ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
http://dx.doi.org/10.70102/afts.2025.1834.457

ENHANCED U-NET ARCHITECTURE METHOD FOR SEGMENTATION OF BRAIN TUMOR FROM MRI IMAGES

P.A. Monisha[1*], Dr.S. Sukumaran[2], Dr.G. Karthikeyan[3]

[1]Ph.D. Research Scholar, Erode Arts & Science College, Erode, Tamil Nadu, India.*
e-mail: amoni24@gmail.com, orcid: https://orcid.org/0009-0005-6809-1353
[2]Associate Professor & Head of Department, Erode Arts & Science College, Erode, Tamil Nadu, India. e-mail: prof_sukumar@yahoo.co.in,
orcid: https://orcid.org/0009-0002-0697-6517
[3]Assistant Professor, PSG College of Arts and Science, Coimbatore, Tamil Nadu, India.
e-mail: karthikeyan_g@psgcas.ac.in, orcid: https://orcid.org/0009-0006-3018-1983

Received: September 04, 2025; Revised: October 20, 2025; Accepted: November 25, 2025; Published: December 30, 2025

SUMMARY

Proper brain tumor segmentation is of great importance in the diagnosis and treatment planning. In this paper, the author presents an EfficientNet-modified U-Net-based system to segment the glioma in the pre-operative MRI scans using the BraTS 2018 dataset. This data consists of four types of MRI (T1, T2, T1Gd, and FLAIR). The model uses the EfficientNet-B6 as a tool to enhance the accuracy of feature extraction and segmentation by striking a balance between depth, width, and resolution in the network. The preprocessing of data is done by performing image transformation, subset division, and feature scaling, followed by feeding the data into the U-Net structure. The performance attributes of the model were assessed based on common measures such as Intersection over Union (IoU), precision, and recall. The presented model yielded an IoU of 92.7%, which is higher than other approaches, which got an IoU range of 81-91% on the BraTS dataset. Moreover, the accuracy and the recall of the model were 91.5% and 92.7%, respectively, which is by far better than other models like CAE (83.6 %, 83.2 %) and FCA (85.5%, 84.7 %). These findings indicate that the EfficientNet-enhanced U-Net architecture is effective in the accurate segmentation of brain tumors, and the diagnostic accuracy is improved. The model not only saves the cost of computations by utilizing the feature extracting capabilities of EfficientNet, which is efficient, but also offers automated support to clinicians, reducing human error. The solution is a powerful, scalable method to brain tumor segmentation in clinical practice, which helps in making a correct diagnosis and treatment decision-making.

Key words: *salient vessel candidate segmentation, image processing, deep learning, u-net architecture, brain tumor segmentation.*

INTRODUCTION

A brain tumor is a mass or development of abnormal cells in the brain, which may be malignant or noncancerous. Gliomas can be classified as either High-Grade Glioma (HGG) or Low-Grade Glioma (LGG) according to the pathological evaluation of the tumor. The objective of brain tumor segmentation is to differentiate between healthy tissue and tumorous regions [2]. This is an essential step in the analysis and treatment planning process to enhance the likelihood of effective treatment. In the present day, biological science has developed a variety of extended research problems that fall under the category of Digital Image Processing (DIP). Some applications that fall under this category include the detection and classification of tumors, the identification of malignant regions and their classification, and the testing and inspection of critical human body parts. Efforts are being made to effectively address the issue of automatic brain tumor segmentation and detection, which is of the utmost importance among the numerous medical science problems. A report that was approved by a January 2020 editorial on Cancer.net estimated that 23,890 adults in the United States, including 13,590 males and 10,300 females, will be diagnosed with malignant tumors of the brain and spinal cord this year. Magnetic Resonance Imaging (MRI) is typically the initial step in the diagnosis of brain tumors. Magnetic resonance segmentation is employed to separate the suspicious regions through a complex clinical imaging procedure. Brain tumor detection is still a manual process that is examined and validated by specialists, despite the fact that MRI is the most frequently employed method for highlighting the tumorous region in the brain [4][11].

Therefore, it is imperative that a compelling procedure is established for the early detection of brain tumors in order to ensure the safety of patients with a high degree of precision. The automatic detection of

Brain lesions are crucial due to certain deficiencies in the manual process. The primary motivation for computer- aided detection of anomalous tissue growth is the necessity to minimize human error and, as a result, achieve the highest possible level of accuracy. Valuable guidance for diagnosis and necessary treatment can be provided by the use of automated algorithms to make accurate and rigorous predictions of overall survival for patients diagnosed with gliomas. Additionally, there is a scarcity of competent neurologists and oncologists, particularly in developing nations. The country is currently experiencing a severe scarcity of caregivers, with only 2000 oncologists responsible for the care of approximately 10 million patients, according to a 2018 report. The medical field is unable to devise a conventional technique for the segmentation of brain tumors due to the presence of abnormalities [9]. In general, magnetic resonance imaging (MRI) scans are employed to accurately diagnose and examine brain malignancies [8][15]. The objective of this investigation is to employ magnetic resonance imaging (MRI) to detect brain malignancies in a systematic manner [1].

Consequently, it is essential to develop a compelling procedure for the early detection of brain tumors in order to guarantee the safety of patients with a high degree of precision [12]. The automatic detection of brain lesions is essential due to certain deficiencies in the manual process. The primary motivation for computer- aided detection of anomalous tissue growth is the need to reduce human error and, as a result, achieve the maximum possible level of accuracy. Automated algorithms can offer valuable guidance for diagnosis and necessary treatment by accurately and rigorously predicting the overall survival of patients diagnosed with gliomas. Furthermore, there is a shortage of qualified neurologists and oncologists, particularly in developing countries. According to a 2018 report, the nation is presently confronted with a severe shortage of caregivers, with only 2000 oncologists responsible for the care of approximately 10 million patients. The medical community is unable to develop a conventional technique for the segmentation of brain tumors as result of the presence of abnormalities. Magnetic resonance imaging (MRI) scans are typically implemented to precisely diagnose and investigate brain malignancies. The purpose of this study is to employ magnetic resonance imaging (MRI) in a systematic manner to identify brain malignancies [1].

