Machine Learning for Human Intelligence
2025, Vol. 3, Issue 1

Research Article

Enhanced Brain Tumor Segmentation Using Transfer Learning- Based Residual U-Net Architecture

Muhammad Saeed[1], Hanzla Ahmad[2], and Haris Naveed[3,⋆]

[1,2,3]Computer Science Department, NFC Institute of Engineering and Technology, 59030, Multan, Punjab, Pakistan.; Email: saeed217@gmail.com, hanzlaahmad39@gmail.com, harislagahri@gmail.com
⋆Corresponding author: Haris Naveed (harislagahri@gmail.com)

Article History

Academic Editor:
Dr. Muhammad Sajid
Submitted: December 21, 2024
Revised: February 01, 2025
Accepted: March 01, 2025

Keywords:
Image Segmentation; Convolutional Neural Networks; Residual Skip Connections; Knowledge Transfer; Machine Learning

Abstract

Beyond assisting medical practitioners in tumor detection and quantification, advanced imaging techniques contribute to the development of more effective treatment and rehabilitation protocols. Contemporary MRI- based brain tumor segmentation methodologies have emphasized U-Net architectural designs to integrate high-level semantic features with low-level spatial information for enhanced segmentation precision. Traditional fully convolutional networks employed for this application demonstrate limitations in successful image reconstruction via the decoder pathway due to insufficient low-level feature propagation from the encoder components. Enhanced optimization of low-level feature transmission from encoding to decoding pathways is essential for improved image reconstruction capabilities. This research proposes a transfer learning-enhanced residual U-Net framework that integrates U-Net and VGG-16 architectures. VGG-16 integration within the encoder pathway enhances image reconstruction performance. Furthermore, residual pathways within skip connections are incorporated to emphasize critical feature characteristics while suppressing noisy and irrelevant feature responses. The model undergoes training using The Cancer Imaging Archive (TCIA) and BraTS 2018 datasets, demonstrating improved performance in segmenting small-scale brain tumors. The proposed methodology exhibits competitive performance compared to existing brain tumor segmentation approaches.

1 Introduction

The advancement of medical technology enables healthcare professionals to develop more efficient diagnostic and treatment systems. Numerous medical domains benefit from e-healthcare implementations. Computer vision applications in biomedical imaging are gaining prominence due to their critical role in providing diagnostic information to radiologists for enhanced patient care. Various medical imaging modalities including X-ray, Magnetic Resonance Imaging (MRI), Ultrasound, and Computed Tomography (CT) significantly influence patient diagnosis and treatment protocols [1]. Brain tumors develop when abnormal cell clusters form within or adjacent to brain tissue, as illustrated in Figure 1. The rapid proliferation of malignant cells severely impacts patient health outcomes. The analysis, diagnosis, and treatment of brain imagery using sophisticated medical imaging technologies represent crucial research areas for physicians, radiologists, and clinical specialists. Brain imaging research is considered

Â© 2025 Copyright by the Authors.

fundamental since brain malignancies are lethal conditions responsible for substantial mortality rates in developing regions. For instance, approximately 29,000 individuals in the United States are diagnosed with brain tumors annually according to the National Brain Tumor Foundation, with 13,000 fatalities occurring yearly [2]. Specialized MRI techniques including MR Spectroscopy (MRS), Perfusion MR, and Diffusion Tensor Imaging (DTI) have been utilized for comprehensive brain tumor examination. Brain malignancies are categorized into two primary types: malignant (cancerous) and benign (non-cancerous) tumors. The World Health Organization classifies malignant tumors into Grades I through IV. Grade I and Grade II tumors are considered less aggressive, while Grade III and Grade IV represent malignant forms with severe health implications that may result in patient mortality.

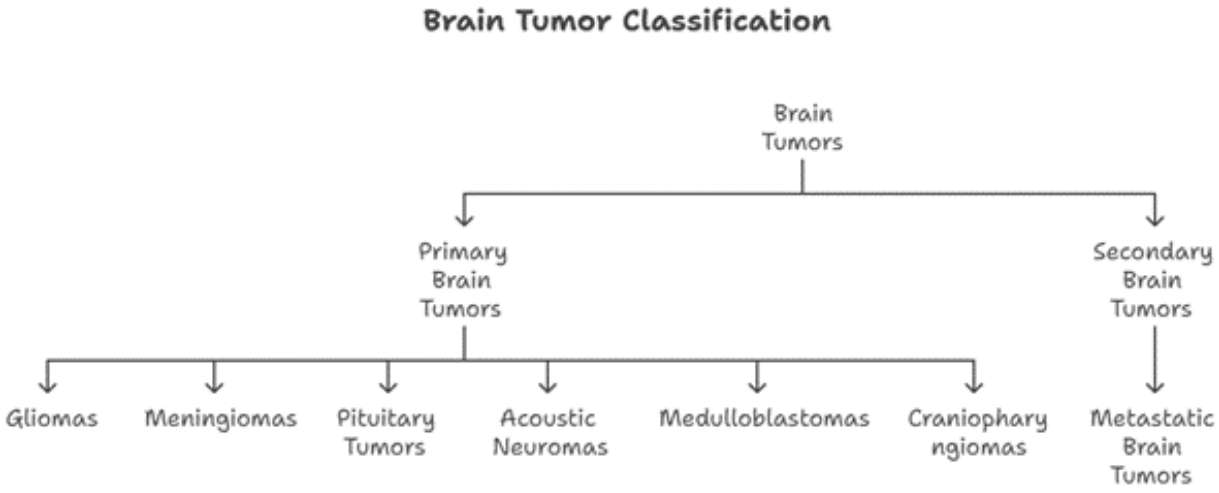

Figure 1: Classification of Brain Tumor Types.

Multiple imaging techniques and methodologies are employed for brain tumor management and prognosis [3]. Within image processing applications, segmentation constitutes the primary step for extracting abnormal brain regions from MRI scans, as demonstrated in Figure 2. Tumor region segmentation is essential for prediction, treatment planning, and therapeutic outcome evaluation. MRI encompasses various sequence techniques for segmentation including T1, T1c, T2-weighted, and FLAIR modalities. MRI data contains multiple characteristics advantageous for segmentation, including structural tensor eigenvalues, local histogram information, and texture features. Deep learning methodologies demonstrate superior performance in image analysis domains such as object detection, image classification, and semantic segmentation. Deep learning approaches have achieved high accuracy levels for automated brain tumor segmentation using multimodal MRI data. Convolutional Neural Networks (CNNs) primarily assist in patient classification, segmentation, and survival prediction for brain tumor cases. The development of deep learning frameworks for brain tumor analysis motivated our comprehensive examination of brain tumor research domains. Despite numerous recent efforts to support medical professionals, challenges remain in precision, robustness, and optimization [4]. Although various methodologies have been proposed for brain tumor segmentation with ongoing research efforts, further improvements are necessary.

Several challenges exist in the literature regarding brain tumor segmentation: rapid tumor development makes early-stage segmentation difficult, segmentation complexity arises from noise, unclear boundaries, unwanted background, irregular shapes, and intensity variations. Semantic gaps exist due to conventional skip connections in U-Net architecture that cannot effectively aggregate features; therefore, residual skip connections are employed for improved feature aggregation. Additional challenges include overfitting, class imbalance, and feature insufficiency. These complications make brain tumor identification challenging. Our objective is to propose a methodology that enhances image quality for accurate brain tumor segment detection and improves precision using state-of-the-art techniques.

