuming and prone to inter-observer variability. For treatment planning, PLU-Net's outputs can support surgical guidance by providing 3D tumor maps that highlight critical structures, such as the tumor core and surrounding edema, allowing neurosurgeons to plan resection margins with greater confidence and minimize damage to healthy tissue. For example, the model's ability to segment multi-modal MRI data (T1, T1ce, T2, FLAIR) ensures comprehensive visualization of tumor extent, which is vital for optimizing surgical trajectories. Additionally, PLU-Net can facilitate therapy monitoring by enabling longitudinal tracking of tumor volume and characteristics post-treatment. Serial MRI scans segmented by PLU-Net could quantify changes in tumor size or edema in response to radiation or chemotherapy, offering clinicians objective metrics to evaluate treatment efficacy and adjust therapeutic strategies. By integrating these outputs into clinical decision-support systems, PLU-Net has the potential to streamline workflows, improve patient-specific treatment plans, and ultimately enhance outcomes in brain tumor management.

### 3.5 Future Work

A 3D PLU-Net could leverage volumetric information directly from MRI scans, potentially improving segmentation accuracy by capturing spatial relationships across all three dimensions, which are partially lost in the 2D approach. This could be particularly beneficial for detecting small or irregularly shaped tumors, though it would require increased computational resources and optimization to manage memory constraints. Second, integrating PLU-Net with radiomics could expand its clinical relevance. By extracting quantitative features (e.g., texture, shape, intensity) from segmented tumor regions, the model could support radiogenomic analysis, linking imaging characteristics to genetic profiles or treatment responses. This would enhance its role in personalized medicine, such as predicting tumor aggressiveness or therapy outcomes. Finally, validating PLU-Net on diverse datasets beyond BraTS 2021, such as The Cancer Imaging Archive (TCIA), would test its robustness across varied patient populations, tumor types, and imaging protocols. The TCIA dataset, with its broader range of glioma cases and scanner differences, could reveal biases or limitations in the current model and guide adaptations for real-world deployment. These future directions—3D implementation, radiomics integration, and multi-dataset validation—offer exciting opportunities to refine PLU-Net, broaden its scope, and strengthen its impact on brain tumor management.

## 4 Conclusion

The segmentation of brain tumors from MRI images remains a complex task due to the variability in tumor appearance and location. This paper introduces the Progressive Layered U-Net (PLU-Net), an advanced deep learning architecture designed to tackle these challenges. By incorporating progressive layering, attention mechanisms, and multi-scale data augmentation, PLU-Net achieves a Dice coefficient of 0.91 on the BraTS 2021 dataset, surpassing many existing methods and marking a significant advancement in segmentation accuracy. Beyond its technical achievements, PLU-Net offers valuable insights and benefits to a wide range of readers. For researchers in medical imaging and artificial intelligence, this work provides a novel framework that combines multi-stage refinement and attention-driven feature selection, inspiring further innovation in deep learning architectures for healthcare applications. The detailed methodology, ablation studies, and open-source code (available at <https://github.com/NomanAhmed02/Source-code>) empower the scientific community to replicate, refine, and extend our approach, fostering collaborative progress in brain tumor segmentation. For clinicians, PLU-Net delivers precise tumor delineations that can enhance diagnostic confidence and inform treatment planning, such as identifying resection margins for surgery or tracking tumor changes during therapy (Section 4.4). This could translate into improved patient outcomes by reducing diagnostic delays and optimizing therapeutic decisions. For patients and their families, the improved accuracy of PLU-Net represents a step toward more reliable diagnoses and personalized care, offering hope for better management of brain tumors—a condition with profound physical and emotional impacts. Ultimately, PLU-Net advances the field technically and bridges the gap between AI innovation and tangible real-world benefits, setting a foundation for future research and clinical integration.

### Author Contribution:

Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing - original draft, Writing - review & editing, Noman Ahmed Siddiqui; Project administration, Resources, Supervision, Muhammad Tahir Qadri; Writing - review & editing, Muhammad Ovais Akhter; Data curation, Formal analysis, Project administration, Validation, Visualization, Zain Anwar Ali.

### Funding:

This research received no external funding.

