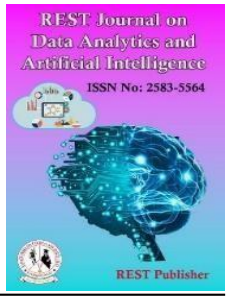
REST Journal on Data Analytics and Artificial Intelligence
Vol: 3(2), June 2024
REST Publisher; ISSN: 2583-5564
Website: http: //restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/3/2/11



Advanced U-Net Model for Precise Brain Tumor Segmentation in MRI Images

[*1]Saritha Dasari, [2]A Rama Mohana Reddy, [3]B. Eswara Reddy
[1]*JNTUA, Ananthapuramu, Andhra Pradesh, India.*
[2]*S.V. University, Tirupati, Andhra Pradesh, India.*
[3]*JNTU, Kalikiri, Andhra Pradesh, India*
**Corresponding author: sarithadasari123@gmail.com*

Abstract: *The accuracy of medical image segmentation is crucial for diagnosis and treatment planning in the modern healthcare system. Deep learning methods, like CNNs, UNETs, and Transformers, have completely changed this industry by automating labor-intensive manual segmentation procedures that were previously done by hand. However, problems like complex architectures and blurry characteristics continue, which causes issues with accuracy. Researchers are working hard to overcome these obstacles to fully realize the potential of medical image segmentation in the revolution of healthcare. Our paper presents an enhanced U-Net model specifically designed for brain tumour MRI image segmentation to improve precision. There are three primary components to our strategy. First, we prioritize feature augmentation using methods like CLAHE in the picture preprocessing phase. Second, we modify the U-Net model's architecture with an emphasis on a customized layered design in order to improve segmentation outcomes. Finally, we use a CNN model for post-processing to further optimize segmentation results using further convolutional layers. A total of 3,064 brain MRI pictures were used to test (612 images), validate (612 images), and train (1,840 images) our model. We obtained exceptional recall (93.66%), accuracy (97.79%), F-score (93.15%), and precision (92.66%). The Dice coefficient's training and validation curves showed little variation, with training reaching roughly 93% and validation 84%, suggesting good generalization ability. High accuracy was validated by visual review of the segmentation findings, albeit occasionally little mistakes like false positives were noticed.*

Keywords: *Medical Image Segmentation, UNet, CLAHE, MRI Images, Healthcare Diagnosis, Deep Learning*

1. INTRODUCTION

Brain tumors present a considerable health challenge, potentially deadly at any stage of detection (Nehra, 2021). Brain cancer can impact people of any age and gender, with more than 100 distinct types of brain tumors identified and classified as primary or metastatic (Zhang A. S.-S., 2017). Primary brain tumors can be malignant or non-cancerous and may grow inside or surrounding brain structures. Conversely, secondary (metastatic) tumors are usually malignant and arise from other body regions before spreading to the brain. Primary brain tumor survival rates are dependent on several variables, including age, location, ethnicity, kind, and molecular features of the tumor (Kiran, 2024).

MRI (Magnetic Resonance Imaging) (Abdelatty, 2024), which has become the gold standard, can now evaluate a wide range of cerebral pathologic abnormalities because of its radiation-free nature. MRI is beneficial and superior to CT in detecting acute ischemia lesions (Mahajan, 2024). However, the applicability of MRI in emergencies is limited due to the relatively long acquisition time, which makes rapid and accurate diagnosis important (Altmann, 2024). Brain MRI segmentation is an essential medical imaging process that scans the brain's MRI into distinct regions of its structure (Kumar P. R., 2024). This method is necessary for the diagnosis (G{\"o}rgec, 2024), treatment planning, and monitoring of neurological disorders, including multiple sclerosis, brain tumors, Alzheimer's disease, and other abnormalities of the brain (Desale, 2024). Similarly, MRI is crucial for quantitative brain volumetry analyses in prenatal diagnosis and examining early human brain development (Ciceri, 2024).