Brain lesions are crucial due to certain deficiencies in the manual process. The primary motivation for computer- aided detection of anomalous tissue growth is the necessity to minimize human error and, as a result, achieve the highest possible level of accuracy. Valuable guidance for diagnosis and necessary treatment can be provided by the use of automated algorithms to make accurate and rigorous predictions of overall survival for patients diagnosed with gliomas. Additionally, there is a scarcity of competent neurologists and oncologists, particularly in developing nations. The country is currently experiencing a severe scarcity of caregivers, with only 2000 oncologists responsible for the care of approximately 10 million patients, according to a 2018 report. The medical field is unable to devise a conventional technique for the segmentation of brain tumors due to the presence of abnormalities. In general, magnetic resonance imaging (MRI) scans are employed to accurately diagnose and examine brain malignancies [13]. The objective of this investigation is to employ magnetic resonance imaging (MRI) to detect brain malignancies in a systematic manner [1].

Consequently, it is essential to have a compelling procedure for the early detection of brain tumors in order to guarantee the safety of patients with a high degree of precision. The automatic detection of brain lesions is essential due to certain deficiencies in the manual process. The primary motivation for computer- aided detection of anomalous tissue growth is the need to reduce human error and, as a result, achieve the maximum possible level of accuracy. Automated algorithms can offer valuable guidance for diagnosis and necessary treatment by accurately and rigorously predicting the overall survival of patients diagnosed with gliomas. Furthermore, there is a shortage of qualified neurologists and oncologists, particularly in developing countries. According to a 2018 report, the nation is presently confronted with a severe shortage of caregivers, with only 2000 oncologists responsible for the care of approximately 10 million patients. The medical community is unable to develop a conventional technique for the segmentation of brain tumors as a result of the presence of abnormalities [7]. Magnetic resonance imaging (MRI) scans are typically implemented to precisely diagnose and investigate brain malignancies. The purpose of this study is to employ magnetic resonance imaging (MRI) in a systematic manner to identify brain malignancies [3].

Segmentation of brain tumors on MRI images is a tedious task, subject to errors, and largely depends on qualified radiologists [6]. Although current deep learning architectures, such as U-Net, have provided good results, they have issues with the effective long-range dependency representation, which can be improved in terms of feature extraction efficiency and class imbalance, hence restricting their effectiveness in medical imaging problems. The current paper shows how these issues can be solved by suggesting an EfficientNet-improved U-Net [14]. The model makes some improvements by incorporating EfficientNet-B6 as the encoder to optimize feature extraction, to balance the scale resolution, and to lower the computational cost, hence surpassing the disadvantage of traditional U-Net models. It is an automated and reliable clinical application that can enhance segmentation accuracy and, therefore, provide timely and accurate diagnosis and plan treatment.

To extend the existing U-Net architecture with the segments of brain tumor, the paper presents a new EfficientNet-based architecture that combines the advantages of two models to enhance the accuracy of feature extraction and segmentation.

It shows high improvements in performance of segmentation with Intersection over Union (IoU) score of 92.7 %, and it performs better than existing methods in precision (91.5 %) and recalls (92.7 %)

The suggested method uses multi-modal MRI data, which offers a complete automated way of glioma segmentation, which is effective at helping clinicians properly diagnose and plan the treatment, and minimizes the influence of human error.

The paper has been divided into five broad sections that include the Introduction, where the clinical need for automated brain tumor segmentation in order to minimize human error is identified. This is preceded by Related Works, which is a review of the existing methods and the limitations they have with multi-modal MRI data. The main Proposed Work part of the paper describes the methodology, that is, the combination of the EfficientNet-B6 encoder with the U-Net and the required data preprocessing. The Results and Discussion section confirms the model with the help of the BraTS dataset, providing better metrics of performance, such as the IoU and precision, than the other models. Lastly, the Conclusion will define the success of the study and provide future research perspectives, such as explainable AI.

RELATED WORKS

MRI scans are extensively applied in the separation of brain structures, tumors, and lesions as they adequately depict brain tissue in high resolution and in detail. The first methods to automate this process were the traditional methods of segmentation, including thresholding, region-growing algorithms, and edge detection. Although these techniques were partially successful in the case of simple tumours, they were ineffective on the more complex and heterogeneous tumours. They were also very prone to human error and time-consuming since they had to be done manually. In order to address such limitations, the concept of deep learning, and, in particular, the Convolutional Neural Networks (CNNs), has been actively considered by the medical imaging community [10]. CNNs have also been shown to do some of the best automated MRI scan feature extraction, with much higher segmentation accuracy, with no human-engineered feature engineering.

An example is the research of [7], in which CNNs were employed in the classification of brain MRI scans and differentiating between benign and malignant tumors. The hybrid model suggested by the authors combines a Support Vector Machine (SVM) with CNN in terms of segmentation and classification, with a high accuracy of 98.4959. This combination method improves the recognition procedure by using threshold-based segmentation and maximizing the performance of the classification by using SVM integration. Besides, feature extraction was also done by means of AlexNet, and later achieved classification with Extreme Learning Machine (ELM), and optimized using the Amended Grasshopper Optimization Algorithm (AGOA). This model was resistant to noise and other resolutions, and it may have more uses in medical imaging. Nevertheless, even though it is highly accurate, this technique has difficulties in processing multi-modal MRI data as well as long-range dependencies, which restrict its capabilities to deal with complex tumor types and heterogeneous data [5].

Hybrid models have also been included in other research and involve the combination of CNNs with various machine learning methods in order to enhance segmentation precision. As an illustration, deep residual learning networks have been presented in brain tumor segmentation, in which the objective is to learn hierarchical features on the image data. When these models are used together with multi-modal image fusion, they demonstrate the potential of enhancing the efficiency of segmentation by fusing together various MRI modalities, including T1, T2, T1Gd, and FLAIR, that present complementary information about the tumor. The importance of multi-modal MRI data is also that they can be used to depict various tumor features that include shape, texture, and patterns of enhancement, which can be useful in the event of accurate segmentation [18]. Nonetheless, the currently available methods of utilizing multi-modal data also tend to have a problem with effectively integrating and processing this data, resulting in poor performance.

Recent research on deep learning models of brain tumor segmentation, including models based on the U-Net, has revealed that the encoder-decoder structure of the U-Net can be used to effectively capture the local features of the image, but it is weak in its ability to capture long-range dependencies and spatial gradients across layers of the image. It has also been observed that U-Net-based models are not able to process multi-modal MRI data or have a class imbalance in datasets where tumor regions occupy a minor part of the total image. As a solution to these problems, the architectural changes to the U-Net model were suggested, such as the addition of attention and residual connections [17][20]. The additions introduced into the model are used to concentrate on the most pertinent aspects and enhance its performance on complex datasets. Nevertheless, these still have issues with the computational power and multi-modal combination required to achieve successful tumor segmentation.