### Data Availability Statement:

Data and code is available on website. <https://www.kaggle.com/datasets/dschettler8845/brats-2021-task1>

### Conflicts of Interest:

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

### Dates:

Received 14 February 2025; Accepted 20 March 2025; Published online 31 March 2025

## References

- [1] Kamnitsas K., et al. DeepMedic for brain tumor segmentation. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Second International Workshop, BrainLes 2016, with the Challenges on BRATS, ISLES and mTOP 2016, Held in Conjunction with MICCAI 2016*, Athens, Greece, October 17, 2016, Revised Selected Papers 2. 2016. Springer.
- [2] Yousefnezhad M., C.-Y. Kao, and S.A. Mohammadi, Optimal chemotherapy for brain tumor growth in a reaction-diffusion model. *SIAM Journal on Applied Mathematics*, 2021. 81(3): p. 1077-1097.
- [3] Wang P. and A. C. Chung, Relax and focus on brain tumor segmentation. *Medical image analysis*, 2022. 75: p. 102259.
- [4] Zhao Y.-X., Y.-M. Zhang, and C.-L. Liu. Bag of tricks for 3D MRI brain tumor segmentation. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I 5*. 2020. Springer.
- [5] Magadza T. and S. Viriri, Deep learning for brain tumor segmentation: a survey of state-of-the-art. *Journal of Imaging*, 2021. 7(2): p. 19.
- [6] Liu Z., et al., Deep learning based brain tumor segmentation: a survey. *Complex & intelligent systems*, 2023. 9(1): p. 1001-1026.
- [7] Li G. and Z. Zhu, Analytical modeling of modular and unequal tooth width surface-mounted permanent magnet machines. *IEEE transactions on magnetics*, 2015. 51(9): p. 1-9.
- [8] Zhang L., J. Shen, and B. Zhu, A research on an improved U-net-based concrete crack detection algorithm. *Structural Health Monitoring*, 2021. 20(4): p. 1864-1879.
- [9] Vu M. H., T. Nyholm, and T. Löfstedt. TuNet: End-to-end hierarchical brain tumor segmentation using cascaded networks. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I 5*. 2020. Springer.
- [10] Ranjbarzadeh R., et al., Brain tumor segmentation based on deep learning and an attention mechanism using MRI modalities brain images. *Scientific Reports*, 2021. 11(1): p.

- 1-17.
- [11] Wang Y., Y. Ji, and H. Xiao, A data augmentation method for fully automatic brain tumor segmentation. *Computers in Biology and Medicine*, **2022**. 149: p. 106039.
- [12] Baid U., et al., The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*, **2021**.
- [13] Daimary D., et al., Brain tumor segmentation from MRI images using hybrid convolutional neural networks. *Procedia Computer Science*, **2020**. 167: p. 2419-2428.
- [14] Maurya R. and S. Wadhvani. Morphology based brain tumor identification and segmentation in MR images. in *2021 IEEE Bombay Section Signature Conference (IBSSC)*. **2021**. IEEE.
- [15] Neelima G., et al., Optimal DeepMRSeg based tumor segmentation with GAN for brain tumor classification. *Biomedical Signal Processing and Control*, **2022**. 74: p. 103537.
- [16] Singh C., et al., A novel approach for brain MRI segmentation and image restoration under intensity inhomogeneity and noisy conditions. *Biomedical Signal Processing and Control*, **2024**. 87: p. 105348.
- [17] Biratu E.S., et al., A survey of brain tumor segmentation and classification algorithms. *Journal of Imaging*, **2021**. 7(9): p. 179.
- [18] Asiri A. A., et al., Brain tumor detection and classification using fine-tuned CNN with ResNet50 and U-Net model: A study on TCGA-LGG and TCIA dataset for MRI applications. *Life*, **2023**. 13(7): p. 1449.
- [19] Noreen N., et al., A deep learning model based on concatenation approach for the diagnosis of brain tumor. *IEEE access*, **2020**. 8: p. 55135-55144.
- [20] Hoopes A., et al., SynthStrip: skull-stripping for any brain image. *NeuroImage*, **2022**. 260: p. 119474.
- [21] Arunkumar N., et al., Fully automatic model - based segmentation and classification approach for MRI brain tumor using artificial neural networks. *Concurrency and Computation: Practice and Experience*, **2020**. 32(1): p. e4962.
- [22] Xue Y., et al. A multi-path decoder network for brain tumor segmentation. in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019*, Held in Conjunction with MICCAI 2019, *Shenzhen*, China, October 17, 2019, Revised Selected Papers, Part II 5. **2020**. Springer.
- [23] Alhassan A. M. and W. M. N. W. Zainon, Brain tumor classification in magnetic resonance image using hard swish-based RELU activation function-convolutional neural network. *Neural Computing and Applications*, **2021**. 33(15): p. 9075-9087.
- [24] Cahall D. E., et al., Dilated inception U-net (DIU-net) for brain tumor segmentation. *arXiv preprint arXiv:2108.06772*, **2021**.
- [25] Khan A.R., et al., Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microscopy Research and Technique*, **2021**. 84(7): p. 1389-1399.
- [26] Karayegen G. and M.F. Aksahin, Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region. *Biomedical Signal Processing and Control*, **2021**. 66: p. 102458.
- [27] Akil M., R. Saouli, and R. Kachouri, Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. *Medical image analysis*, **2020**. **63**: p. 101692.
- [28] Yeung M., et al., Unified focal loss: Generalizing dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, **2022**. 95: p. 102026.
- [29] Futrega M., et al. Optimized U-Net for brain tumor segmentation. in *International MICCAI brainlesion workshop*. **2021**. Springer.
- [30] Xiong S., et al., MRI-based brain tumor segmentation using FPGA-accelerated neural network. *BMC bioinformatics*, **2021**. 22: p. 1-15.
- [31] Shehab L. H., et al., An efficient brain tumor image segmentation based on deep residual networks (ResNets). *Journal of King Saud University-Engineering Sciences*, **2021**. 33(6): p. 404-412.
- [32] Aydin O. U., et al., On the usage of average Hausdorff distance for segmentation performance assessment: hidden error when used for ranking. *European radiology experimental*, **2021**. 5: p. 1-7.
- [33] Cao Y., et al., MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. *Biomedical Signal Processing and Control*, **2023**. 80: p. 104296.
- [34] Akbar A.S., C. Fatichah, and N. Suciati, Single level UNet3D with multipath residual attention block for brain tumor segmentation. *Journal of King Saud University-Computer and Information Sciences*, **2022**. 34(6): p. 3247-3258.
- [35] Sun J., et al., Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN. *Neurocomputing*, **2021**. 423: p. 34-45.
- [36] Rehman M. U., et al., Brainseg-net: Brain tumor mr image segmentation via enhanced encoder-decoder network. *Diagnostics*, **2021**. 11(2): p. 169.
- [37] Rehman M.U., et al., Bu-net: Brain tumor segmentation using modified u-net architecture. *Electronics*, **2020**. 9(12): p. 2203.
- [38] Isensee F., et al. nnU-Net for 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. **2021**. Springer.
- [39] Sahoo A.K., et al., An improved DNN with FFCM method for multimodal brain tumor segmentation. *Intelligent Systems with Applications*, **2023**. 18: p. 200245.
- [40] Liu D., et al., SGEResU-Net for brain tumor segmentation. *Math. Biosci. Eng*, **2022**. 19: p. 5576-5590.