One of the critical concerns in clinical research methods is segmenting MRI images of the human brain. In many image analysis applications, segmenting medical images is essential (Yellu, 2024). In medical image analysis, automated segmentation techniques are more accurate than manual segmentation. Using clinically collected MRI data, automated brain tissue segmentation is one of the critical phases in the quantitative analysis of the brain. It facilitates accurate quantitative assessment of the brain, assisting in identifying, diagnosing, and categorizing diseases. Therefore, the effectiveness of the segmentation strategy is essential for disease detection and treatment planning (Kumar, 2024). The segmentation process aims to separate and specify various anatomical structures, including white matter (WM), gray matter (GM), cerebrospinal fluid (CSF), and pathological regions. Automated segmentation of perinatal brain MRI remains difficult due to considerable changes in the brain's global shape and substantial variations in image intensity, reflecting the quick tissue maturation occurring around birth.

Accurate and automated segmentation of the brain is critical for different brain analysis tasks, especially brain tumors, including high-resolution reconstruction and cortical surface analysis. Although Artificial Intelligence (AI) algorithms for medical image segmentation can perform super-human accuracy on average, many radiologists remain skeptical (Fidon, 2024). This skepticism partly arises from the fact that AI algorithms might malfunction and generate errors that contradict expert knowledge about the segmentation task, mainly when applied across various imaging protocols and anatomical pathologies (Cai, 2024). Traditional diagnostic methods are time-consuming and prone to errors (Solanki, 2023). The need for robust medical image segmentation algorithms has increased due to a demand for a more expert workforce in this specialized field. Traditional image segmentation methods, such as thresholding (Sharif, 2024), edge-based (Al Garea, 2024), and region-based methods (Reddy, 2024), face challenges due to restrictions in medical image acquisition, pathology types, and biological variations (Soppari, 2024). Brain MRI segmentation is exceptionally challenging due to constraints in image procurement, the nature of brain pathology, and biological variations. Another challenge in MRI segmentation is addressing the opacity in pixel values, which intuitionistic-based clustering algorithms aim to solve (Arora, 2024).

Our study employs U-Net for Brain MRI segmentation due to its many advantages over traditional deep learning (DL), machine learning (ML), and clustering methods. U-Net addresses the complex structures of medical images using its encoder-decoder architecture, which helps precisely locate and segment features. Unlike conventional ML and clustering techniques that depend on handcrafted features and vast manual refinement, U-Net automatically learns appropriate features through its convolutional layers. The skip connections in U-Net are also valuable for the preservation of spatial information, generating segmentation findings that are more accurate, particularly in complex and varied anatomical structures. This makes U-Net specifically effective for brain MRI and fetal brain segmentation, overwhelming challenges connected with motion artifacts and high anatomical variability.

2. REVIEW OF LITERATURE

Shahzad et al. (Shahzad, 2022) introduced several progress in medical image segmentation employing neural networks by proposing a modified U-Net architecture for deployment on the Intel/Movidius Neural Compute Stick 2 (NCS-2). The motivation behind U-Net is its demonstrated efficacy in medical image segmentation tasks, especially when handling limited dataset sizes. Their modified U-Net variant significantly reduced the parameter count from 30 million in the original U-Net to 0.49 million, desiring to optimize resource utilization without compromising performance. Promising results were obtained from experimental evaluations conducted on three different medical imaging datasets: Ziehl-Neelsen sputum smear microscopy (ZNSDB), heart MRI, and brain MRI (BraTs). For the BraTs, heart MRI, and ZNSDB datasets, the proposed approach obtained maximum dice scores of 0.96, 0.94, and 0.74, demonstrating strong performance and facilitating compelling inference on the NCS-2 platform.

