IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 February 2024, accepted 12 May 2024, date of publication 20 May 2024, date of current version 29 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3402947

RESEARCH ARTICLE

TransResUNet: Revolutionizing Glioma Brain Tumor Segmentation Through Transformer-Enhanced Residual UNet

NOVSHEENA RASOOL[1], JAVAID IQBAL BHAT[1], NIYAZ AHMAD WANI[2], (Member, IEEE), NAVEED AHMAD[3], (Member, IEEE), AND MOHAMMED ALSHARA[3,4]

[1]Department of Computer Science, Islamic University of Science and Technology, Kashmir, Jammu and Kashmir 192122, India
[2]Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology, Patiala, Punjab 147004, India
[3]College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia
[4]College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University, Riyadh 11432, Saudi Arabia

Corresponding author: Novsheena Rasool (novsheena.rasool@iust.ac.in)

This work was supported by Prince Sultan University funded for the Article Processing Charges (APC).

ABSTRACT Accurate segmentation of brain tumors from MRI (Magnetic Resonance Imaging) sequences is essential across diverse clinical scenarios, facilitating precise delineation of anatomical structures and disease-affected areas. This study presents an innovative deep-learning method for segmenting glioma brain tumors, utilizing a hybrid architecture that combines ResNet U-Net with Transformer blocks. The proposed model adeptly encompasses both the local and global contextual details present in MRI scans. The architecture includes an encoder based on ResNet for extracting hierarchical features, followed by residual blocks to enhance feature representation while maintaining spatial information. Additionally, a central transformer block, incorporating multi-head attention mechanisms, enables the modeling of long-range dependencies and contextual comprehension, progressively refining feature interactions. To handle structural scale variations within MRI images, skip connections are utilized during the decoding phase. Transposed convolutional layers in the decoder upsample feature maps, retaining details and incorporating contextual information from earlier layers. A rigorous assessment of the model's functionality was carried out with the BraTS2019 dataset, employing a comprehensive set of evaluation metrics including accuracy, IOU score, specificity, sensitivity, dice score, and precision. The evaluation focused on individual tumor classes, namely the whole, core, and enhancing tumor regions. During validation, the suggested model demonstrated remarkable dice scores of 0.91, 0.89, and 0.84 for the whole tumor, core tumor, and enhancing tumor, respectively, yielding an impressive overall accuracy rate of 98%.

INDEX TERMS Glioma, segmentation, MRI, ResNet-transformer, deep learning.

I. INTRODUCTION

In recent years, advances in medical imaging leveraging deep learning (DL) have substantially enhanced various sectors [1], [2], [3] including the healthcare sector, with a focus on refining the accuracy and effectiveness of automated disease detection, segmentation, and classification [4], [5]. This progress is driven by the pressing need to improve diagnostic procedures and treatment strategies for patients afflicted with various ailments. Within this domain, two primary categories of approaches have emerged as focal points of innovation: those that employ generative approaches and those that rely on discriminative approaches [6]. Generative models, such as diverse neural networks, strive to grasp the underlying data distribution of diseases, thereby generating highly lifelike disease analyses [7]. These methods hold

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Zhou.

tremendous promise for capturing intricate disease nuances and may even assist in crafting realistic 3D reconstructions for surgical planning. Conversely, discriminative models harness sophisticated machine learning (ML) techniques to directly pinpoint and outline disease regions within medical images. By adeptly distinguishing between diseased and healthy areas, discriminative models excel in precise analysis tasks, thereby facilitating accurate disease diagnosis and monitoring. As the field advances, the synergy between generative and discriminative models, bolstered by the integration of cutting-edge technologies like deep learning and artificial intelligence, promises to revolutionize disease analysis into a more precise, efficient, and accessible process. Ultimately, this advancement has the potential to elevate the prognosis and quality of life for patients grappling with these challenging medical conditions.

Glioma is a tumor that originates within the central nervous system's glial cells, which are essential for nurturing and protecting nerve cells. Gliomas typically comprise different regions, including the whole tumor (WT), the core tumor (CT) containing necrotic cells, and the enhancing tumor region (ET) with heightened signal intensity. These distinct regions are identified through advanced imaging techniques and contribute significantly to the diagnosis process and the planning of treatment. Gliomas present significant diversity in terms of aggressiveness, localization, and prognosis [8]. Central to glioma classification is the grading system established by the World Health Organization (WHO) [9]. Gliomas are classified into 4 groups spanning from I to IV, where grade I represents the least severe form and grade IV denotes the most aggressive. Low-grade gliomas, encompassing grades I and II, manifest slower growth rates and are associated with a more favorable prognosis. In contrast, high-grade gliomas, comprising grades III and IV, exhibit rapid growth and an increased likelihood of malignancy [10]. High-grade gliomas, particularly grade IV, are commonly referred to as glioblastomas (GBM). Glioblastomas represent the most severe and malignant subtype among gliomas. They exhibit swift expansion, invasive tendencies, and resilience against conventional therapies [11]. Unfortunately, glioblastomas are associated with a poor prognosis, with most patients surviving only around 12 to 18 months following diagnosis [12], even with aggressive treatment approaches such as surgery, radiation therapy, and chemotherapy. The challenges in treating high-grade gliomas stem from their ability to invade surrounding brain tissue and their resistance to complete surgical removal or targeted therapies.

Glioma segmentation constitutes a crucial task in medical imaging, aimed at the automated identification and delineation of various tumor areas, such as the core tumor (the central tumor region) containing necrotic cells, the enhancing tumor area harboring active cells from the tumor, and the whole tumor region [13]. This segmentation process relies on the analysis of diverse multi-modal MRI modalities obtained through advanced medical imaging methodologies, which often incorporate deep learning-based techniques. Commonly employed MR imaging sequences for glioma analysis encompass contrast-enhanced T1-weighted (T1ce), T1-weighted, T2-weighted, and Fluid Attenuation Inversion Recovery (FLAIR) modalities [14]. These modalities provide supplementary data for examining various subregions within gliomas. FLAIR and T2 modalities are particularly effective in highlighting the tumor and surrounding edematous tissue [15]. Conversely, T1 and T1ce sequences are adept at emphasizing the enhancing tumor region harboring active tumor cells, excluding the surrounding edema. Additionally, within the tumor core observed in T1c images, there may exist an area with enhanced signal intensity, denoted as the enhancing tumor core. Each of these regions provides vital information about the tumors' size, shape, and specific characteristics. This information is invaluable in brain tumor diagnosis and monitoring, as it assists medical professionals in understanding the nature and severity of the tumor. Moreover, it is vital for planning treatment and helping clinicians decide on the most appropriate interventions, whether they involve surgery, radiation therapy, or chemotherapy.

Accurately segmenting gliomas and their internal tumor structures in MRI modalities is paramount, not only for subsequent follow-up assessments but also for devising distinct treatment planning strategies tailored to different regions of the tumor, thereby further emphasizing the critical need for precise segmentation [16]. However, the conventional approaches involving segmentation by hand are undeniably time-consuming and prone to both inter- and intra-rater discrepancies that can be challenging to quantify. Consequently, physicians often resort to approximate measures when evaluating glioma boundaries and characteristics [17]. Despite the challenging diversity in the form, arrangement, and positioning of abnormal pathological features, as well as the complicating effect of tumor masses on neighboring normal tissues, there is an urgent demand for more precise, and efficient fully automatic segmentation methods. These methods must address issues such as tumor complexity, location uncertainty, class imbalance, and imaging data variability [6]. Moreover, inherent issues in MRI images, such as intensity inhomogeneity and variations in intensity ranges across sequences and acquisition scanners [14], contribute to the complexity of the segmentation process. Although significant progress has been made in segmenting these tumors [18], [19], [20], [21], [22], [23], persistent challenges necessitate multi-stage treatment for patients. This study introduces a new deep-learning approach that integrates ResNet U-Net with Transformer blocks to address these challenges. Our model enhances contextual understanding by gathering both local and global details in MRI images, thereby improving segmentation accuracy, the proposed system model is depicted in Fig.1. The paper is organized as follows: The introduction section presents the importance and challenges of brain tumor segmentation; the related work reviews existing methodologies; the proposed methodology includes a dataset description, preprocessing procedure, network architecture, training process, and experimental

results; Finally, the conclusion and future research directions are discussed.