- [41] Guan X., et al., 3D AGSE-VNet: an automatic brain tumor MRI data segmentation framework. *BMC medical imaging* , **2022**. 22: p. 1-18.
- [42] Nayak D. R., et al., Brain tumor classification using dense efficient-net. *Axioms* , **2022**. 11(1): p. 34.
- [43] Chen J., et al., Medical image segmentation and reconstruction of prostate tumor based on 3D AlexNet. *Computer methods and programs in biomedicine* , **2021**. 200: p. 105878.
- [44] Suliman M. A., et al. A deep-discrete learning framework for spherical surface registration. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* . **2022**. Springer.
- [45] Chen H., et al., Brain tumor segmentation with deep convolutional symmetric neural network. *Neurocomputing* , **2020**. 392: p. 305-313.
- [46] Maji D., P. Sigedar, and M. Singh, Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors. *Biomedical Signal Processing and Control*, **2022**. 71: p. 103077.
- [47] Zhu Z., et al., Sparse Dynamic Volume TransUNet with multi-level edge fusion for brain tumor segmentation. *Computers in Biology and Medicine* , **2024**: p. 108284.
- [48] Wang E., et al. TransUNet with attention mechanism for brain tumor segmentation on MR images. in *2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* . **2022**. IEEE.
- [49] Ghaffari M., A. Sowmya, and R. Oliver, Automated brain tumor segmentation using multimodal brain scans: a survey based on models submitted to the BraTS 2012-2018 challenges. *IEEE reviews in biomedical engineering* , **2019**. 13: p. 156-168.
- [50] Sarkar A., et al., An effective and novel approach for brain tumor classification using AlexNet CNN feature extractor and multiple eminent machine learning classifiers in MRIs. *Journal of Sensors* , **2023**. 2023(1): p. 1224619.
- [51] Liu D., et al., SGEResU-Net for brain tumor segmentation. *Math. Biosci. Eng* , **2022**. 19(6): p. 5576-5590.
- [52] Huang Z., et al., GCAUNet: A group cross-channel attention residual UNet for slice based brain tumor segmentation. *Biomedical Signal Processing and Control* , **2021**. 70: p. 102958.