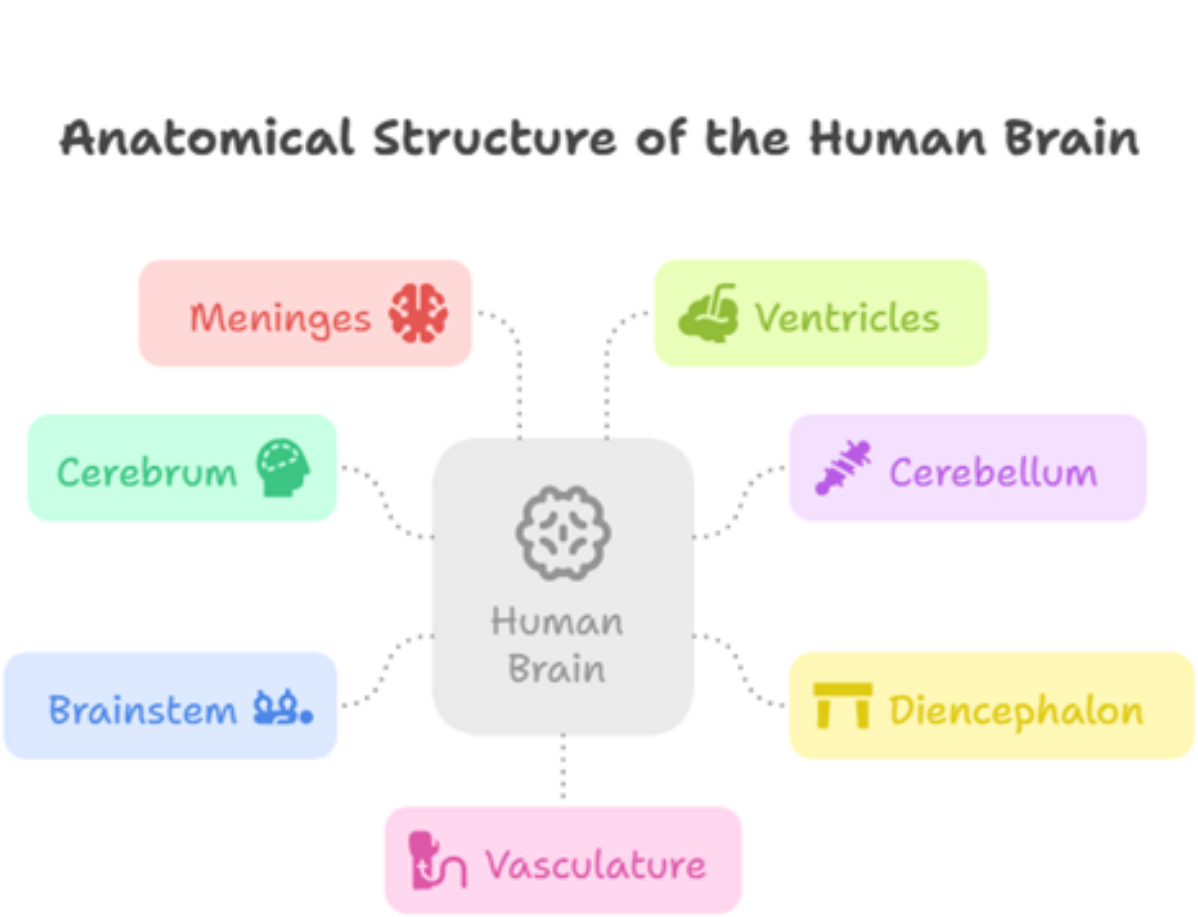

Figure 2: Brain Anatomical Structure.

1.1 Research Objectives

Numerous methodologies have been proposed for brain tumor segmentation with extensive ongoing research, yet further improvements remain necessary. Brain tumor segmentation presents challenges including noise interference, indistinct boundaries, unwanted background elements, complex geometries, and intensity fluctuations. Various difficulties associated with rapid tumor development make early-stage segmentation challenging. Our goal is to develop an approach that enhances image quality and precisely segments brain tumors while increasing accuracy through advanced methodologies.

1.2 Organization

The document organization follows: Section 2 describes previous research in brain tumor segmentation, Section 3 presents the methodology with detailed operational descriptions, Section 4 covers experimental setup and results analysis, and the final section provides conclusions summarizing the research.

2 Related Work

Various architectural designs have been proposed for brain tumor segmentation aimed at enhancing system accuracy. Tumor segmentation can be accomplished through pixel-wise classification or dense pixel classification approaches. This section examines novel strategies employed by researchers for brain tumor segmentation tasks. Authors in [5] utilized Wide Residual Networks (WRN) for automatic

glioma segmentation to extract features from multi-modal brain tumor samples. Subsequently, PPNet obtains global representations at various levels based on WRN features to scale recovery units, providing the network with original data again. This generates pixel-level predictions identical in size to input data, though limitations include overfitting issues and feature loss problems in WRN modules. When layers increase, performance degrades, leading to a fixed setting of 4 layers, which may not consistently provide optimal results for all images.

In [6], authors proposed a novel cascaded U-Net for brain tumor segmentation where complete tumors are segmented first, followed by inner tumor region segmentation. Residual blocks provide auxiliary supervision, facilitating gradient flow during training and transmitting high-resolution information for improved tumor localization from shallow to deep network layers. However, this approach encounters overfitting due to additional gradients causing model overfitting. The densely connected U-Net cascading architecture encodes and decodes tumor segments and sub-regions between dilated convolution layers for deep feature learning and improved tumor localization. While different augmentation techniques address overfitting issues for better segmentation accuracy, class imbalance problems remain unaddressed, limiting tumor accuracy improvements compared to previous works [7].

Subsequently, [8] employs this model for encoding identified lower-level characteristics and decoding obtained higher-level characteristics. This integrates with a newly developed Global Attention model for learning deep feature maps with high-level error precision. However, ignoring overfitting and class imbalance issues results in inadequate intra-structure segmentation of brain tumors. To address this, authors in [9] proposed a Middle Supervision Deep Residual Dilated Network (RDM-Net) for learning deep network layers without resolution loss while eliminating vanishing gradient problems. Various spatial fusion blocks preserve detailed information for small tumors, and middle supervision modules with proposed pyramid and multi-hierarchical loss reduce training time and information path distances, providing better segmentation results. However, this approach is unsuitable for different input image sizes and loses contextual information.

In [10], a novel multimodal three-dimensional segmentation algorithm based on nested dilation networks (NDNs) enhances low-level features but is limited by class imbalance issues and inability to segment edema portions of brain tumors. Later in [11], inception modules allow networks to learn richer representations, though data imbalance affects the proposed model for whole-brain slices since tumor pixels are relatively few, impairing the system's segmentation capability. AGResU-Net models combining attention gate units and residual blocks in U-Net architecture prove useful for brain tumor segmentation. Since Two-dimensional U-Net models have limitations in fully utilizing 3D MRI data information, AGResU- Net loses significant background information and local details between different slices. Additionally, limited labeled images for Deep Learning systems create differences between brain tumor segmentation approaches and new MRI scanning systems [12].

This attention gate residual U-Net utilizes attention gate modules to modify skip connections, benefiting small-scale brain tumor segmentation and enhancing salient features. The 2D HTTU-Net architecture, an automated technique introduced in [13], includes two tracks accounting for various tumor sizes, each using different kernel sizes and varying convolution blocks. However, memory limitations exist due to multimodal MRI images and complex structures, with time-consuming training stages. The Multi-Features Refinement and Aggregation architecture in [14] creates new platforms for effective semantic feature information utilization through feature re-extraction and efficient data creation. For tumor enhancement, it performs poorly because the proposed end-to-end model determines entire images, while patches are primarily based on other counterpart techniques. Recently, [15] authors proposed automated brain tumor segmentation using encoder and decoder v-net for multi-scale feature extraction. Squeeze and excitation blocks added to each encoding and decoding stage focus on feature recalibration processes, enhancing useful feature maps and salient spatial locations to improve segmentation accuracy with reduced computation time. However, their model loses contextual information, reducing segmentation accuracy.