II. RELATED WORK

A. TUMOR SEGMENTATION

Recent breakthroughs in healthcare have ignited the development of innovative methodologies and tools, leading to significant advancements in the domain of healthcare image analysis. In this section, we will explore the current approaches that have propelled the field forward.

Huang et al. [20] introduced NLSE-VNet, a new approach designed to segregate gliomas in MRI images and forecast the survival duration of glioma patients. NLSE-VNet exhibited noteworthy performance, achieving an average Dice coefficient of 79% for brain tumor segmentation tasks and demonstrating a low average RMSE of 311.5 for survival prognosis across the BraTS 2019 and 2020 databases. González et al. [24] introduced a group of asymmetric U-Net-style algorithms aimed at segmenting and enhancing tumor regions. Additionally, they employed the DenseNet method for the prognosis of survival. The assessment conducted on the BraTS 2020 test set resulted in dice scores of 0.80, 0.87, and 0.80 for the enhancing tumor region, whole tumor region, and tumor core region, respectively, with an overall dice coefficient of 0.82. Tamilarasi [25] introduced a convolutional neural network (CNN) classification technique aimed at distinguishing between gliomas and normal brain MRI images. Initially, the brain MRI image undergoes adaptive histogram equalization to accentuate abnormal pixels relative to their surroundings. Subsequently, the enhanced brain image is subjected to multidirectional scaling using the Gabor transform, facilitating feature extraction. These extracted features are then utilized for training and classification via the CNN deep learning algorithm. Finally, tumor regions are delineated through morphological operations. This proposed segmentation technique employing CNN classification achieved notable performance metrics, including 96.9% sensitivity, 99.3% specificity, and 99.2% accuracy. Kumar et al. [26] employed residual models and structured them into a UNet architecture for tissue segmentation. They devised a new residual algorithm to serve as the foundation for constructing the UNet architecture from scratch, facilitating the segmentation of multispectral images. The resulting model achieved a Jaccard index (IoU) of 0.82 and a dice coefficient of 0.90. These advancements underscore the continuous evolution of medical image analysis, with each contribution pushing the boundaries of knowledge and paving the way for improved methodologies and tools in healthcare.

B. MOTIVATION

The precise segmentation of brain tumors from MRI images is a significant task in neuroimaging and clinical diagnostics, with profound implications for patient care in the domain of neuro-oncology. Conventional methods relying on manual delineation are fraught with time constraints and subjective interpretations, leading to inconsistencies and variability between observers. This variability undermines the reliability of diagnostic assessments and treatment planning, highlighting the urgent need for more robust and efficient segmentation techniques. Our work is motivated by the pressing need to address these challenges and progress to the forefront of glioma tumor segmentation. Leveraging recent advancements in deep learning, automated segmentation techniques offer the promise of improved accuracy and efficiency compared to manual methods. However, existing approaches often struggle to capture the complex morphological and textural characteristics [18] of brain tumors, particularly gliomas, which are highly heterogeneous and infiltrative.

C. CONTRIBUTION

In this endeavor, our study contributes to the following key aspects:

1) NOVEL HYBRID ARCHITECTURE

We propose an innovative method for segmenting glioma tumors by integrating ResNet U-Net with Transformer blocks. This innovative architecture combines the strengths of ResNet blocks for feature extraction with the attention mechanism of transformers, enabling more precise delineation of tumor boundaries.

2) ENHANCED PERFORMANCE

Through extensive evaluation of the BraTS2019 HGG dataset, our proposed methodology demonstrates exceptional performance, surpassing current cutting-edge approaches to accuracy and effectiveness. This improved performance is crucial for reliable diagnostic assessments and treatment planning in neuro-oncology.

III. PROPOSED METHODOLOGY

A. DATASET DESCRIPTION

The BraTS dataset from 2019 stands as a crucial asset in propelling the field of medical imaging forward, particularly in the field of brain tumor segmentation and treatment. It is a diverse compilation of brain MRI scans sourced from various clinical centers, encompassing multiple MRI techniques such as T1-weighted, fluid-attenuated inversion recovery, T2-weighted, and post-contrast T1-weighted scans. These scans were acquired using diverse clinical rules and a variety of scanners from various institutions (n = 19), and all of them are available in nifti file format. An important feature of this dataset is the meticulous manual segmentation carried out by 1 to 4 raters, all precisely adhering to the same annotation process [14]. These annotations, which have received validation from experienced neuro-radiologists, encompass crucial regions of the tumor with enhanced contrast (referred to as ET with label 4), the surrounding edema (designated as ED with label 2), central area of necrosis and non-enhancing tumor (commonly denoted as NCR/NET with label 1) and the background with label 0. Furthermore,the data has undergone rigorous pre-processing,

including aligning with a standardized structural design, adjusting to a consistent resolution of $1mm^3$, and removing the skull for consistent analysis [21]. This dataset also includes crucial information on overall survival (OS), which is quantified in days and stored in a CSV file format. This CSV document contains entries related to pseudo-identifiers of imaging data, patient ages, and details regarding resection status. This dataset focuses primarily on individuals who have undergone gross total resection (GTR) as their resection status, highlighting them for evaluation purposes.

B. PREPROCESSING

Our preprocessing pipeline encompassed the following key steps:

1) N4ITK BIAS FIELD CORRECTION

Our initial procedure commenced with the partitioning of 259 BraTS 2019 high-grade glioma images into an 80-20 split, allocating 80% for training and 20 for validation purposes. Subsequently, we applied the N4ITK (N4 Bias Field Correction) technique to the MRI images. N4ITK is a well-established and widely used method for addressing non-uniformities in image intensity, particularly in MRI data [27]. These non-uniformities can arise from various factors during MRI acquisition, such as magnetic field variations or imperfections in the machine. Correcting these intensity irregularities is crucial as it ensures that the MRI images maintain consistency and accuracy in their intensity levels [28]. Furthermore, bias correction is essential for preventing misleading or inaccurate interpretations of the images, making them reliable and valuable tools for diagnosis and treatment planning within the field of healthcare.

2) Z-SCORE NORMALIZATION

Following N4ITK bias field correction, we performed Z-score normalization on the MRI modalities (T1, T1ce, Flair, T2). Z-score normalization involves centering data around a mean of 0 and adjusting its scale to have a standard deviation of 1 [27]. This step enhances comparability and stability for subsequent analysis. However, we exempted the mask files from this normalization process. These masks contain multiple labels, with label 0 representing the background, labels 1 and 2 indicating the core tumor and edema, respectively, and label 3 signifying the enhancing tumor. To maintain consistency and alignment with the task at hand, we replaced mask label 4 with label 3.The z-score normalization equation is given by:

$$z_{ijk} = \frac{V_{ijk} - \mu}{\sigma}$$

where, z_{ijk} is the standardized intensity value (z-score) of the voxel at position (i, j, k), V_{ijk} is the original intensity value of the voxel at position (i, j, k), μ is the mean intensity value of the entire volume and σ is the standard deviation of the intensity values of the entire volume.

3) PATCH EXTRACTION FOR LOCALIZED REGION GENERATION

The BraTS 2019 dataset presents three-dimensional MRI images with dimensions of $240 \times 240 \times 155$, offering valuable information [14] but also poses challenges in memory and computation. To address these concerns effectively, after normalization, we stack the three dimensional volumes and crop them to focus on relevant anatomical structures. Subsequently, we randomly select two depth indices within each cropped multi-channel volume to ensure diversity and extract two patches per volume. For each selected depth index, random starting coordinates for height and width were generated, defining the boundaries of the extracted patches. The resulting patches, representing smaller localized volumes, were saved for further analysis. Each extracted patch has a size of $128 \times 128 \times 128 \times 4$, where the last dimension (4) represents the number of channels or modalities. This ensures that all modalities (e.g., T1, T1ce, Flair, and T2) are preserved in the extracted patches, enabling comprehensive analysis and feature extraction. This systematic approach streamlines preprocessing efforts, facilitating efficient utilization of computational resources while retaining critical information within the brain tumor MRI volumes.