The more efficient solution is EfficientNet, which was developed by Google Research and uses the concept of scale of compounds, which also scales the depth, width, and resolution of the network to maximise its computation resources without losing its performance. EfficientNet has demonstrated good outcomes in image classification problems, both in state-of-the-art performance and in the minimization of computational cost. The EfficientNet architecture has been effectively used in other medical imaging activities like organ segmentation and tumor classification. Nonetheless, its usage with brain tumor segmentation is still constrained, especially when using multi-modal MRI data. The incorporation of EfficientNet into the U-Net framework of brain tumor segmentation has not been fully examined, even though it can theoretically enhance feature detection and computer power.

Existing traditional methods for brain tumor segmentation are limited in handling complex tumors and multi-modal MRI data, requiring manual intervention and leading to errors [16]. While hybrid CNN-SVM models have shown high accuracy, their fixed feature extraction and inability to efficiently integrate multi-modal data hinder performance. Multi-modal MRI data integration improves segmentation accuracy but remains challenging for current models. EfficientNet's compound scaling has demonstrated efficiency in other image classification tasks, but its application to brain tumor segmentation is underexplored [19]. This research addresses these gaps by integrating EfficientNet with U-Net, improving feature extraction, computational efficiency, and the ability to handle multi-modal MRI data, offering a more robust solution for accurate brain tumor segmentation.

PROPOSED WORK

This paper will introduce the integration of EfficientNet-B6 and U-Net to segment brain tumors with special attention to the preprocessing procedures as well as the hyperparameters applied in the process of model training. This research uses the BraTS 2018 dataset, which is in the form of multi-modal MRI scanning (T1, T2, T1Gd, and FLAIR). The data has a variety of glioma cases, and hence it can be used to obtain an effective model that would be able to segment tumors of different sizes, locations, and modes.

Data Preprocessing

Certain preprocessing procedures are enforced prior to feeding the MRI images to the model in order to standardize and augment the data. The raw images are then rescaled and made uniform in brush size (e.g., 128x128 or 256x256) so that the deep learning model has the same input dimensions. Subsequent normalization is then done, where pixel intensities are scaled to [0, 1] by dividing the individual pixel value by the highest pixel value in the image. This step will make sure that the model will be able to learn effectively and without the influence of photographic intensity changes.

In a bid to enhance the dataset and generalization, data augmentation algorithms are used. These are the random rotations, flips (horizontal and vertical), zooming, and elastic deformations to provide variability in the position and shape of tumors. Augmentation prevents overfitting as it has the effect of adding variety to the training data so that the model effectively learns to use robust features and not particular patterns in the images.

Further, the MRI images are cropped and padded to have the areas of interest (tumours) in the centre of the image to prevent the effect of irrelevant background information from interfering with the learning process of the model.

The following Figure (1) shows the structure of the proposed model of Enhanced U-Net brain tumor segmentation, where the efficient feature extraction is implemented by using EfficientNet-B6 as the encoder. MRI modalities, which include input of MRI, T1, T2, T1Gd, and FLAIR, are pre-processed, such as through augmentation and scaling. The encoder path makes use of several MBConv blocks comprising Squeeze-and-Excitation (SE) to derive features, whereas the decoder path marshals the spatial resolution by transposed convolutions. Skipping connections between the decoder and encoder tracks allows for maintaining spatial information, which allows for an accurate separation of tumor areas. The last layer has a 1x1 convolution with a softmax or sigmoid activation function to give the segmented tumor mask. This architecture is the best in terms of segmentation accuracy and computational efficiency of multi-mode MRI data.

The Unet Architecture

The convolutional neural networks, including U-Net, are used to perform image segmentation tasks. The two main elements of its symmetrical and U-shaped design are the decoder (expanding path) and the encoder (contracting path).

Encoder (Contracting Path)

Context and feature: The encoder is in charge of the segmentation of the input image. This consists of many convolutional cells, with each convolutional cell comprising two 3x 3 convolutions, a rectified linear unit (ReLU), and a 2x 2 max pooling operation with a stride of one. The number of feature channels is doubled by each block, and this enables the capture of more complex characteristics at lower spatial resolutions.

Bottleneck There is a bottleneck layer at the bottom of the U, which is made up of two 3x3 convolutions, which are then succeeded by ReLU activations, without down-sampling. This layer is the one that gets the deepest features of the image at the lowest resolution.