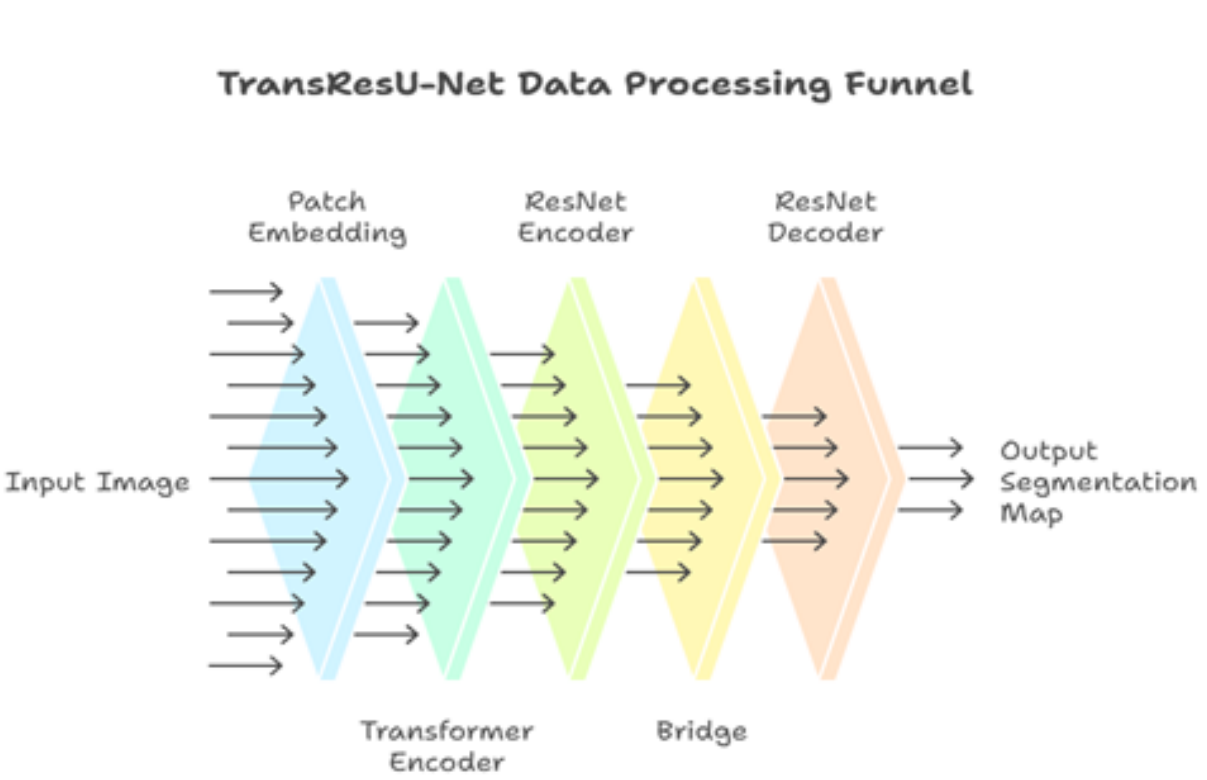

Figure 3: TransResU-Net Model Workflow Diagram.

3 Materials and Methods

This study proposes a strategy for brain tumor classification. Recently, numerous brain tumor segmentation techniques have been proposed, but improvements remain necessary. Brain tumor segmentation processes involve several complexities due to rapid development [16]. Tumor segmentation identification in initial stages represents a challenging task due to critical noise, unclear boundaries, unwanted context, complex shapes, and intensity variations. Semantic gaps exist due to conventional skip connections in U-Net architecture that cannot effectively aggregate features; therefore, residual skip connections are employed for improved feature aggregation. To address these issues, we propose the TransResU-Net model trained on brain tumor datasets. We first provide a brief model overview, then detailed discussions of various model modules in subsequent sub-sections [17]. The model utilizes transfer learning with residual skip connection U-Net architecture trained on the TCIA dataset as shown in Figure 3.

3.1 Enhanced U-Net Architecture

We propose a Trans U-Net architecture for model implementation by merging U-Net and VGG16 architectures with residual connections, implemented on Google Colab with GPU: 1xTesla K80, featuring 2496 CUDA cores and 12 GB GDDR5 VRAM space. This model was developed using brain MR images from The Cancer Imaging Archive dataset, which includes manual FLAIR abnormality segmentation masks. The Cancer Imaging Database (TCIA) provided the image collection, discussing 110 patients from lower-grade glioma patient groups with available fluid-attenuated inversion recovery (FLAIR) sequencing and genomic cluster data in The Cancer Genome Atlas (TCGA) [18], shown in Figure 4. Image samples are preoperative for 110 patients, providing MRI post-contrast, FLAIR, and pre-contrast samples. Manual segmentation on FLAIR MRI images produced corresponding tumor masks. MRI images are stored in TIF format files with corresponding manually segmented tumor masks available publicly. Missing series exist for some pre and post-contrast images (6 and 9 respec-

tively) including FLAIR sequences. With 256 x 256 pixel resolution, MRI data per patient varies from approximately 20 to 88 images, displaying brain boundary areas [19].

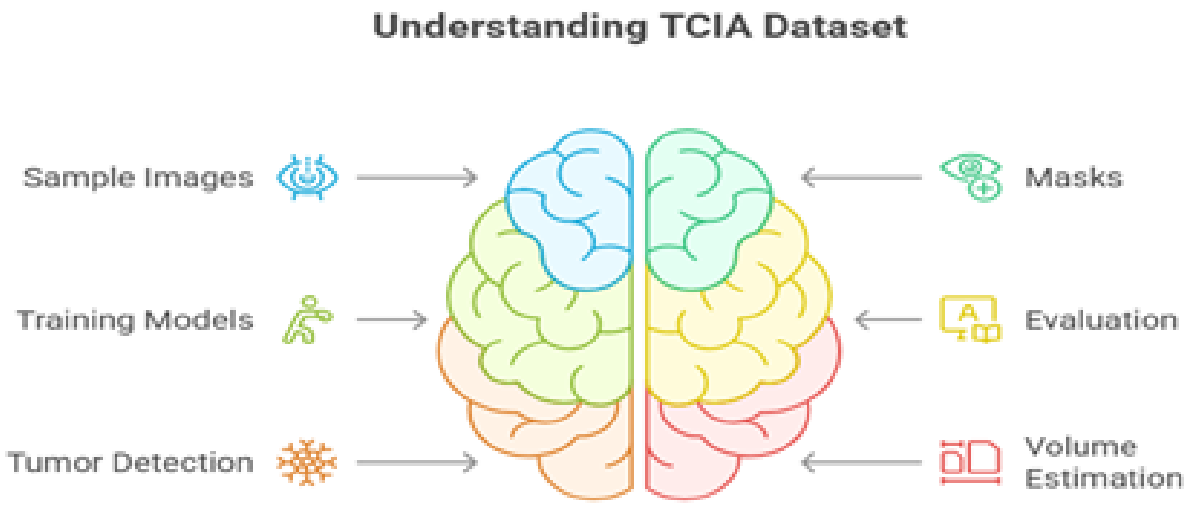

Figure 4: TCIA Dataset Sample Images with Corresponding Masks.