C. NETWORK ARCHITECTURE AND TRAINING

This section outlines the architecture of our suggested three-dimensional brain tumor segmentation model, building upon the foundational U-Net architecture [29] while incorporating enhancements to improve its performance. Departing from the 2D U-Net framework, we transition to a 3D U-Net design. Moreover, instead of solely employing basic convolution layers, we incorporate ResNet blocks and transformer blocks into the architecture. The model architecture comprises the following components:

1) ENCODER WITH RESIDUAL BLOCKS

We input the central two patches from each preprocessed multimodal MRI, each sized $128 \times 128 \times 128 \times 4$, into the encoder section of our network. These images are considered multimodal as they consist of stacked T1, T2, T1CE, and FLAIR sequences, which are commonly utilized in medical imaging for different contrasts and tissue types. The dimensions of $128 \times 128 \times 128$ represent the spatial size of the central input patch, while the '4' signifies the number of channels representing each sequence. Each channel corresponds to a different imaging sequence, allowing our model to capture diverse information for accurate segmentation. The encoder section of our network employs a sequence of ResNet blocks to capture hierarchical features from the input image data. Each ResNet block comprises convolutional layers followed by batch normalization and activation functions, allowing the model to capture increasingly abstract features. These ResNet blocks serve the purpose of extracting hierarchical characteristics from the MRI scans while also reducing the spatial dimensions

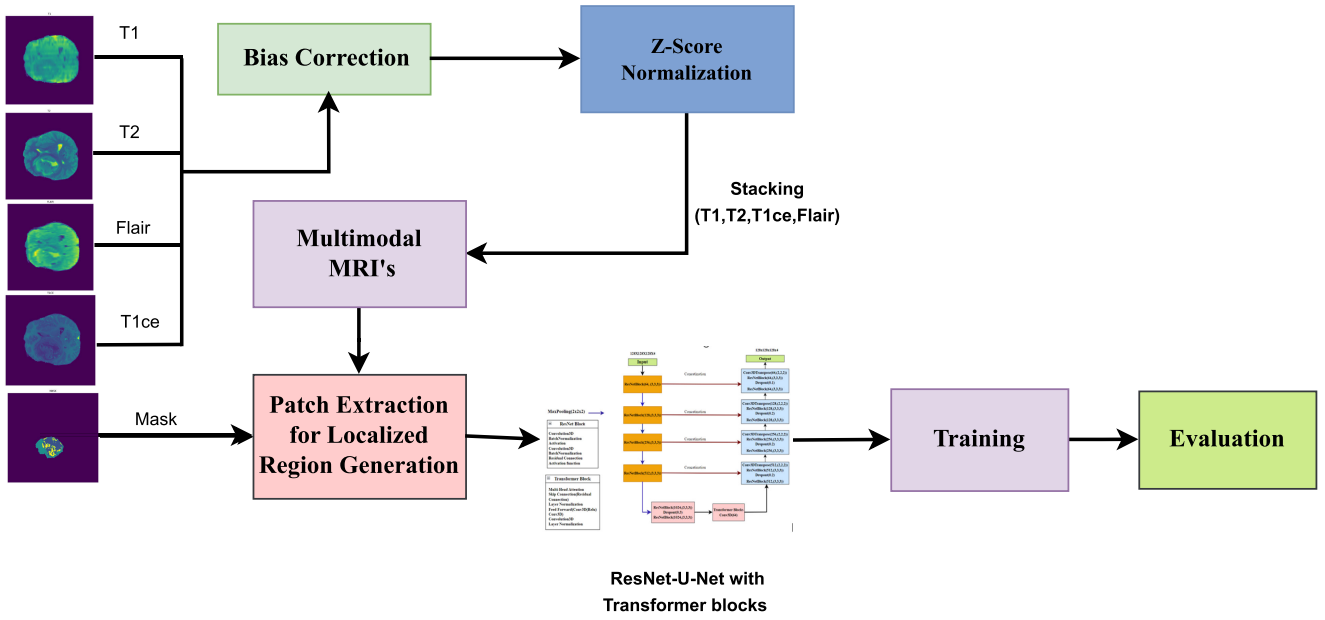


FIGURE 1. System model.

through the utilization of max-pooling layers (represented by a blue arrow), as shown in Fig. 2. We opt for ResNet blocks over direct CNN layers because they effectively address a significant challenge encountered when training deep image segmentation networks—the vanishing gradient problem [30]. The encoder blocks comprise a pair of convolutional layers featuring a progressive increment in the number of filters (64, 128, 256, 512), accompanied by residual connections to facilitate gradient flow during training. Through the integration of residual connections, the model ensures smooth gradient flow across its many layers, facilitating the learning of intricate features from complex volumetric image data. These residual connections enable the network to capture residual information, representing the discrepancy between input and output, which is essential for preserving fine details and enhancing segmentation accuracy.The operations within the ResNet block can be described as follows:

1) Convolution and Activation (Conv1):

$$\text{Conv1}(X) = \text{Act}\left(\text{BN}\left(\text{Conv3D}(X, F, K, S, P)\right)\right) \quad (1)$$

The input tensor X undergoes a 3D convolution operation (Conv3D) with parameters F filters, K kernel size, S strides, and P padding. The result is passed through batch normalization (BN) and then an activation function (Act), typically ReLU.

2) Convolution and Batch Normalization (Conv2):

$$\text{Conv2}(\text{Conv1}(X)) = \text{BN}\left(\text{Conv3D}\left(\text{Conv1}(X), F, K, P\right)\right) \quad (2)$$

The output of the first convolutional layer (Conv1) is passed through another Conv3D operation with the same filter size, kernel size, and padding but without stride. The result is normalized using batch normalization (BN).

3) Residual Connection (Residual):

$$\text{Residual}(X) = \begin{cases} \text{Conv3D}(X, F, (1,1,1), S, P) \\ \quad \text{if } S \neq (1,1,1) \text{ or } X.\text{shape}[-1] \neq F \\ X \\ \quad \text{otherwise} \end{cases} \quad (3)$$

If the strides (S) in Conv1 are not (1, 1, 1), or the number of output channels of Conv1 (X.shape$[-1]$) is not equal to F, then a $1 \times 1 \times 1$ convolution (Conv3D) is applied to adjust the dimensions. Otherwise, the input tensor X remains unchanged.

4) Final Output (Output):

$$\text{Output}(X) = \text{Act}\left(\text{Add}\left(\text{Residual}(X), \text{Conv2}\left(\text{Conv1}(X)\right)\right)\right) \quad (4)$$

The output of the Residual block and the output of Conv2 are added together. The sum is passed through an activation function (Act) to produce the final output.

2) MIDDLE SECTION: INTEGRATION OF RESIDUAL BLOCKS AND TRANSFORMER BLOCKS

In the middle section of the model, we integrate two consecutive residual blocks, each containing 1024 filters. These blocks consist of a pair of convolutional layers, succeeding batch normalization, and rectified linear unit (ReLU) activation, facilitating feature extraction and non-linear transformations. Additionally, residual connections within each block aid in mitigating the vanishing gradient problem and enhancing optimization during training. Following the residual blocks, we incorporate four transformer blocks, each applying multi-head self-attention mechanisms [31] to encompass distant relationships within the feature maps. This

allows the model to selectively focus on different regions (whole, core, and enhancing tumor) within the input data. By computing attention scores between positions within the feature maps, relevant spatial relationships and dependencies are prioritized. To ensure smooth information flow across different model layers and to mitigate potential information loss during processing, we integrate skip connections within the transformer blocks. These connections allow us to preserve valuable information from earlier processing stages, thereby enriching our model's ability to encompass both local and global contextual information. Specifically, skip connections are established, linking the input feature maps with the output of the multi-head self-attention mechanism, followed by layer normalization. In addition to the self-attention mechanism, feed-forward layers play a crucial role in introducing non-linear transformations to feature representations. These layers consist of convolutional operations, incorporating ReLU activation functions to capture intricate patterns and relationships within the data, as illustrated in Fig. 2. Also, to ensure stability during training and enhance generalization to unseen data, layer normalization is utilized within the transformer blocks to ensure consistent distributions of inputs across each layer, thereby reducing the likelihood of vanishing or exploding gradients. Layer normalization is applied after both the multi-head self-attention and the feed-forward layers. Furthermore, a supplementary convolutional layer, featuring 64 filters and employing a $3 \times 3 \times 3$ kernel size, is utilized to modulate the channel count within the feature maps. Batch normalization and activation of ReLU are applied to ensure stable activations and introduce non-linearity to the feature transformations, ultimately facilitating the generation of accurate segmentation masks.