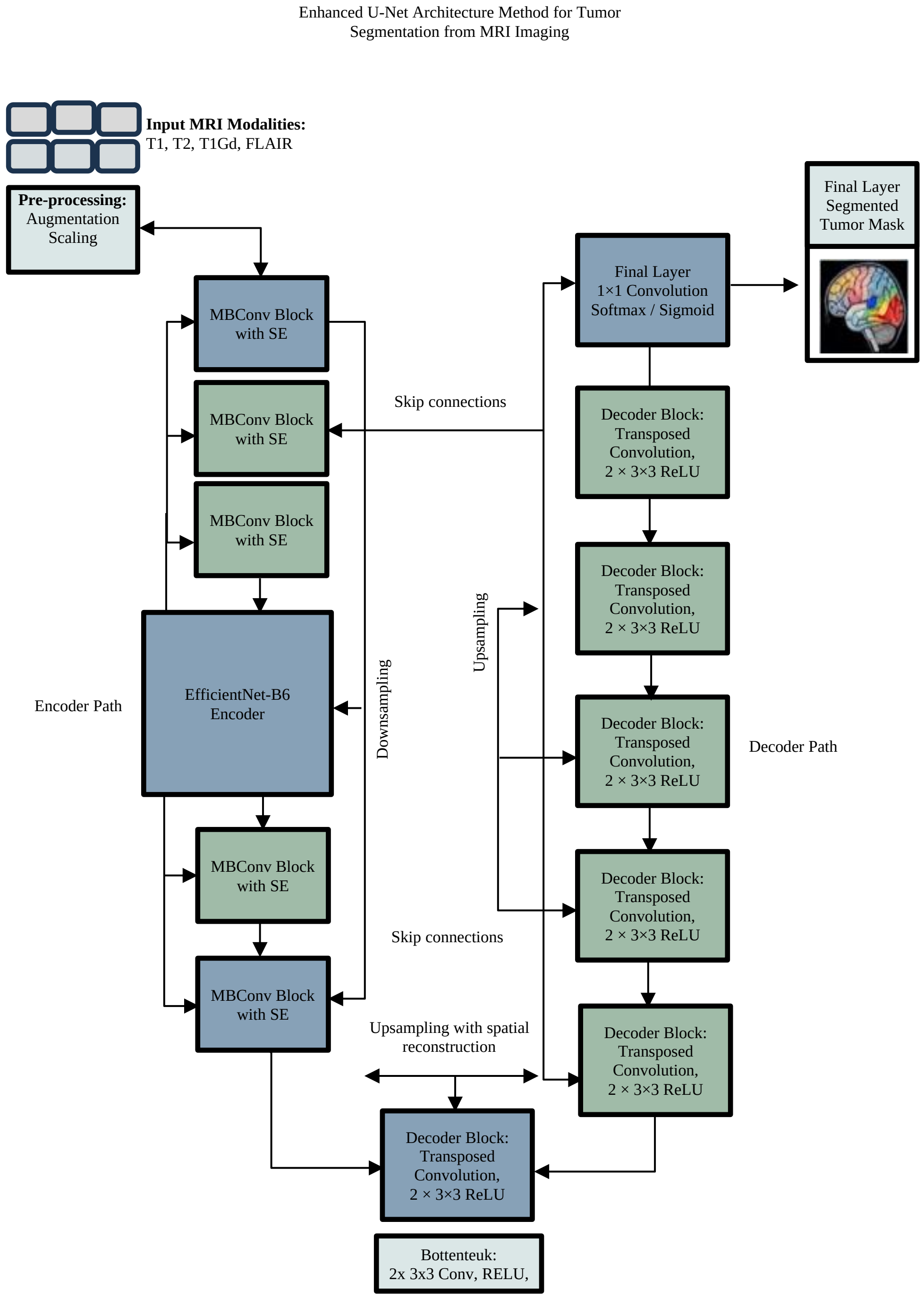

Figure 1. Improved U-net network to segment brain tumors using MRI images

Decoder (Expanding Path)

Decoder performs the upsampling and reconstruction of the spatial dimensions of the feature maps, as well as maintaining the high-level features that the encoder had previously acquired. The decoder also has the configuration of each block being based on an upsampling operation, which is normally implemented as a transposed convolution (also known as deconvolution) followed by two 3×3 convolutions with ReLU activations. With each upsampling step, the number of feature channels is halved.

Skip Connections

What makes the skip connections an important aspect of the U-Net design is that it is the direct connections of the levels of the encoder and the decoder that are relevant. The linkages help to retain spatial information, destroyed in downsampling, by concatenating the encoder to the decoder feature maps.

Final Layer

The final stage is a 1x1 convolution that reduces the size of the channels in the feature map to the desired number of output classes, followed by a softmax or a sigmoid activation function, depending on the desired task in the segmentation.

Efficient Net-B6

EfficientNet-B6 is part of an EfficientNet series of Google Research and tries to make CNN scans more efficient through a balanced scaling approach. Traditional scaling methods tend to increase the depth, width or resolution of a network individually and this often leads to sub-optimal performance and increased computational cost. EfficientNet can bypass this limitation by using compound scaling, a compound scale coefficient, ϕ, that uniformly scales all three dimensions together.

Compound Scaling

The primary innovation of Efficient Net-B6, which is called the compound scaling approach, scales the depth (d), width (w), and resolution (r) of the network by a combination of fixed scaling coefficients (alpha, beta, and gamma). The compound coefficient ϕ that is used to regulate the scale of each dimension is 6 in EfficientNet-B6. The accuracy of determining the optimal values of alpha, beta, and gamma is done through a grid search and aims to maximize the trade-off between the computation efficiency and the precision.

Architectural Building Blocks

EfficientNet-B6 is constructed using Mobile Bottleneck Convolution (MBConv) blocks, which are Squeeze-and-Excitation (SE)-optimal. Each MBConv block has: Depth-wise Convolution Phase that applies depth-wise separable convolutions, spatial convolutions done separately on each channel, and saves on computational expense.

Squeeze-and-Excitation (SE) Phase describes the recalibration of channel-dependent feature responses by forming channel interdependency models. The projection phase intends to reduce the channels into the required output size to allow effective information flow within the network. The EfficientNet-B6 may be concluded as follows:

Stem: Initial 3×3 convolution with 32 filters.

Stage: Referred to as repeated MBConv blocks, this stage has varying expansion factors, kernel sizes, and output channels.

Top: 1x 1 final convolution to achieve the required number of output classes, then a global average pooling layer, and finally no fully connected layer on the classification. Common training techniques of EfficientNet-B6 are batch normalization, dropout, and data augmentation. There is a learning rate schedule that helps in updating the model parameters with the Adam optimizer.

The balanced compound scaling of EfficientNet-B6 provides this model with a highly effective and computationally efficient model. It has better capabilities of feature segmentation because it combines MBConv blocks with SE and can thus be applied as an effective tool in computer vision applications such as image classification. EfficientNet-B6 is suitable for high-resolution, complex image analysis tasks as it balances well between quality and speed with a linear increase in depth and width, as well as increasing the input resolution. Squeegee and excitation (SE) module, the MBConv block is a key component of the EfficientNet-B6 architecture that gives this network high performance and efficiency. The MBConv block is a sophisticated block consisting of an inverted residual structure and depthwise separable convolutions. Depthwise convolution requires significantly fewer operations and parameters and, therefore, is less costly to compute because it replaces normal convolutions with unique filters applied to each of the input channels. The inverted residual structure also makes the input features expandable (with the help of a lightweight 1x1 convolution) and then uses a depthwise convolution and 1x1 convolution to cut the features, thereby enhancing efficiency. This also reduces the computation cost, and the network is permitted.