Images from the TCIA dataset undergo preprocessing where the initial stage crops unwanted regions from MRI scans referring to non-tissue areas, with each patient's corresponding tumor mask images cropped per patient. After cropping, to preserve image aspect ratios, cropped images are padded with zeros and resized to 256 x 256 pixels [20]. MRI images and tumor masks complete cropping, padding, and resizing processes. Intensity level scaling between 0 and 255 is applied, then pixel values are adjusted between 1 and -1, where -1 represents non-tissue regions. In all tumor masks, non-zero pixel values equal 1. After processing, 3121, 1373, and 1748 represent total numbers of images, tumor images, and normal tissue images respectively across all patients. Data augmentation techniques improve training data size by generating updated images from original datasets. This approach introduces variations that enhance model learning ability and better generalization for future unseen data [21]. The model generalizes by adding training dataset variations and becomes less prone to overfitting. Various data augmentation techniques are employed: horizontal flip set to true, height shift at 5%, rotation at 20%, width shift at 5%, shear at 5%, zoom at 5%, and fill mode set to nearest, as shown in Table 1. After applying preprocessing and data augmentation to MRI images, they proceed through TransResU-

Table 1: Data Enhancement Techniques (Model Implementation)

TECHNIQUES	**VALUES**
HORIZONTAL FLIP	True
HEIGHT ADJUSTMENT	0.05
ROTATION ANGLE	0.2
WIDTH ADJUSTMENT	0.05
SHEAR TRANSFORMATION	0.05
ZOOM FACTOR	0.05
FILL METHOD	Nearest

Net Architecture where actual tumor segmentation occurs. TransResU-Net Architecture consists of Encoder (upsampling), Decoder (downsampling), and Residual paths between encoder and decoder [22]. In TransResU-Net, pre-trained encoders are used in U-Net Architecture. Pre-trained models are employed in transfer learning approaches. Given VGG-Net and U-Net architecture similarities,

this study selects a hybrid of these powerful architectures where VGG-Net replaces U-Net's encoder path. Using pre-trained encoders enhances training processes. The VGG-16 model was pre-trained as an encoder on the TCIA dataset. Since lower levels focus more heavily on basic feature extraction, lower-level weights have greater impact on transfer learning assignments. Higher layer weights are dedicated to advanced feature extraction for target missions. Given these details, we utilize only the first 7 layers of VGG-16 in encoder modules to maintain lightweight architecture [23]. When training data is used, this UNet-VGG16 method freezes contraction layers to prevent weight layer changes. Instead, we utilize VGG16 model convolution layer weights. The goal is to reduce computational processes and shorten model training time. Encoder (downsampling) paths include convolutional layers relying on skip connections using three convolution layers. Network size remains minimal using 2×2 max-pooling functions. ReLU activation functions remove negative values from networks. Initially, grayscale images are converted to 3-channel images, then entered into main encoders for pre-trained encoder compatibility. The first two encoder blocks contain two sets of 3×3 convolutions with ReLU activation and max-pooling layers [24]. Final encoder blocks differ by having three 3×3 convolutions instead of two sets. Input layer size is $256 \times 256 \times 1$, and output layer size is $256 \times 256 \times 1$, representing sigmoid activation feature convolution layers. Image segmentation visualization processes under current proposed architectures for contracting VGG-16 layers in encoder paths.

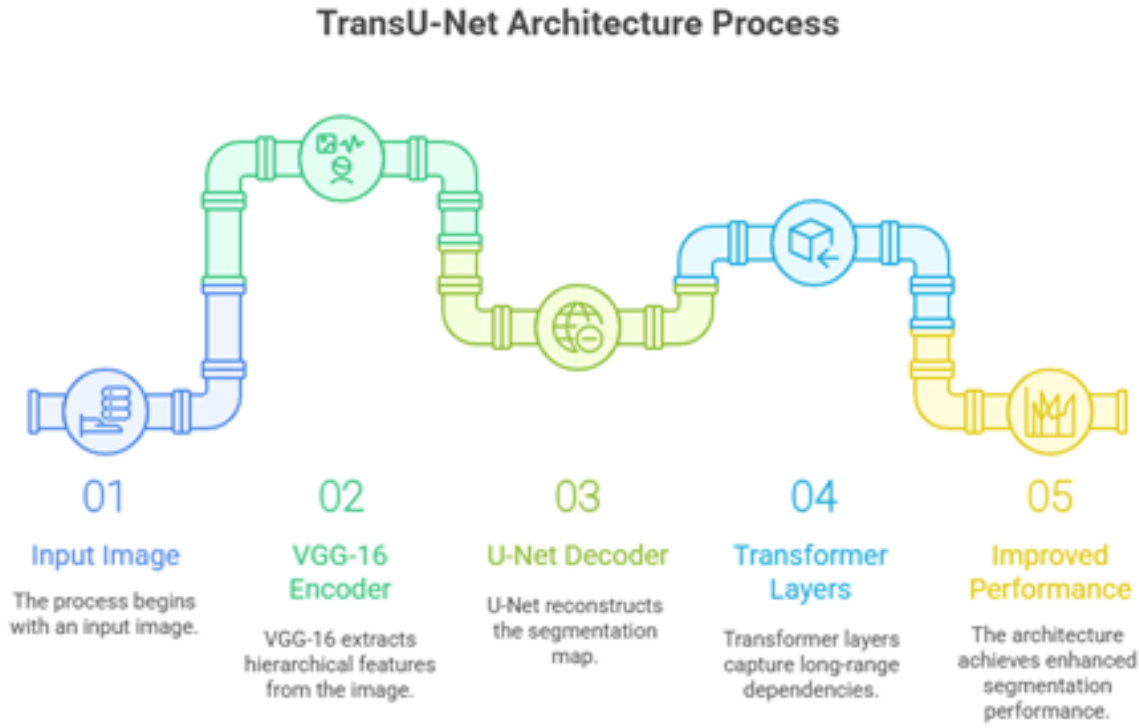

Figure 5: Encoder and Decoder Path.

Another module called Modified Skip Connection, termed Residual Skip Connection, is utilized. The innovative skip connection method propagates spatial details from encoder to decoder paths, introduced by U-Net architecture. Encoder layers until max-pooling and corresponding decoder layers after upsampling are typically connected by these connections [25]. Encoders and decoders suffer semantic gaps when original skip connections are used and feature aggregation is not performed accurately, despite preserving diffused spatial features. Basically, after final upsampling and before first max-pooling, decoders are connected by skip connections. In this context, despite undergoing less computation, encoder module features are considered relatively primitive features. Conversely, fea-

tures generated from decoder units are claimed to become more developed as they undergo extensive processing. Using convolution block chains within skip connections to eliminate semantic gaps is an excellent idea. Residual paths are connections that reduce these problems, then three decoder blocks from decoder modules as shown in Figures 5 and 6. Two 3×3 convolution layers with ReLU activation are used in both upsampling blocks. We use nearest neighbor upsampling as it is efficient for restricted items at desired output form surfaces. Consequently, relevant encoded features from residual