The operations within the transformer block can be described as follows: Let x be the input data to the transformer block

1) MultiHead Attention Mechanism:

$$\text{attn\_output} = \text{MultiHeadAttention}(x, x) \tag{5}$$

This operation applies the multi-head attention mechanism to the input tensor x, producing the output tensor attn_output. It computes attention scores between positions in the feature maps to prioritize relevant spatial relationships and dependencies.

2) Layer Normalization:

$$\text{out1} = \text{LayerNormalization}(x + \text{attn\_output}) \tag{6}$$

Layer normalization is applied to the sum of the input tensor x and the output of the multi-head attention mechanism (attn_output), resulting in the output tensor out1. This operation normalizes the activations across each layer of the transformer block.

3) FeedForward Layers:

$$\text{ffn} = \text{Conv3D}(\text{ReLU}(\text{Conv3D}(\text{ff\_dim}, 1, \text{out1}))) \tag{7}$$

The feed-forward layers consist of two convolutional layers with ReLU activation functions applied to the output tensor out1. The first convolutional layer reduces the dimensionality to ff_dim, followed by another convolutional layer that restores the original number of channels.

4) Additional Convolution Layer:

$$\text{output} = \text{LayerNormalization}(\text{out1}) + \text{Conv3D}(x.shape[-1], 1, \text{ffn}) \tag{8}$$

Finally, the output tensor is obtained by adding the output of the feed-forward layers to the original input tensor x, followed by layer normalization. An additional convolutional layer is then applied to adjust the number of channels back to the original value.

3) DECODER

In the decoder phase of our model, we aim to reconstruct high-resolution segmentation masks by increasing the resolution of the feature maps to match the dimensions of the original input, while integrating information from the corresponding encoder layers via skip connections. We begin the decoder with the first upsampling layer, which doubles the spatial dimensions of the feature maps using a transposed convolution operation. Simultaneously, we extract feature maps from the matching encoder layer and merge them with the upsampled feature maps. This concatenation enables us to recover spatial details lost during the downsampling process. Following the concatenation, the combined feature maps are passed through two residual blocks. These residual blocks are comprised of two convolutional layers, succeeded by batch normalization and ReLU activation functions, facilitating the learning of residual mappings. Subsequently, the upsampling layers follow a similar pattern, progressively increasing the spatial resolution of the feature maps and concatenating them with the corresponding encoder layers. These operations establish skip connections, allowing us to directly access low-level features captured in the encoder layers. This mechanism helps in preserving fine-grained details crucial for accurate segmentation. Each upsampling layer is accompanied by two residual blocks, maintaining the integrity of the feature representations while refining the segmentation mask. The final convolution layer serves as the culmination of the segmentation process, pivotal for generating segmentation masks that delineate various regions of interest within MRI images, including the whole tumor, core tumor, and enhancing tumor. Through convolutional operations, this layer transforms refined feature representations from earlier layers into segmentation masks, utilizing learnable filters to extract spatial patterns indicative of different classes or regions within the input volume. By incorporating the softmax activation function, the output values are normalized across multiple classes, ensuring that the predicted probabilities sum up to one for each voxel in the segmentation mask. This normalization facilitates the interpretation of model predictions as confidence scores for

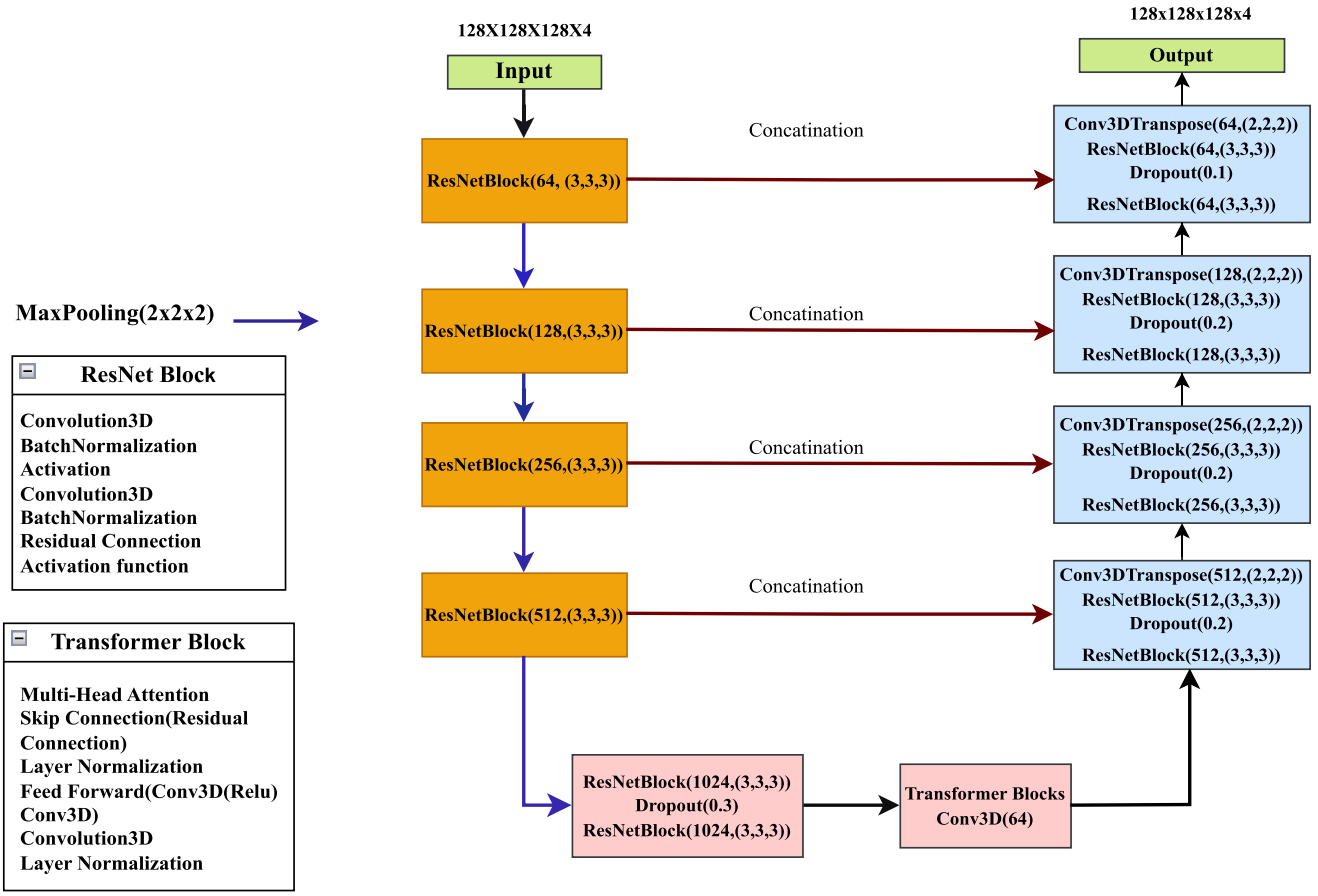


FIGURE 2. Proposed ResNet-UNet with transformer blocks for brain tumor segmentation.