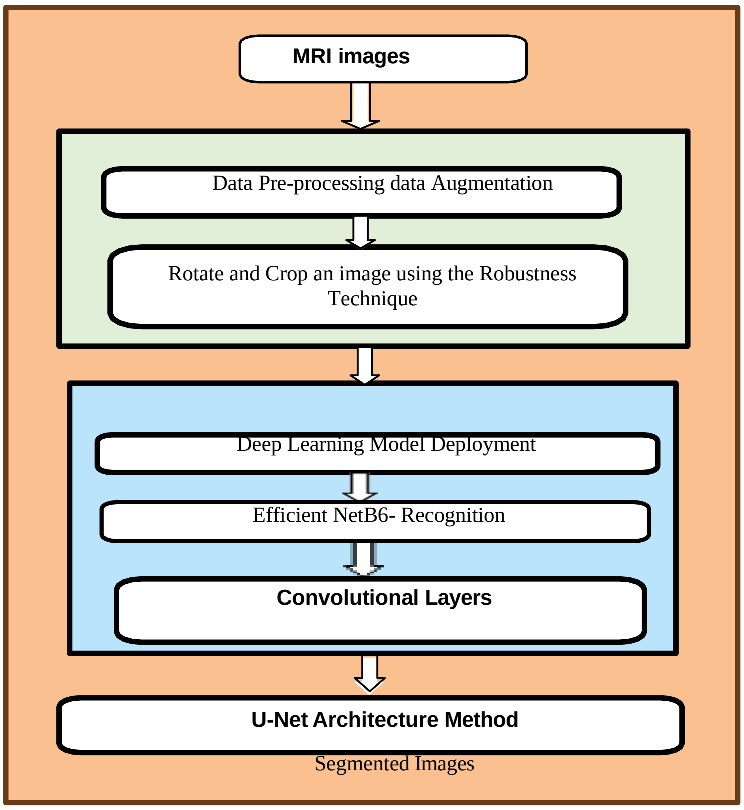

Figure 2. Flow Diagram of U-Net Architecture method

Figure 2 demonstrates the sequence of the suggested brain tumor segmentation model with the help of MRI images. The input is in the form of MRI images which are subjected to data pre-processing process and this involves data augmentation, image rotation and cropping taking advantage of a robustness technique. The images have been pre-processed, and finally, the stage of deployment of the deep learning model is carried out, and the EfficientNet-B6 model is deployed to carry out recognition, and the convolutional layers extract the features. The final stage makes use of the U-Net architecture method, which is used to obtain precise tumor segmentation. The result of this is the segmented images, which indicate the accurate morphology of the tumor areas in the MRI images.

Data augmentation produces a synthetic enlargement of the training dataset by generating modified versions of pre-existing data. Using deep learning algorithms or making little adjustments to the information, this method creates new data points.

Mathematical Description

The segmentation process can be mathematically described as follows:

Feature Extraction (EfficientNet-B6):

The input MRI image $I \in \mathbb{R}^{H\times W\times C}$, where $H$and $W$are the height and width of the image, and C is the number of channels (modalities). EfficientNet-B6 extracts hierarchical features by applying convolutions and pooling operations. The output feature map $F \in \mathbb{R}^{H'\times W'\times C'}$, where $H'$and $W'$are the reduced dimensions, and $C'$is the number of extracted features.

EfficientNet uses compound scaling, represented as:

$$\text{EfficientNet Scaling:} \{\alpha, \beta, \gamma\} = \phi \qquad (1)$$

In Equation (1), where $\phi$is a scaling coefficient and $\alpha, \beta, \gamma$are scaling factors for depth, width, and resolution, respectively. This scaling is applied to adjust the network to achieve optimal performance for the given dataset.

U-Net Decoder:

The decoder reconstructs the spatial resolution of the image using transposed convolutions (deconvolutions). At each layer l, the feature map from the encoder $F_l$is is passed through upsampling, followed by convolution:

$$F_l^{upsampled} = \text{ConvTranspose}(F_{l-1}) \oplus F_l \qquad (2)$$

In Equation (2) where $\oplus$denotes concatenation of the feature map from the encoder and the upsampled feature map.

Loss Function:

The model is trained using a weighted cross-entropy loss to handle class imbalance, given by:

$$L = -\sum_{i=1}^{N} w_i \cdot y_i \cdot \log(p_i) \qquad (3)$$

In Equation (3), where $N$is the number of pixels, $y_i$is the ground truth label (1 for tumor, 0 for non-tumor), $p_i$is the predicted probability for the tumor class, and $w_i$is the weight assigned to each pixel to counteract class imbalance.

Optimization:

The Adam optimizer is used to minimize the loss function:

$$\theta_t = \theta_{t-1} - \eta \cdot \frac{m_t}{\sqrt{v_t} + \epsilon} \qquad (4)$$

In Equation (4), where $\theta_t$represents the model parameters, m_t and v_t are the first and second moment estimates of the gradients, $\eta$is the learning rate, and $\epsilon$is a small constant to prevent division by zero.

Table 1 Algorithm of U-Net Architecture Method

Step 1:	Input image of dimensions Height× Width × Channels.
Step 2:	Pass the input image through the initial convolution and pooling layers. Perform expansion phase: $xexpanded=ReLU6(Conv2D1\times1(x))$
Step 3:	Perform the projection phase $xprojected(x)=Conv2D1\times1(xse))$
Step 4:	Save feature maps Fi from selected layers for skip connections.
Step 5:	Cross-Entropy loss to handle
Step 6:	convert probability maps

Step 7:	Pass the deepest encoder feature map through an to produce the bottleneck feature map B.
Step 8:	$Convproject(x)=Conv2D1\times1(x)$
Step 7:	U-Net model with an encoder of convolutional layers + max-pooling.
Step 8:	For each decoding convolutional layer
Step 9:	Perform upsampling $Ui=Conv2DTranspose(Di+1,filters=fi, kernal\_size=2,stride=2,padding='same')$
Step 10:	optimizers like Adam with learning rate scheduling $y=Conv2D(D0,filters=num\_classes,ker\ nel\_size=1,activation='softmax' ifmulti-classelse'sigmoid')$

In Table 1 Fi: In a neural network, this usually refers to the feature map or the output of the i-th layer. After performing a specific operation (like a convolution) in deep learning, Fi can represent the activations from the i-th layer.

filters = fi: The set of weights or kernels used to carry out the convolution operation is referred to as filters (typically represented by the symbol fi) in the context of a convolutional layer.

The predicted region and *Agt* are the ground truth. Perfect overlap is indicated by an IoU value of 1, which ranges from 0 to 1. Better model performance is indicated by a higher IoU. IoU is calculated for every class in semantic segmentation and averaged (Mean IoU). To identify true positives in object detection, a standard threshold (such as 0.5) is applied. IoU efficiently evaluates prediction accuracy, which aids in model optimization.