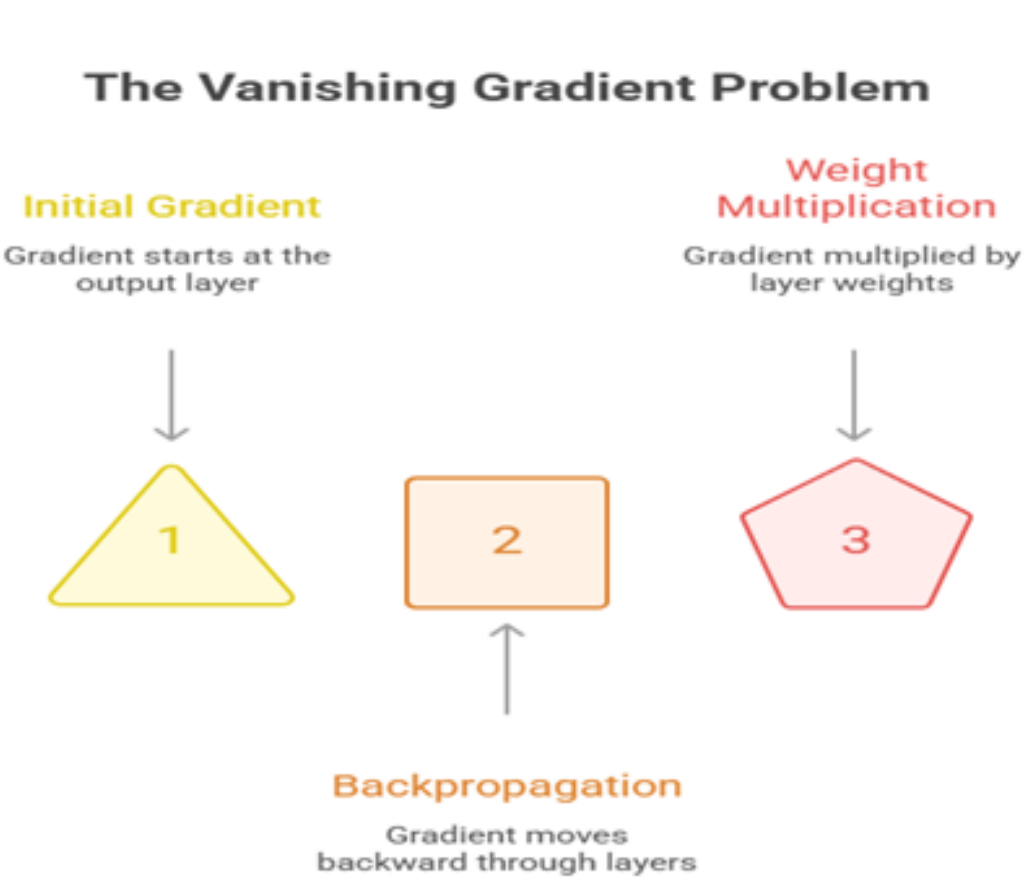

Figure 6: Residual Path used in Skip Connection.

connections are combined with features from final convolution layers within decoder blocks. Expected output masks are then produced by final decoder layers, together with 1x1 convolutions and sigmoid activation. In segmentation model FLAIR, masks are used in processed MRI data. Processed manual segmentation masks serve as ground truth for models. Processed images and masks are divided into training and test data for segmentation model training and validation. With learning rates equal to 5×10^{-5}, Adam optimizers are used. Loss functions are represented by model metrics, dice similarity coefficients (DSC), with negative values. Batch size was 8, with 40 different epochs. Image groups are imbalanced since normal tissue images overall exceed tumor images. We construct loss functions with weights representing ratios of pixel levels in manual segmentation masks provided for ground truth to reduce class imbalance concerns and improve tumor image mask estimation. In MRI images, segmentation models can recognize tumors at pixel levels by providing tumor mask predictions displaying tumor location, shape, and size. Tumor detection systems are proposed at image levels to identify image data as normal cells or tumors using corresponding tumor masks generated by modeling frameworks. MRI images are classified as normal tissue using this method if linked tumor masks have null values for every image pixel, and if related tumor masks have non-zero image pixels. Average DSC values are then obtained for test data. Complete model working using TransResU-Net.

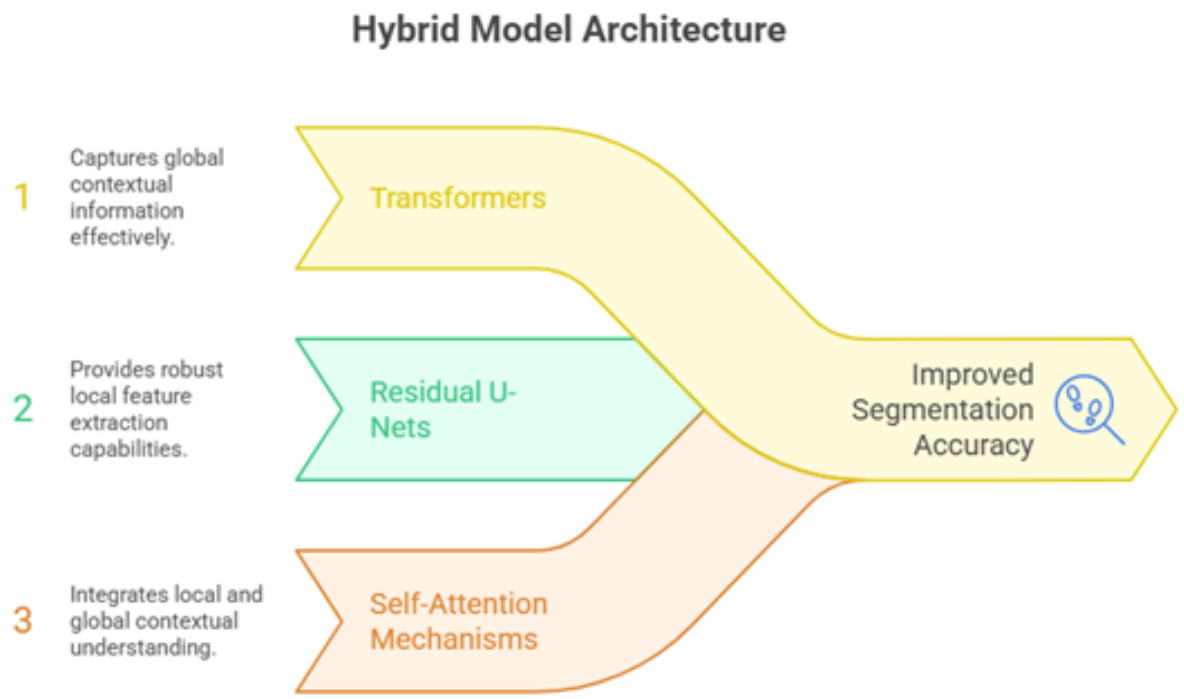

Figure 7: Complete TransResU-Net Model Architecture.

In training for parameter estimation, Adaptive Moment Estimator (Adam) optimization is employed in this study. Adam represents a hybrid of RMSProp and AdaGrad approaches. To preserve single learning rates for all weight changes, Adam uses stochastic gradient descent. Consequently, learning rates are guaranteed to remain constant during training procedures. Learning rates are preserved and separately modified as learning for each network weight (parameter), as shown in Figure 7. Loss function binary classification setups use binary cross-entropy. Binary cross-entropy blends sigmoid activation and cross- entropy loss. Results are displayed in graphs for both training and validation. Training and validation dice coefficient loss for the model is 0.0043, as shown in Table 2, with training and validation dice coefficients at 99%.