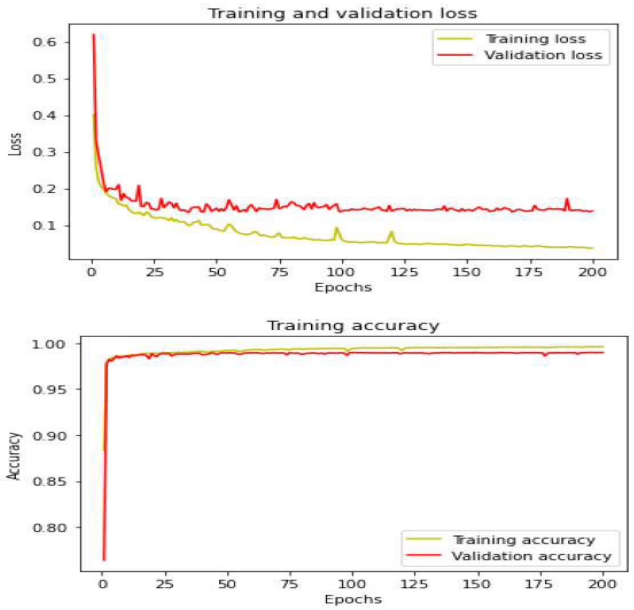


FIGURE 3. Training performance of proposed model.

each class, enhancing the reliability and interpretability of the segmentation results.

Furthermore, we strategically employ dropout layers to counteract the risk of overfitting and bolster the model's capacity for generalization. dropout, a widely used regularization technique, randomly deactivates a portion of neurons during training, thereby discouraging the network from relying too heavily on specific pathways or features that may be present in the training data but not necessarily generalize to unseen samples. Throughout the architecture, we strategically position dropout layers to introduce stochasticity and promote robust learning. In the encoder section, a dropout layer with a dropout rate of 0.3 follows the first residual block, while in the transformer blocks and subsequent layers of the decoder, dropout layers with rates of 0.3 and 0.2 are implemented. These dropout layers encourage the model to learn more invariant and representative features, ultimately enhancing its performance on segmentation tasks. Additionally, prior to the final output layer, a dropout layer with a rate of 0.1 is applied, further refining the model's predictions and guarding against overfitting, particularly crucial in tasks where precise delineation of boundaries is paramount. The flexibility of adjusting dropout rates based on dataset characteristics and model performance underscores the adaptability of our approach to diverse segmentation challenges. Through the strategic integration of dropout layers, we demonstrate improved robustness and generalization capabilities in our ResNet U-Net model with transformer blocks, paving the way for more effective and reliable segmentation results.

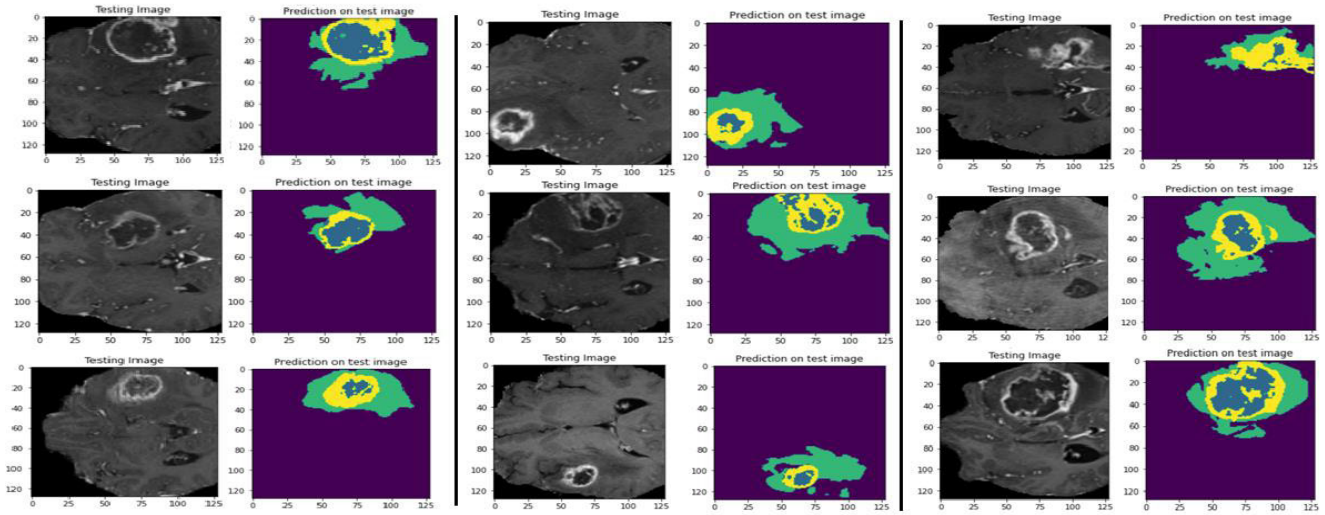


FIGURE 4. Proposed model's predicted segmentations on test images (Color-coded: Core-Blue, Enhancing-Yellow, Edema-Green, BACKGROUND-PURPLE).

TABLE 1. Hyperparameters and their corresponding values used in the experimental setup.

Hyperparameters	Values
Learning Rate	0.0001
The dimensionality of feed-forward network (ff_dim)	64
Batch Size	2
Number of Attention Heads	8
Dropout	0.3
Epochs	250

4) TRAINING

In the model training process, we begin by configuring the network using the widely favored Adam optimizer, a prominent choice in deep learning. We employ various evaluation metrics, encompassing dice coefficient [19], sensitivity, precision, specificity, accuracy, and the IOU (Intersection over Union) score, to comprehensively assess the model's performance in segmenting brain tumors.The hyperparameters of the proposed model are detailed in Table 1, while the overall performance metrics are presented in Table 2.To efficiently manage our dataset during training, we rely on a data generator. This generator dynamically loads preprocessed data as needed, a crucial practice, especially when working with large and complex datasets. Addressing the challenge of class imbalance within MRI segmentation data, we extend our approach beyond utilizing a custom dice coefficient loss function alone. Proactively, we employ the 'compute class weight' function from scikit-learn to calculate class weights. These weights are thoughtfully determined to establish a balanced weighting scheme for the different classes in the dataset. This deliberate weighting ensures consistent attention to minority classes or specific image regions, ultimately augmenting the model's capacity to precisely delineate these critical areas during the training process,the training performance is shown in Fig.3. Furthermore, using the proposed model, we generate predicted segmentation's for the test images, as shown in Fig. 4.

D. EVALUATION METRICS

The mathematical equation of various evaluation metrics are shown below:

a) Dice Coefficient: The Dice coefficient measures similarity between ground truth and predicted masks in image segmentation and is expressed as:

$$\text{Dice} = \frac{2 \cdot |A \cap B|}{|A| + |B|} \tag{9}$$

where:

A = represents ground truth mask
B = represents predicted mask
$|A \cap B|$ = intersection of ground truth and predicted masks
$|A|$ = number of elements in the ground truth mask
$|B|$ = number of elements in the predicted mask

b) Dice Coefficient Loss: The dice coefficient loss function is defined as:

$$\text{Dice\_Loss} = 1 - \text{Dice} \tag{10}$$

c) Sensitivity: Sensitivity measures the accuracy of an algorithm in diagnosing abnormal regions by quantifying the proportion of correctly diagnosed abnormal tissue and is calculated as:

$$\text{Sensitivity} = \frac{\sum(\text{TP})}{\sum(\text{TP} + \text{FN})} \tag{11}$$

where,TP & = true positives, FN = false negatives

*d) Specificity:*Specificity evaluates algorithm accuracy in identifying normal regions by assessing the proportion of correctly diagnosed normal tissue.

$$\text{Specificity} = \frac{\sum(\text{TN})}{\sum(\text{TN} + \text{FP})} \tag{12}$$

where,TN & = true negatives, FP = false positives

*e) Precision:*Precision evaluates algorithm accuracy in identifying relevant regions by measuring the proportion of correctly diagnosed abnormal tissue among all identified abnormalities.