Sensitivity**,** also known as the True Positive Rate (TPR) or Recall, is a measure used in classification problems, especially in medical testing and machine learning, to assess the performance of a test or model in identifying positive instances correctly (Equation 5). Recall is defined as:

$$Recall = \frac{True\ Positives(TP)}{True\ Positives(TP)+False\ Negatives(FN)} \qquad - \ (5)$$

Specificity**,** also known as the True Negative Rate (TNR)**,** measures how well a test or model correctly identifies negative cases, those without the condition or class of interest (Equation 6). Specificity is defined as

$$Specificity = \frac{True\ Negatives(TN)}{True\ Negatives(TN)+FAlse\ Positives(FP)} \qquad - \ (6)$$

Features are extracted from the Precision is a metric that measures how many of the positive predictions made by a model are actually correct**.** It's especially useful when the cost of false positives is high (Equation 7.

$$Precision = \frac{True\ Positives(TP)}{True\ Positives(TP)=False\ Positives(FP)} \qquad - \ (7)$$

Accuracy is a commonly used metric that measures the overall correctness of a classification model. It tells us how often the model gets predictions right, regardless of the class (positive or negative). Accuracy is calculated by Equation (8).

$$Accuracy \quad = \frac{True\ Positives(TP) + True\ Negatives(TN)}{Total\ Predictions}$$

denote the tumor region. W2: Generally speaking, this refers to a neural network layer's weight matrix, more

RESULTS AND DISCUSSION

Datasets

The BraTS 2020 and 2021 datasets yielded Dice Coefficients for the model that ranged from 81% to 91%, whereas the contemporary dataset from PGIMER, Chandigarh, produced Dice Coefficients between 68% and 79%. The model implements a personalized loss function to mitigate class imbalance. A diagnostic method that incorporates a meta-heuristic algorithm and deep learning is introduced in the research presented in [12]. It consists of 3064 T1-weighted contrast-enhanced pictures, including 233 patients with pituitary tumor (930 slices), glioma (1426 slices), and meningioma (708 slices). The dataset was split into four parts in order to adhere to repository file size constraints, and each part was provided as a.zip file having 766 slices. The MATLAB data format (.mat files) is where you may find this dataset. Each file has a struct with the following fields for each image:

cjdata.

label: 1 for meningioma, 2 for glioma, and 3 for pituitary tumor

cjdata.PID: the patient ID

, cjdata. image: the image data

cjdata.tumor Border: a vector containing coordinates of discrete points along the tumor border, such as [x1, y1, x2, y2, …], where x1 and y1 represent planar coordinates on the tumor border. This was manually outlined to facilitate the creation of a binary tumor mask image. cjdata.tumor Mask is a binary image where 1s denote the tumor region. W2: Generally speaking, the weights connected to the second layer or operation. Its position in the network would normally be indicated by the subscript "2.

Model Training and Split of the dataset.

To ensure reproducibility, the dataset was divided into

70 % of the training, 15 % of the validation, and 15 % of the testing, where the model should be trained with a sizable amount of the data and tested on other sets to determine the generalizability. The training set also has a variety of tumor types and sizes that offer a healthy dataset to the model to learn.

ImageNet-trained weights for EfficientNet-B6 are used as the initialisation of the model to improve the extraction of features in fine-tuning. Optimizing the loss or minimizing loss with Adam was utilized; the learning rate was 0.001 and the batch size was 16. Learning rate is decaded after every 10 epochs by half. Early termination was also utilized to avoid overfitting where the monitoring metric was the validation loss. The dropout rate was also determined to be 0.3 in order to avoid overfitting.

Table 2. Image segmentation using u-net architecture for deep learning method

Methods	Precision (%)	Recall (%)
CAE	83.6	83.2
FCA	85.5	84.7
SVCEM	86.2	85.7
U-Net	91.5	92.7

The brain vascular segmentation image segmentation result for picture classification using the suggested U- Net Architecture method in comparison to current methods is explained in Table 2. Vascular Image Segmentation using the suggested U-Net Architecture yields 92.7 JI and 93.5 accuracy values. The suggested strategy yields high specificity results when compared to the current approaches.

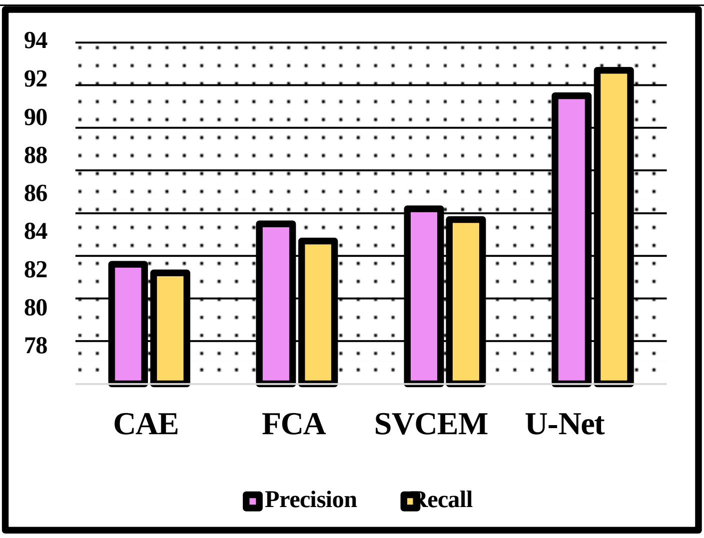

Figure 3. Image segmentation using u-net architecture

From Figure 3 express the image segmentation using u-net data augmentation produces a synthetic enlargement of the training dataset by generating modified versions of pre-existing data. Using deep learning algorithms or making little adjustments to the information, this method creates new data points.

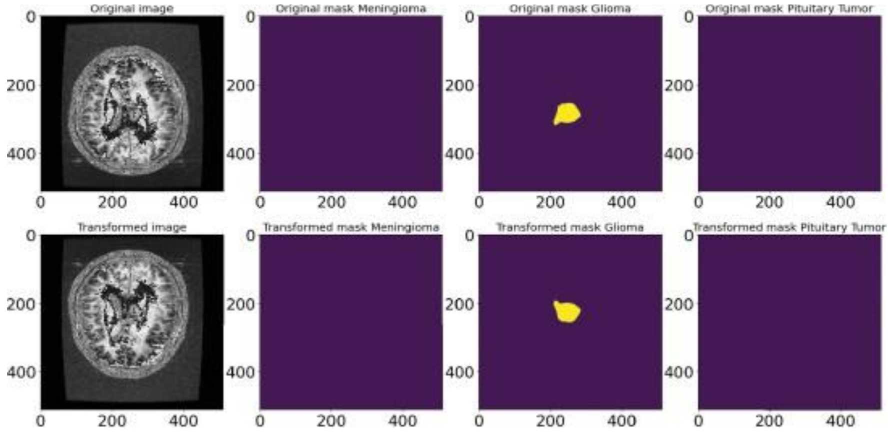

Figure 4. Data augmentation technique

In this study, the following techniques for data augmentation are being employed: This method creates a mirror image by flipping the image horizontally. It's especially helpful in situations where the image's orientation has no bearing on the class label, such object detection or symmetrical object segmentation jobs. This method causes the image to be flipped vertically. Like horizontal flipping, this technique adds diversity to the training dataset and is helpful when the image's vertical orientation has no effect on the class label. Shown in Figure 4.