3.2 Performance Evaluation

For reducing class imbalance problems, we use two loss functions: WCE and GDL, abbreviated as Weighted Cross Entropy and Generalized Dice Loss. These are described below:

$$\mathrm{WCE} = \frac{1}{N} \sum_{i} \sum_{k=1}^{L} W_k g_{ik} \log(p_{ik}) \tag{1}$$

$$\mathrm{GDL} = 1 - \frac{2 \left(\sum_{i=1}^{L} w_i \sum_k g_{ik} p_{ik} \right)}{\sum_{i=1}^{L} w_i \sum_k (g_{ik} + p_{ik})} \tag{2}$$

where L represents total label numbers and K denotes batch size. Weight assigned to the i^th label is indicated by w_i. For network model assessment, we compared our proposed system's Dice Score with other latest architectures. Dice score values closer to 1 indicate greater precision, while values closer to 0 reflect less precision. Dice Scores reflect correlations between model-segmented images and Ground

Truth segmented images. In the 0-1 range, segmented maps (S) and ground truth (P) lie. Below, the Dice Score Equation is shown.

$$\text{Dice Score}(S, P) = \frac{2|S \cap P|}{|S| + |P|} \tag{3}$$

$$\text{Dice Score} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \tag{4}$$

Equation numerator (3) represents elements found in two sets, while subsets of sets belonging to value 1 are S and T. For network detection precision, Equations 4 and 5 represent dice scores at true positive and false negative accuracy levels. Accuracy is the most widely used metric for model evaluation, known as scores generated during class generalization. How accurately can models generalize themselves?

$$\text{Dice Score} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

As indicated in Table 3, we separated datasets into three sections for system training: training data,

Table 2: Training and Validation Performance Metrics

PARAMETER	TRAINING	VALIDATION
Dice Coefficient	99.4	99.1
Loss Value	0.0043	0.0041

validation data, and test data. Individual MRI patches were created and subsequently separated into training, validation, and testing data sets. We had initial learning rates of 0.01.

- We used ADAM optimizers for device initialization. Generally, Adam optimizers are used as substitutes for classical stochastic gradient missions. This enables networks to refine weights iteratively. We used ADAM optimizers because they are well-suited to large dataset challenges.
- We have cross-entropy for loss functions. For this assignment, binary cross-entropy fits best.

Table 3: Model Hyperparameters

Stages	Hyper Parameters	Values
Regularization	Batch Normalization(2D)	Mean=0 Standard Deviation=1.1
Parameters of System	Epochs	16
	Batch Size	16
	Optimizer	Adam=0.01

4 Results

We initially introduce three brain tumor segmentation benchmarks used for model assessment, followed by basic explanations of data pre-processing methods. Then, assessment criteria and implementation descriptions are provided in subsequent sub-sections. Finally, comparative experimental results of models are reported and addressed on three benchmarks.

4.1 Dataset Description

The Cancer Imaging Archive dataset was utilized in our research. This dataset includes brain MR images with manual FLAIR abnormality segmentation masks from TCIA. The Cancer Imaging Database (TCIA) provided image collections. They reference 110 patients who are part of lower-grade glioma patient groups with at least FLAIR sequences and genomic cluster data available in The Cancer Genome Atlas (TCGA). Tumor genomic clusters and patient information are found in data.csv files. Preoperative image samples are for 110 patients and include pre-contrast, post-contrast, and FLAIR images from MRIs. [34] performed manual segmentation and correspondingly generated tumor masks using FLAIR MRI images. Relevant manually segmented tumor masks and MRI images are offered in TIF format files, with MRI samples made publicly available for post-contrast and pre-contrast images respectively. Missing series exist for some pre and post-contrast images (6 and 9 respectively). It includes FLAIR sequences. With 256×256 pixel resolution, MRI data per patient differs from approximately 20 to 88 images, displaying brain boundary areas.

4.2 Data Processing Pipeline

For TCIA datasets, initial steps involve removing non-tissue portions from MRI by cropping corresponding tumor mask images from each patient's images. Following cropping, cropped images are padded with zeros to maintain image aspect ratios before resizing to 256 by 256 pixels. MRI images and tumor masks complete cropping, padding, and resizing processes. Intensity level scaling between 0 and 255 is applied, then pixel values are adjusted between 1 and -1, where -1 represents non-tissue regions in images. In all tumor masks, non-zero pixel values equal 1. After processing, 3121, 1373, and 1748 represent totals of all patient images, tumor images, and normal tissue images respectively.

4.3 Implementation Details

We conducted experiments for model 1 on Google Colab for proposed model evaluation. With GPU: 1 x Tesla K80, featuring 2496 CUDA cores, Google Colab provides 12 GB GDDR5 VRAM space. In TransResU-Net architecture for parameter estimation, Adaptive Moment Estimator (Adam) optimization is used in this research. In this research, we compared segmentation results of our TransResU-Net architecture with other state-of-the-art networks. Segmentation results of our proposed TransResU-Net architecture are shown: first is input images with provided segmentation masks (areas where tumors are present) and third images are final segmentation predicted results as shown in Figures 8 and 9. UNet (baseline) architecture represents 80% Dice Score. UNet models provide promising results for Dice scores and give top results for whole tumor accuracy. Other architectures using transfer learning have good dice scores at 84 percent but compared to our architecture, their accuracy lags at 92 percent. Other CNN architectures do not provide promising dice score and accuracy results at 78 and 71 percent respectively as given in Table 4. In Trans-VGG16 architecture, their accuracy is 96%. Our proposed architecture shows increases in terms of accuracy and dice score for whole tumors. In

Table 4: Performance Evaluation and Accuracy Assessment

METHODOLOGY	DICE COEFFICIENT	ACCURACY
U-Net (Baseline)	80	85
U-Net (DLA) [26]	82	87
Trans-Net [27]	84	92
CNN [28]	78	71
Trans-VGG16 [29]	-	96
TransResU-Net (proposed)	83	99.76

this research, our proposed architecture achieves promising results with other state-of-the-art architectures. We divided datasets into 4 parts where sets 3 and 4 achieve dice coefficients of 97 and 99 percent respectively. Wide Residual Network and Pyramid Pool Network (WRN-PPNET)-based multimodal