$$\text{Precision} = \frac{\sum(\text{TP})}{\sum(\text{TP} + \text{FP})} \tag{13}$$

where,TP & = true positives, FP = false positives

*f) Accuracy:*Accuracy assesses overall correctness of an algorithm's predictions and is calculated as:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total Number of Predictions}} \tag{14}$$

*g) IOU-Score:*The IoU (Intersection over Union) score measures the overlap between predicted and ground truth regions. is given by:

$$\text{IOU} = \frac{\text{Intersection}}{\text{Union}} \tag{15}$$

IV. EXPERIMENTAL RESULTS

The segmentation network proposed is utilized on the MRI dataset for high-grade glioma (HGG) brain tumors from the BraTS 2019 database to precisely delineate tumor boundaries within the scans. After segmentation using our proposed algorithm, the resulting outputs undergo rigorous assessment against ground truth annotations using a diverse range of evaluation metrics. As depicted in Fig.4, we present screenshots showcasing a selection of input test images fed into the model, juxtaposed with the corresponding predicted segmentation outcomes. Notably, the predicted segmentation results are vividly displayed in color, offering a clear visual representation of the identified tumor areas. In contrast, the ground truth images are presented in grayscale, providing a reference point for comparison and validation. Our findings demonstrate the efficacy of the proposed network in accurately identifying all tumor regions within the brain scans, effectively pinpointing their precise locations. Specifically, the core tumor is depicted in blue, edema in yellow, and enhancing tumor regions in green. Moreover, the surrounding navy blue area serves to delineate the background, enhancing the overall clarity of the segmentation outcomes.The visual representation of proposed method's DSC with others is depicted in Fig.5

A. STATE-OF-THE-ART-TECHNIQUES

Table 3 provides a comparison between the proposed network and other methods, considering a range of pertinent criteria.

Chang et al. [14] introduced DPAFNet, a 3D segmentation model combining dual-path and multi-scale attention fusion modules. Dual-path convolution broadens network scale, while residual connections prevent degradation. The attention-fusion module aggregated global and local information and enhanced semantic understanding. A module employing 3D iterative dilated convolutions has been utilized to broaden the receptive field, enhancing contextual awareness.

Cao et al. [19] proposed MBANet, a three-dimensional convolutional neural network with a novel multi-branch attention mechanism. The network employs an optimized shuffle unit to construct the basic unit (BU) module, employing group convolution and channel shuffling for enhanced feature extraction. MBANet incorporates a unique multi-branch three-dimensional Shuffle Attention (SA) module for attention processing, facilitating both channel and spatial attention. Additionally, a 3D SA module is integrated into skip connections to improve resolution recovery.

Berkley et al. [21] suggested a three-dimensional U-Net model for segmentation of brain tumors employing their clinical dataset. The dataset contained MRIs with varied tumor types and resolutions. Ground-truth segmentations were provided by expert radiation oncologists.

Lu et al. [22] proposed GMetaNet, a three-dimensional multi-scale Ghost convolutional neural network with a MetaFormer decoding path. The proposed approach combined CNN's local modeling with the Transformer's long-range representation for semantic information extraction. GMetaNet introduces three novel Ghost modules: GSP, GSA, and DRG, optimizing multi-scale features and capturing long-range dependencies. A global decoder incorporating MetaFormer effectively merges local and global features. Deep supervision enhances convergence by ensembling outputs.

Peng et al. [23] suggested a novel glioma tumor segmentation method using AD-Net on multimodal MRI data. They utilized two learnable parameters for combining convolutional feature maps at different scales, dynamically adjusted through gradient backpropagation. The Jensen–Shannon divergence constrained feature map distributions, enhancing regularization across downsampling. Deep supervision training expedited model convergence.

Liu et al. [26] introduced a segmentation method leveraging learnable group convolution, reducing network parameters while enhancing communication between convolution groups. Incorporating skip connections between convolution modules improved segmentation precision. Deep supervision was employed to merge output images, reducing overfitting and enhancing recognition capabilities. Evaluation of the BraTS 2018 dataset revealed superior segmentation performance compared to leading methods in the challenge.

Raza et al. [32] introduced dResU-Net, a hybrid framework combining deep residual and U-Net models, for automating three-dimensional brain tumor segmentation (BTS). The model addressed the vanishing gradient issue, leveraging residual networks as encoders and U-Net as decoders to employ low- and high-level features simultaneously. Shortcut connections maintain lower-level features, whereas skip connections accelerate training.

Zhang et al. [33] introduced a novel framework for brain tumor segmentation from multimodal MRI data. The proposed approach utilized deep feature learning across modalities to address data scale limitations. The framework included Cross-Modality Feature Transition (CMFT)

TABLE 2. Performance metrics employed in our experimentation.

Dataset	Dice Coeff.	IOU Score	Loss	Accuracy	Label	Precision	Specificity	Sensitivity	DSC.
Training	0.97	0.92	0.02	0.99	Whole Tumor(WT)	0.99	0.99	0.99	0.96
					Core Tumor (CT)	0.96	0.99	0.97	0.96
					Enhancing Tumor(ET)	0.92	0.99	0.94	0.93
Validation	0.85	0.75	0.14	0.98	Whole Tumor(WT)	0.98	0.99	0.98	0.91
					Core Tumor (CT)	0.92	0.99	0.88	0.89
					Enhancing Tumor (ET)	0.82	0.99	0.87	0.84

TABLE 3. Comparative analysis of brain tumor segmentation methods.

Reference	Method	Dataset	Dice Similarity Coefficient		
			WT	CT	ET
Chang et al. [14]	DPAFNet	BraTS 2019	0.89	0.81	0.78
Yuan Cao et al. [19]	MBANet	BraTS 2018	0.89	0.85	0.80
He Huang et al. [20]	NLSE-VNet	BraTS 2019/BraTS 2020	0.87/0.88	0.74/0.79	0.70/0.73
Berkley et al. [21]	3D U-Net model	In-house clinical data	0.76	0.64	0.61
Yao Lu et al. [22]	GMetaNet	BraTS 2018	0.90	0.82	0.78
Yanjun Peng et al. [23]	AD-Net	BraTS 2020	0.90	0.80	0.76
Hengxin Liu et al. [26]	DMFNet	BraTS 2018	0.90	0.86	0.80
Raza et al. [32]	dResU-Net	BraTS 2021	0.86	0.84	0.82
Dingwen Zhang et al. [33]	Cross-Modality DFL	BraTS 2018	0.90	0.83	0.79
Proposed Study	ResNet-Transformer	BraTS 2019	0.91	0.89	0.84

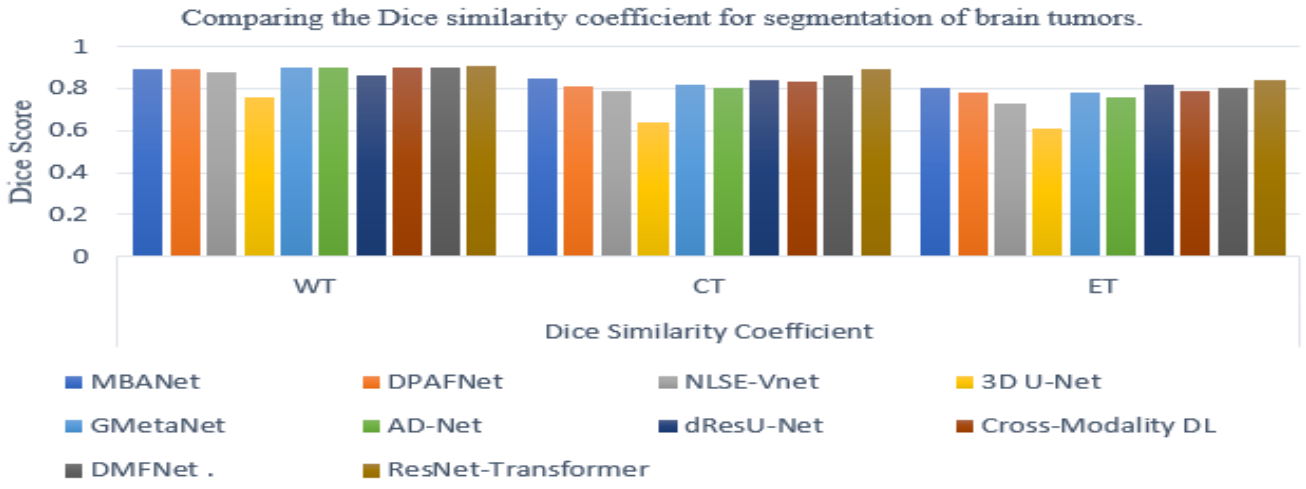


FIGURE 5. Comparative assessment of dice similarity coefficient (DSC) for brain tumor segmentation techniques.

and Cross-Modality Feature Fusion (CMFF) processes. CMFT transfers knowledge between modalities, while CMFF integrates knowledge for comprehensive representations.