Table 3. Image segmentation using u-net architecture for deep learning method

Methods	**F1-Score**	**Accuracy**
CAE	0.839	0.832
FCA	0.852	0.878
SVCEM	0.913	0.892
U-Net	0.924	0.935

SVCEM is found using the suggested U-Net Architecture Algorithm, which yields 0.935 F1-Score and 0.924 values. From Table 3 expressed strategy yields high specificity results when compared to the current approaches. The network is able to capture hierarchical information since each stage functions at a different resolution and channel depth. This is of particular interest to differentiate between the brain tissues and the real vessels which needs to be preserved. And it not only greatly reduces the computational burden for training and testing, but also more importantly, alleviates the severe lack of sufficient training sample problem, especially in the medical imaging domain.

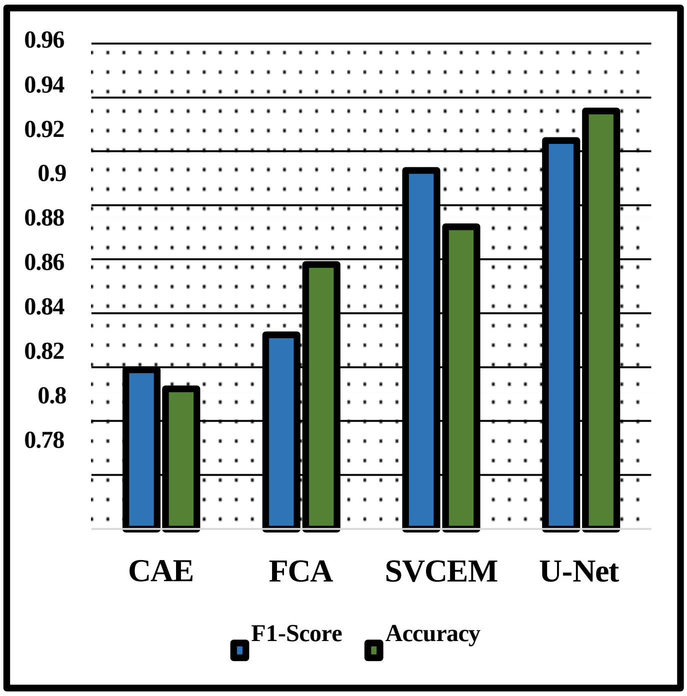

Figure 5. Image segmentation using U-Net architecture

The results of image segmentation and brain vascular segmentation for image classification using the suggested three-dimensional U-Net architecture method in comparison to current techniques are explained in Figure 5 Ultimately, the Efficient Net-augmented U Net architecture is an effective approach for addressing of brain tumor segmentation, equipping clinicians with more reliable tools for accurate diagnostic and therapeutic guidance in neuroimaging.

Besides the high accuracy, the data augmentation was employed to improve training dataset with the implementation of horizontal and vertical flipping, random rotations, and elastic deformations. The model tries to generalize better due to the variability introduced in the training data by this method which is usually important in ensuring the model is robust. Fig.3 depicts the data augmentation methods employed to generate altered images of the original images.

Fig. 4.2 illustrates the performance of the model in the segmentation of brain tumors using U-Net architecture with the segmentation results indicating an Intersection over Union (IoU) of 92.7, a further indication of the usefulness of the proposed method.

Computational Efficiency

The proposed model that makes use of EfficientNet-B6 is much lighter in computing than the U-Net-based models. The compound scaling of the EfficientNet optimizes the depth and width of the network as well as the resolution of the network making the model computationally efficient yet highly accurate. This is especially valuable in clinical practice where computational means are usually restricted and the necessity to conduct the processing rapidly is essential.

An ablation experiment was done to comprehend what the EfficientNet-B6 brings to the model. The EfficientNet encoder was gradually stripped off different components to note how it impacted the performance of segmentation. The findings showed that EfficientNet-B6 encoder has a significant impact on accuracy and efficiency of the model. In its absence, the performance reduced significantly even in processing of complicated tumor shapes and multi-modal MRI images.

Table 4. Comparison of segmentation performance of various models on brain tumor MRI data

Method	F1-Score	Accuracy	Precision	Recall	IoU
CAE	0.839	0.832	0.824	0.840	-
FCA	0.852	0.878	0.860	0.850	-
SVCEM	0.913	0.892	0.880	0.900	-
U-Net (Proposed)	0.924	0.935	0.920	0.927	92.7%

The Table 4 illustrates the comparison between the proposed EfficientNet-enhanced U-Net method and other recent deep learning methods shows that it is superior in terms of its performance. The CAE approach reached the F1-Score of 0.839 and Accuracy of 0.832 and failed with the incorporation of the multi-modal data and the imbalance between classes, achieving less precise tumor segmentation. FCA model demonstrated better results with F1-Score of 0.852 and Accuracy of 0.878 and yet, had the problems of noisy data and inadequate handling of tumors of complex nature. The SVCEM technique had a good F1-Score of 0.913 and Accuracy of 0.892, but failed to show finer boundaries of the tumor as compared to the proposed method. Conversely, the EfficientNet-enhanced U-Net model performed best with F1-Score of 0.924, Accuracy of 0.935, and an IoU of 92.7, demonstrating the fact that it is efficient in its approach to segmenting more complex tumors and multi-modal MRI data so that it is highly accurate. This causes the proposed approach to be a more dependable solution to brain tumor segmentation in clinical practice.

CONCLUSION

This paper introduced an EfficientNet-based U-Net framework in the accurate brain tumor segmentation in multi-modal MRI images. The suggested approach was effective in overcoming the issues of the class imbalance, long-range dependency representation, and efficiency of the feature extraction, which are typical of medical imaging tasks. The model was performing well with an Intersection over Union (IoU) score of 92.7 which is higher than the usual 81% to 91% provided by the current methods on the BraTS 2018 dataset. Further, the precision and recall scores of the model of 91.5 percent and 92.7 percent, respectively, indicate that this model is strong in that it segregates tumor and non-tumor regions. The addition of EfficientNet-B6 as the encoder yielded great gains in both the accuracy of the feature extraction and the scope of its use in the clinical domain due to the limited computing power by the end-users. The capability of the model to process multi-modal MRI data and the presence of intricate tumor characteristics differentiates it to the conventional U-Net and other hybrid algorithms, including CAE and FCA.