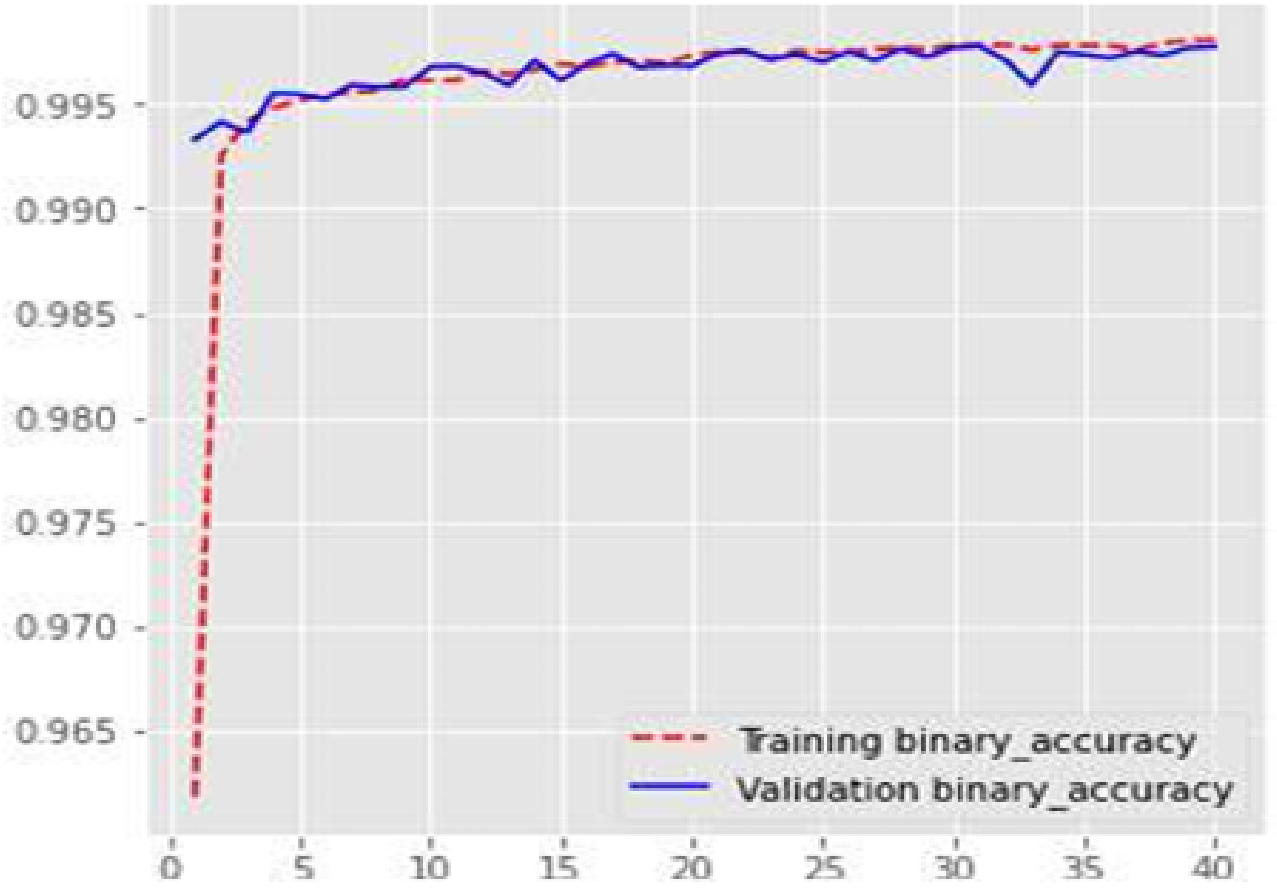

Figure 8: Training Accuracy.

brain tumor image segmentation compared to other approaches using pixel-wise prediction to segment glioma results in 91% accuracy. In other U-NET using Multi-Level Upsampling in decoders, they use multi- level upsampling to obtain more features resulting in 87% accuracy for segmented whole tumors. Nested Dilation Networks (U-NET) used dilated convolutions instead of plain convolutions to learn better features and improved tumor localization resulting in 70% dice coefficients because they ignore class imbalance problems. Further, Attention Gate ResU-Net and HTTU-Net achieve dice coefficients of 87 and 86 percent respectively. As comparison results given in Table 5, our method is better than other state-of- the-art methods, achieving 98 percent dice coefficient for whole tumors.

Table 5: Dice Score Performance Comparison

METHODOLOGY	DICE COEFFICIENT
Nested Dilation Networks (U-NET) [10]	70%
Attention Gate ResU-Net [12]	87%
HTTU-Net [13]	86%
ResU-Net (Proposed)	98%

5 Conclusion

Several recent studies have used fully convolutional networks for brain tumor segmentation; however, in these networks, decoder paths lack sufficient low-level features from encoder paths for proper image reconstruction. In this study, we proposed a transfer learning residual U-Net model with two architectures: U-Net and VGG-16. VGG-16 is integrated into encoders for better image reconstruction and adds residual paths in skip connections for illustrating salient feature details, thus disabling unwanted

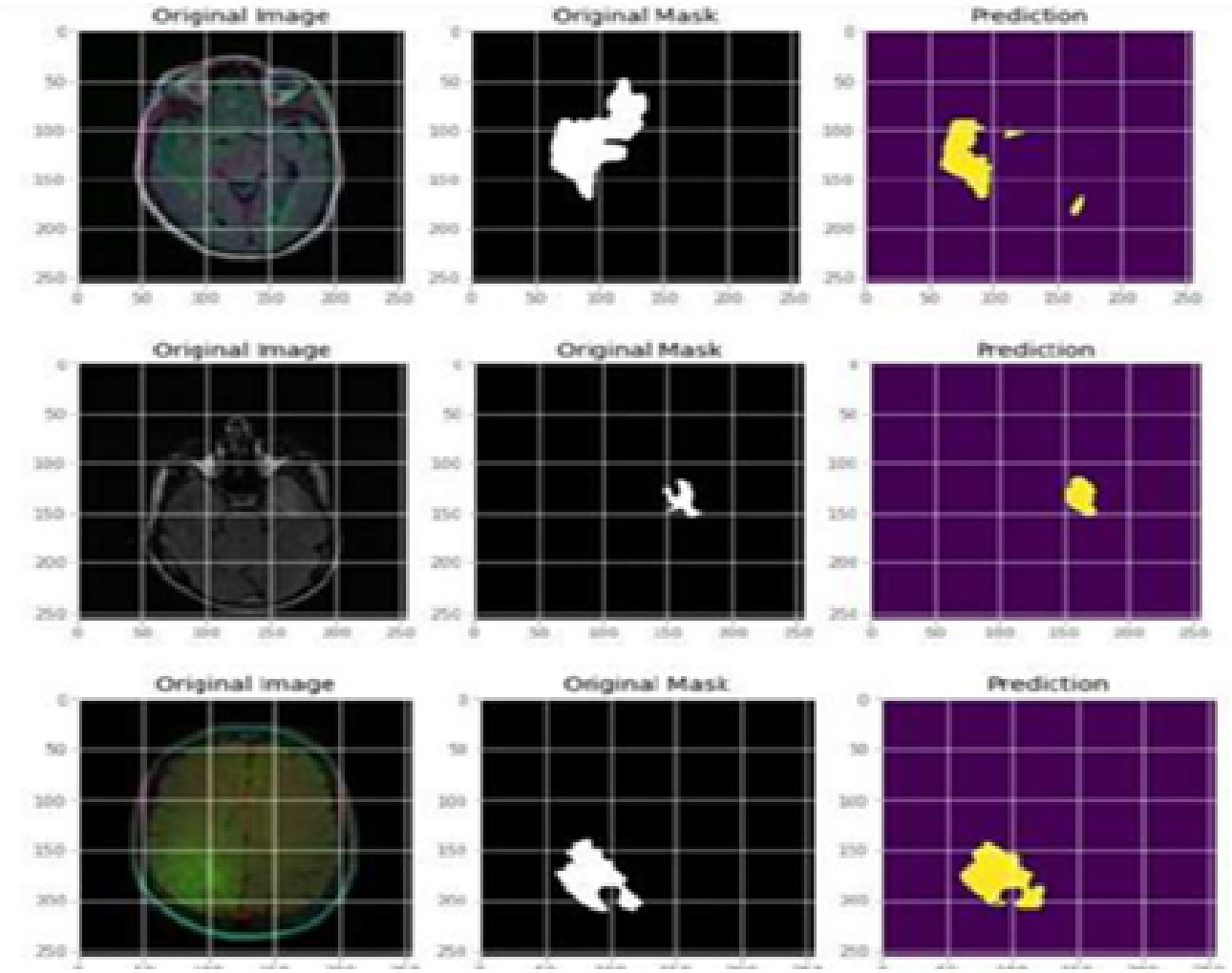

Figure 9: Segmentation Results.

and distracting feature responses that facilitate segmenting small brain tumors. These were trained on different datasets and achieve competitive performance with other brain tumor segmentation methods. In ResU-Net with modified residual skip connections and batch normalization layers, we use original convolutional layers. Experiments prove that ResU-Net achieves greater metric values than FCNN. When compared to existing U-Net models, proposed models perform better with lowest loss values and highest Dice Coefficient values. Segmentation outcomes using proposed methodologies frequently and extremely effectively address Region of Interest objectives for each brain tumor MRI image. By using 2D U-Net models, some contextual information is lost, so future work will explore 3D U-Net for better segmentation performance. Future research could use unique architectures or convolution block scenarios to find additional alternative models.