V. CONCLUSION AND FUTURE SCOPE

Precise identification of glioma brain tumors in MRI images is essential for accurate diagnosis, treatment strategy, and evaluation of therapy effectiveness. Manual segmentation is laborious and subject to variability, highlighting the necessity for more accurate automated approaches, particularly considering the diverse geometries of tumors, spatial variations, and intricacies of MRI imaging. Advancements in deep learning have the potential to transform glioma segmentation, enhance accuracy, and eventually improve patient prognosis and quality of life. This paper presents a deep learning architecture tailored for accurate segmentation of glioma brain tumors. We propose a novel framework that includes many preprocessing processes and an advanced model architecture. Our preprocessing method involves bias field correction, z-score normalization, stacking, cropping, and random patch extraction from multi-channel MRI volumes. Firstly, bias field correction is employed to mitigate intensity changes induced by magnetic field inhomogeneities, thereby ensuring consistency across images. Additionally, Z-score normalization is implemented to standardize pixel intensities to a common scale, facilitating improved convergence during model training. Moreover, the process of stacking involves merging multiple MRI modalities (e.g., T1-weighted and T2-weighted) to provide complementary data for tumor delineation. Furthermore, cropping is utilized to isolate

specific regions of interest within MRI volumes, reducing computational overhead and enhancing model performance. Finally, random patch extraction, after cropping, aids in creating varied training samples, improving the model's capacity to generalize to unfamiliar data. Our suggested methodology combines ResNet blocks and transformer blocks to efficiently address the challenges of glioma segmentation. ResNet blocks facilitate deep feature extraction by addressing the vanishing gradient issue, enabling the model to capture essential hierarchical characteristics for tumor delineation. Transformer blocks, based on the transformer design used in natural language processing, help in capturing long-distance relationships inside MRI volumes, improving the model's capacity to include spatial and contextual data. The model attains a notable accuracy of 98% by employing the NVIDIA Tesla V100 32GB GPU. It effectively delineates gliomas and their subregions, offering vital information for diagnosis and treatment planning.

While the proposed model offers several advantages, including enhanced segmentation accuracy through the integration of local and global contextual information, it also presents limitations such as its applicability in real-world scenarios, exploration of a limited dataset, potential computational complexity, and lack of longitudinal data analysis. To overcome these limitations, future efforts could entail conducting clinical trials or case studies to integrate the model into diagnostic workflows. Collaborating with medical experts would facilitate evaluating its impact on treatment planning and patient prognosis. Additionally, qualitative assessments could offer insights into clinicians' perspectives and the challenges of implementing the model in routine clinical practice. Furthermore, utilizing additional BraTS datasets and exploring diverse glioma MRI datasets beyond the BraTS series could enhance the model's generalizability and robustness. Diversifying the dataset is expected to enhance the model's performance in real-world scenarios. Moreover, optimizing network architectures and employing techniques like data augmentation and transfer learning could mitigate the computational demands stemming from integrating ResNet and Transformer blocks. Further research might involve collecting longitudinal data, annotating tumor changes, and analyzing glioma progression and treatment response, thereby improving clinical outcomes by enhancing our understanding of tumor dynamics.

ACKNOWLEDGMENT

The authors would like to thank Prince Sultan University for their support.

REFERENCES

[1] S. S. Rangarajan, C. K. Shiva, A. Sudhakar, U. Subramaniam, E. R. Collins, and T. Senjyu, "Avant-garde solar plants with artificial intelligence and moonlighting capabilities as smart inverters in a smart grid," *Energies*, vol. 16, no. 3, p. 1112, Jan. 2023.

[2] S. Ullah, J. Ahmad, M. A. Khan, M. S. Alshehri, W. Boulila, A. Koubaa, S. U. Jan, and M. M. Iqbal Ch, "TNN-IDS: Transformer neural network-based intrusion detection system for MQTT-enabled IoT networks," *Comput. Netw.*, vol. 237, Dec. 2023, Art. no. 110072.

[3] N. Rasool and J. I. Bhat, "Unveiling the complexity of medical imaging through deep learning approaches," *Chaos Theory Appl.*, vol. 5, no. 4, pp. 267–280, Dec. 2023.

[4] N. A. Wani, R. Kumar, and J. Bedi, "DeepXplainer: An interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence," *Comput. Methods Programs Biomed.*, vol. 243, Jan. 2024, Art. no. 107879.

[5] E. Chikhaoui, A. Alajmi, and S. Larabi-Marie-Sainte, "Artificial intelligence applications in healthcare sector: Ethical and legal challenges," *Emerg. Sci. J.*, vol. 6, no. 4, pp. 717–738, May 2022.

[6] X. Dong, R. Huang, X. Wei, Z. Jie, J. Yu, J. Yin, and X. Liang, "UniDiff: Advancing vision-language models with generative and discriminative learning," 2023, *arXiv:2306.00813*.

[7] P. Celard, E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. S. Vieira, and L. Borrajo, "A survey on deep learning applied to medical images: From simple artificial neural networks to generative models," *Neural Comput. Appl.*, vol. 35, no. 3, pp. 2291–2323, Jan. 2023.

[8] L. M. Wang, Z. K. Englander, M. L. Miller, and J. N. Bruce, "Malignant glioma," in *Human Brain and Spinal Cord Tumors: From Bench to Bedside. Volume 2*. Cham, Switzerland: Springer, 2023, pp. 1–30.

[9] N. Rasool and J. I. Bhat, "Glioma brain tumor segmentation using deep learning: A review," in *Proc. 10th Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, Mar. 2023, pp. 484–489.

[10] K. A. van Garderen, S. R. van der Voort, M. M. J. Wijnenga, F. Incekara, A. Alafandi, G. Kapsas, R. Gahrmann, J. W. Schouten, H. J. Dubbink, A. J. P. E. Vincent, M. van den Bent, P. J. French, M. Smits, and S. Klein, "Evaluating the predictive value of glioma growth models for low-grade glioma after tumor resection," *IEEE Trans. Med. Imag.*, vol. 43, no. 1, pp. 253–263, Jan. 2024.

[11] Y. Du, K. E. Pollok, and J. Shen, "Unlocking glioblastoma secrets: Natural killer cell therapy against cancer stem cells," *Cancers*, vol. 15, no. 24, p. 5836, Dec. 2023.

[12] W. Cao, L. Xiong, L. Meng, Z. Li, Z. Hu, H. Lei, J. Wu, T. Song, C. Liu, R. Wei, L. Shen, and J. Hong, "Prognostic analysis and nomogram construction for older patients with IDH-wild-type glioblastoma," *Heliyon*, vol. 9, no. 7, Jul. 2023, Art. no. e18310.

[13] S. R. van der Voort, F. Incekara, M. M. J. Wijnenga, G. Kapsas, R. Gahrmann, J. W. Schouten, R. Nandoe Tewarie, G. J. Lycklama, P. C. D. W. Hamer, R. S. Eijgelaar, P. J. French, H. J. Dubbink, A. J. P. E. Vincent, W. J. Niessen, M. J. van den Bent, M. Smits, and S. Klein, "Combined molecular subtyping, grading, and segmentation of glioma using multi-task deep learning," *Neuro-Oncol.*, vol. 25, no. 2, pp. 279–289, Feb. 2023.