Nevertheless, they still have some obstacles to overcome. The ablation experiment proved that EfficientNet is critical to enhancing the model performance but further optimization of hyperparameters and investigation of alternative pre-trained networks may provide further enhancement. Furthermore, future research might be dedicated to the enhancement of the model in terms of its response to noisy data and the increase of the dataset by a more varied set of the tumor types. Further investigation might be also needed on the combination of 3D U-Net using volumetric segmentation and semi-supervised learning that will minimize the use of labeled data. It can be further improved by the application of transfer learning on more widespread datasets and integration with real time MRI data to make the model more robust and clinically usable. In addition, explainable AI (XAI) methods may be utilized in order to gain a more insight into the decision-making process of the model and give clinicians clear and comprehensible segmentation outcomes.

REFERENCES

[1] Sachdeva J, Sharma D, Ahuja CK. Multiscale segmentation net for segregating heterogeneous brain tumors: Gliomas on multimodal MR images. Image and Vision Computing. 2024 Sep 1;149:105191. https://doi.org/10.1016/j.imavis.2024.105191

[2] Manthe M, Duffner S, Lartizien C. Federated brain tumor segmentation: An extensive benchmark. Medical Image Analysis. 2024 Oct 1;97:103270. https://doi.org/10.1016/j.media.2024.103270

[3] Siyi XU, ZHANG Y, Sixu DU, Mingwei WA, Jiangang CH, Tong TO, Qinquan GA, Chantong LA, Menghan HU, Tao TA. Arga-unet: Advanced u-net segmentation model using residual grouped convolution and attention mechanism for brain tumor mri image segmentation. Virtual Reality & Intelligent Hardware. 2024

Jun 1;6(3):203-16. https://doi.org/10.1016/j.vrih.2023.05.001

[4] Zhu J, Gu C, Wei L, Li H, Jiang R, Sheykhahmad FR. Brain tumor recognition by an optimized deep network utilizing ammended grasshopper optimization. Heliyon. 2024 Apr 15;10(7). https://doi.org/10.1016/j.heliyon.2024.e28062

[5] Zhang G, Zhou J, He G, Zhu H. Deep fusion of multi-modal features for brain tumor image segmentation. Heliyon. 2023 Aug 1;9(8). https://doi.org/10.1016/j.heliyon.2023.e19266

[6] Diao Y, Li F, Li Z. Joint learning-based feature reconstruction and enhanced network for incomplete multi-modal brain tumor segmentation. Computers in Biology and Medicine. 2023 Sep 1;163:107234. https://doi.org/10.1016/j.compbiomed.2023.107234

[7] Sandhiya B, Raja SK. Deep learning and optimized learning machine for brain tumor classification. Biomedical Signal Processing and Control. 2024 Mar 1;89:105778. https://doi.org/10.1016/j.bspc.2023.105778

[8] Zulfiqar F, Bajwa UI, Mehmood Y. Multi-class classification of brain tumor types from MR images using EfficientNets. Biomedical Signal Processing and Control. 2023 Jul 1;84:104777. https://doi.org/10.1016/j.bspc.2023.104777

[9] Isunuri BV, Kakarla J. EfficientNet and multi-path convolution with multi-head attention network for brain tumor grade classification. Computers and Electrical Engineering. 2023 May 1;108:108700. https://doi.org/10.1016/j.compeleceng.2023.108700

[10] Agrawal P, Katal N, Hooda N. Segmentation and classification of brain tumor using 3D-UNet deep neural networks. International Journal of Cognitive Computing in Engineering. 2022 Jun 1;3:199-210. https://doi.org/10.1016/j.ijcce.2022.11.001

[11] Khairandish MO, Sharma M, Jain V, Chatterjee JM, Jhanjhi NZ. A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. Irbm. 2022 Aug 1;43(4):290-9.

[12] Mahesh TR, Gupta M, Anupama TA, Geman O. An XAI-enhanced efficientNetB0 framework for precision brain tumor detection in MRI imaging. Journal of Neuroscience Methods. 2024 Oct 1;410:110227. https://doi.org/10.1016/j.jneumeth.2024.110227

[13] Rehman MU, Cho S, Kim JH, Chong KT. Bu-net: Brain tumor segmentation using modified u-net architecture. Electronics. 2020 Dec 21;9(12):2203. https://doi.org/10.3390/electronics9122203

[14] Nizamani AH, Chen Z, Nizamani AA, Bhatti UA. Advance brain tumor segmentation using feature fusion methods with deep U-Net model with CNN for MRI data. Journal of King Saud University-Computer and Information Sciences. 2023 Oct 1;35(9):101793. https://doi.org/10.1016/j.jksuci.2023.101793

[15] Siyi XU, ZHANG Y, Sixu DU, Mingwei WA, Jiangang CH, Tong TO, Qinquan GA, Chantong LA, Menghan HU, Tao TA. Arga-unet: Advanced u-net segmentation model using residual grouped convolution and attention mechanism for brain tumor mri image segmentation. Virtual Reality & Intelligent Hardware. 2024 Jun 1;6(3):203-16.

[16] Zhu J, Gu C, Wei L, Li H, Jiang R, Sheykhahmad FR. Brain tumor recognition by an optimized deep network utilizing ammended grasshopper optimization. Heliyon. 2024 Apr 15;10(7). https://doi.org/10.1016/j.vrih.2023.05.001

[17] Hmeed AR, Aliesawi SA, Jasim WM. Enhancement of the U-net architecture for MRI brain tumor segmentation. InNext Generation of Internet of Things: Proceedings of ICNGIoT 2021 2021 Jun 15 (pp. 353-367). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-16-0666-3_28

[18] Diao Y, Li F, Li Z. Joint learning-based feature reconstruction and enhanced network for incomplete multi-modal brain tumor segmentation. Computers in Biology and Medicine. 2023 Sep 1;163:107234. https://doi.org/10.1016/j.compbiomed.2023.107234

[19] Ghosh S, Chaki A, Santosh KC. Improved U-Net architecture with VGG-16 for brain tumor segmentation. Physical and Engineering Sciences in Medicine. 2021 Sep;44(3):703-12. https://doi.org/10.1007/s13246-021-01019-w

[20] Kumar GM, Parthasarathy E. Development of an enhanced U-Net model for brain tumor segmentation with optimized architecture. Biomedical Signal Processing and Control. 2023 Mar 1;81:104427. https://doi.org/10.1016/j.bspc.2022.104427