Funding

This research received no external funding.

Data Availability Statement

Data sharing does not apply to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Nadeem, M. W., Goh, H., & Hussain, M. (2020). Brain tumor analysis empowered with Deep Learning: A review, taxonomy, and future challenges. *Brain Sciences*, *10*(2), 118. https://doi.org/10.3390/brainsci10020118

[2] Davis, F. G., McCarthy, B. J., & Bondy, M. L. (2008). Issues of diagnostic review in Brain tumor studies: From the Brain Tumor Epidemiology Consortium. *Cancer Epidemiology, Biomarkers& Prevention*, *17*(3), 484–489. https://doi.org/10.1158/1055-9965.epi-07-0725

[3] DeAngelis, L. M. (2001). Brain tumors. *New England Journal of Medicine*, *344*(2), 114–123. https://doi.org/10.1056/nejm200101113440207

[4] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., ... & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, *35*, 18–31. https://doi.org/10.1016/j.media.2016.05.004

[5] Wang, Y., Li, C., Zhu, T., & Zhang, J. (2019). Multimodal Brain Tumor Image segmentation using WRN-PPNet. *Computerized Medical Imaging and Graphics*, *75*, 56–65. https://doi.org/10.1016/j.compmedimag.2019.04.001

[6] Liu, H., Shen, X., Shang, F., Ge, F., & Wang, F. (2019). Cu-Net: Cascaded U-net with loss weighted sampling for Brain Tumor segmentation. In *Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy* (pp. 102–111). Springer. https://doi.org/10.1007/978-3-030-33226-6_12

[7] Zhang, M., Liu, G., Tina, J., & Liu, Y. (2019, December). Improved U-net with multi-scale cross connection and dilated convolution for Brain tumor segmentation. In *2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE)* (pp. 1-5). IEEE. https://doi.org/10.1109/icmipe47306.2019.9098234

[8] Hu, Y., Liu, X., Wen, X., Niu, C., & Xia, Y. (2019). Brain tumor segmentation on multimodal MR imaging using Multi-level upsampling in Decoder. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (pp. 168–177). Springer. https://doi.org/10.1007/978-3-030-11726-9_15

[9] Ding, Y., Li, C., Yang, Q., Qin, Z., & Qin, Z. (2019). How to improve the deep residual network to segment multi-modal Brain tumor images. *IEEE Access*, *7*, 152821–152831. https://doi.org/10.1109/access.2019.2948120

[10] Wang, L., Zhang, T., Qin, Z., Shi, J., Zhao, P., Yang, D., ... & Ma, C. (2019). Nested dilation networks for brain tumor segmentation based on magnetic resonance imaging. *Frontiers in Neuroscience*, *13*, 285. https://doi.org/10.3389/fnins.2019.00285

[11] Li, H., Li, A., & Wang, M. (2019). A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Computers in Biology and Medicine*, *108*, 150–160. https://doi.org/10.1016/j.compbiomed.2019.03.014

[12] Zhang, J., Jiang, Z., Dong, J., Hou, Y., & Liu, B. (2020). Attention gate RESU-net for automatic MRI brain tumor segmentation. *IEEE Access*, *8*, 58533–58545. https://doi.org/10.1109/access.2020.2983075

[13] Aboelenein, N. M., Songhao, P., Koubaa, A., Noor, A., & Afifi, A. (2020). HTTU-net: Hybrid two track U-Net for Automatic Brain Tumor Segmentation. *IEEE Access*, *8*, 101406–101415. https://doi.org/10.1109/access.2020.2998601

[14] Wu, D., Ding, Y., Zhang, M., Yang, Q., & Qin, Z. (2020). Multi-features refinement and aggregation for medical brain segmentation. *IEEE Access*, *8*, 57483–57496. https://doi.org/10.1109/access.2020.2981380

[15] Liu, P., Dou, Q., Wang, Q., & Heng, P.-A. (2020). An encoder-decoder neural network with 3D squeeze-and-excitation and deep supervision for Brain Tumor Segmentation. *IEEE Access*, *8*, 34029–34037. https://doi.org/10.1109/access.2020.2973707

[16] Zhou, C., Ding, C., Wang, X., Lu, Z., & Tao, D. (2020). One-pass multi-task networks with cross-task guided attention for Brain Tumor segmentation. *IEEE Transactions on Image Processing*, *29*, 4516–4529. https://doi.org/10.1109/tip.2020.2973510

[17] Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014, September). Return of the devil in the details: Delving deep into convolutional nets. In *Proceedings of the British Machine Vision Conference 2014* (pp. 6.1-6.11). BMVA. https://doi.org/10.5244/c.28.6

[18] Cui, S., Mao, L., Jiang, J., Liu, C., & Xiong, S. (2018). Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. *Journal of Healthcare Engineering*, *2018*, 1–14. https://doi.org/10.1155/2018/4940593

[19] Rajinikanth, V., & Satapathy, S. C. (2018). Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy. *Arabian Journal for Science and Engineering*, *43*(8), 4365–4378. https://doi.org/10.1007/s13369-017-3053-6

[20] Tabatabai, G., Weller, M., & Wick, W. (2010). Molecular diagnostics of gliomas: The clinical perspective. *Acta Neuropathologica*, *120*(5), 585–592. https://doi.org/10.1007/s00401-010-0750-6

[21] Anitha, V., & Murugavalli, S. (2016). Brain tumour classification using two-tier classifier with adaptive segmentation technique. *IET Computer Vision*, *10*(1), 9-17.

[22] Buda, M., Saha, A., & Mazurowski, M. A. (2019). Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. *Computers in Biology and Medicine*, *109*, 218–225. https://doi.org/10.1016/j.compbiomed.2019.05.002

[23] Bhandari, A., Koppen, J., & Agzarian, M. (2020). Convolutional neural networks for brain tumour segmentation. *Insights into Imaging*, *11*(1), 77.

[24] Angulakshmi, M., & Lakshmi Priya, G. G. (2017). Automated brain tumour segmentation techniques—a review. *International Journal of Imaging Systems and Technology*, *27*(1), 66-77.

[25] Zheng, P., Zhu, X., & Guo, W. (2022). Brain tumour segmentation based on an improved U-Net. *BMC Medical Imaging*, *22*(1), 199.

[26] Naser, M. A., & Deen, M. J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Computers in Biology and Medicine*, *121*, 103758. https://doi.org/10.1016/j.compbiomed.2020.103758

[27] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, *35*(5), 1240–1251. https://doi.org/10.1109/tmi.2016.2538465

[28] Pravitasari, A. A., Hidayat, R., & Adiwijaya. (2020). UNET-VGG16 with transfer learning for MRI-based brain tumor segmentation. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *18*(3), 1310. https://doi.org/10.12928/telkomnika.v18i3.14753

[29] Atabansi, C. C., Chen, T., Cao, R., & Xu, X. (2021). Transfer learning technique with VGG-16 for near-infrared facial expression recognition. *Journal of Physics: Conference Series*, *1873*(1), 012033. https://doi.org/10.1088/1742-6596/1873/1/012033