[14] Y. Chang, Z. Zheng, Y. Sun, M. Zhao, Y. Lu, and Y. Zhang, "DPAFNet: A residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104037.

[15] M. F. Ahamed, M. M. Hossain, M. Nahiduzzaman, M. R. Islam, M. R. Islam, M. Ahsan, and J. Haider, "A review on brain tumor segmentation based on deep learning methods with federated learning techniques," *Computerized Med. Imag. Graph.*, vol. 110, Dec. 2023, Art. no. 102313.

[16] L. Fang and X. Wang, "Brain tumor segmentation based on the dual-path network of multi-modal MRI images," *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108434.

[17] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. Jafarzadeh Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Comput. Biol. Med.*, vol. 152, Jan. 2023, Art. no. 106405.

[18] P. Salome, F. Sforazzini, G. Brugnara, A. Kudak, M. Dostal, C. Herold-Mende, S. Heiland, J. Debus, A. Abdollahi, and M. Knoll, "MR intensity normalization methods impact sequence specific radiomics prognostic model performance in primary and recurrent high-grade glioma," *Cancers*, vol. 15, no. 3, p. 965, Feb. 2023.

[19] Y. Cao, W. Zhou, M. Zang, D. An, Y. Feng, and B. Yu, "MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104296.

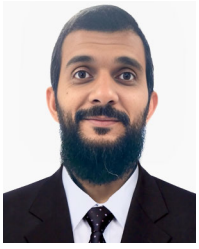
[20] H. Huang, W. Zhang, Y. Fang, J. Hong, S. Su, and X. Lai, "Overall survival prediction for gliomas using a novel compound approach," *Frontiers Oncol.*, vol. 11, Aug. 2021, Art. no. 724191.

[21] A. Berkley, C. Saueressig, U. Shukla, I. Chowdhury, A. Munoz-Gauna, O. Shehu, R. Singh, and R. Munbodh, "Clinical capability of modern brain tumor segmentation models," *Med. Phys.*, vol. 50, no. 8, pp. 4943–4959, 2023.
[22] Y. Lu, Y. Chang, Z. Zheng, Y. Sun, M. Zhao, B. Yu, C. Tian, and Y. Zhang, "GMetaNet: Multi-scale ghost convolutional neural network with auxiliary MetaFormer decoding path for brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104694.
[23] Y. Peng and J. Sun, "The multimodal MRI brain tumor segmentation based on AD-Net," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104336.
[24] S. R. Gonzalez, I. Zemmoura, and C. Tauber, "3D brain tumor segmentation and survival prediction using ensembles of convolutional neural networks," in *Proc. 6th Int. Workshop Brainlesion, Glioma, Multiple Sclerosis, Stroke Traumatic Brain Injuries*. Lima, Peru: Springer, Oct. 2020, pp. 241–254.
[25] M. Tamilarasi, "Performance analysis of glioma brain tumor segmentation using CNN deep learning approach," *IETE J. Res.*, vol. 69, no. 5, pp. 2400–2411, Jul. 2023.
[26] P. Santosh Kumar, V. P. Sakthivel, M. Raju, and P. D. Sathya, "Brain tumor segmentation of the FLAIR MRI images using novel ResUnet," *Biomed. Signal Process. Control*, vol. 82, Apr. 2023, Art. no. 104586.
[27] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: Improved n3 bias correction," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, Jun. 2010.
[28] N. Syamala and Y. Karuna, "Brain MRI image bias correction using generative adversarial network," *Soft Comput.*, pp. 1–13, Jun. 2023.
[29] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*. Munich, Germany: Springer, Oct. 2015, pp. 234–241.
[30] H. Sahli, A. Ben Slama, A. Zeraii, S. Labidi, and M. Sayadi, "ResNet-SVM: Fusion based glioblastoma tumor segmentation and classification," *J. X-Ray Sci. Technol.*, vol. 31, no. 1, pp. 27–48, Jan. 2023.
[31] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
[32] R. Raza, U. I. Bajwa, Y. Mehmood, M. W. Anwar, and M. H. Jamal, "dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 103861.
[33] D. Zhang, G. Huang, Q. Zhang, J. Han, J. Han, and Y. Yu, "Cross-modality deep feature learning for brain tumor segmentation," *Pattern Recognit.*, vol. 110, Feb. 2021, Art. no. 107562.

NOVSHEENA RASOOL received the B.Tech. degree in computer science and engineering from the University of Kashmir, in 2016, and the M.Sc. degree in information technology from the Islamic University of Science and Technology (IUST), Kashmir, Jammu and Kashmir, India, in 2018, where she is currently pursuing the Ph.D. degree with the Department of Computer Science. Her research interests include deep learning, machine learning, and medical image analysis. She has served as the first author on several papers indexed in Scopus and is actively engaged as a reviewer in biomedical signal processing and control journal.

JAVAID IQBAL BHAT received the Ph.D. degree in computer science from the University of Kashmir, Jammu and Kashmir, India. He is currently an Associate Professor and the Director of IT&SS with the Islamic University of Science and Technology (IUST), Kashmir, Jammu and Kashmir. His research interests include vision computing, social network analysis, medical image analysis, and data analytics. He has published more than 25 papers in his research-related field.

NIYAZ AHMAD WANI (Member, IEEE) received the bachelor's degree in computer science from the University of Kashmir, Jammu and Kashmir, India, in 2013, the Master of Science degree from the Islamic University of Science and Technology (IUST), Awantipora, Jammu and Kashmir, in 2016, focusing on the complexities of information technology. He is currently pursuing the Ph.D. degree in computer science with the Thapar Institute of Engineering and Technology, Patiala, Punjab. He has many publications that include internationally reputed journals, conferences, and workshops. His academic and research work established him as a potential scholar in the constantly changing field of computer science and technology. As he progresses in his Ph.D., his dedication to innovation and progress stays strong, paving the way for a future when his contributions to the field are expected to have a significant influence. His research interests include a wide range of topics, such as machine learning, deep learning, image processing, medical data analytics, and smart healthcare. He is serving as a Reviewer for various Elsevier and Springer journals, including *Artificial Intelligence Review*, *Multimedia Tools and Applications*, *Biochemistry*, and *Journal of Cell Biology*.



NAVEED AHMAD (Member, IEEE) received the degree in computer science from the Department of Computer Science, University of Peshawar, Pakistan, in 2007, and the Ph.D. degree in computer science from the Centre for Communication System Research (CCSR), University of Surrey, U.K., in 2013. He is currently an Associate Professor with the College of Computer and information Sciences, Prince Sultan University, Saudi Arabia. Prior to his current position, he was an Assistant Professor with the Department of Computer Science, University of Peshawar. He worked in the area of cyber security, privacy, blockchain technology, and penetration testing. He has managed three research and development grants related to blockchain, transport system for emergency vehicles, and platoon management. He has over 50 publications, that include international reputed journals, conferences, and workshops. He was a reviewer of various research proposals/grants funded by the U.K. Research and Innovation (UKRI). He also served on the program committee for various national conferences and workshops. He remained a reviewer of various IEEE, Elsevier, Springer, and MPDI journals. He has the honor of serving as the Guest Editor for a special issue in the IEEE Internet of Things Journal.

MOHAMMED ALSHARA received the M.S. degree in data science from The University of Texas at Austin, in 2011, and the Ph.D. degree in information technology from the University of North Texas, in 2016. He is an Associate Professor and the Dean of the College of Computer and Information Sciences, Prince Sultan University. He served as the Vice-Dean for academic affairs and the quality and developments and the Chair of the Information Technology Department, College of Computer and Information Sciences, Imam University, Riyadh, from 2016 to 2022. In addition, he spent two years (2008–2009) as a Data Governance Specialist with information technology company Compaq Middle East. His research interests include data science, information security, and information retrieval.