Akter et al. (Akter, 2024) presented a novel deep Convolutional Neural Network (CNN)-based architecture developed for automatically classifying brain images into four different classes alongside a U-Net-based segmentation model by employing six benchmarked datasets to conduct extensive evaluations to compare the influence of segmentation on tumor classification in brain MRI images. Two classification approaches were examined based on accuracy, recall, precision, and AUC parameters, and they surpassed pre-trained alternatives with an exceptional performance standard across all datasets. Specifically, their classification model reached the highest accuracy of 98.7% in a combined dataset and 98.8% when incorporated with segmentation. Notably, the

highest classification accuracy achieved 97.7% across the four individual datasets, underscoring the efficacy of their approach in increasing brain tumor classification and segmentation employing DL techniques.

Shiny et al. (Shiny, 2024) introduced an optimization-driven approach for brain tumor classification from MRI scans. The methodology involved preprocessing pre-operative and post-operative MRI images employing filtering techniques and Region of Interest (RoI) extraction. Tumor segments were created from the preprocessed data by segmenting it using a modified U-Net model. Subsequently, histogram features were extracted, and tumor classification was performed employing a U-Net model trained with the suggested Poor Bird Swarm Optimization algorithm (PRBSA), a hybrid of the Poor and prosperous optimization (PRO) algorithm and Bird Swarm Algorithm (BSA). Finally, pixel change detection employed speeded-up robust features (SURF) on the classified output. The PRBSA-based U-Net model showed excellent performance, reaching the highest accuracy of 94%, sensitivity of 93.7%, and specificity of 94% in tumor classification tasks, showcasing its robustness in optimizing MRI-based brain tumor diagnosis.

Rutoh et al. (Rutoh, 2024) introduced a novel 3D Guided Attention-based deep Inception Residual U-Net (GAIR-U-Net) method developed to address challenges in tumor segmentation from multimodal MRI scans. The GAIR-U-Net incorporated attention mechanisms, an inception module, and residual blocks with dilated convolution to improve feature representation and spatial context understanding. In developing the U-Net architecture, the model used inception and residual connections to capture complex patterns and hierarchical features while extending its width in three-dimensional space without significantly improving computational complexity. Although dilated convolutions allowed learning local and global information, they also enhanced segmentation accuracy and adaptability by prioritizing important regions and suppressing unnecessary features. Experimental evaluations on the BraTS 2020 dataset, including T1-weighted, T1-ce, T2-weighted, and FLAIR sequences, presented promising performance. The GAIR-U-Net acquired dice scores of 0.8796, 0.8634, and 0.8441 for whole tumor (WT), tumor core (TC), and improving tumor (ET), respectively, on the BraTS 2020 validation dataset.

Chen et al. (Chen, 2024) introduced the Adaptive Cascaded Transformer U-Net (ACTransU-Net), a unique architecture for MRI brain tumor segmentation. ACTransU-Net combined Transformer and dynamic convolution within a cascaded U-Net framework to effectively capture global information and local details of brain tumors. The architecture incorporated two stages: initially cascading two 3D U-Nets for coarse-to-fine segmentation. Later, omni-dimensional dynamic convolution modules were integrated into the second-stage shallow encoder and decoder to improve local detail representation by dynamically modifying convolution kernel parameters. Additionally, 3D Swin-Transformer modules are presented into the second-stage deep encoder and decoder to capture long-range dependencies in the images, enhancing the global representation of brain tumors. Experimental results on the BraTS 2020 and BraTS 2021 datasets illustrated the effectiveness of ACTransU-Net, performing average Dice Similarity Coefficient (DSC) scores of 84.96% and 91.37% and 95th percentile Hausdorff Distance (HD95) values of 10.81 mm and 7.31 mm respectively.

3. PROPOSED METHODOLOGY

This paper presents a unique method for segmenting brain MRI data using a customized U-Net model. This model optimizes the arrangement of layers through comprehensive experimentation to achieve excellent segmentation results. The following sections provide an exhaustive explanation of our innovative approach. Figure 1 shows the overall methodology of our research work.

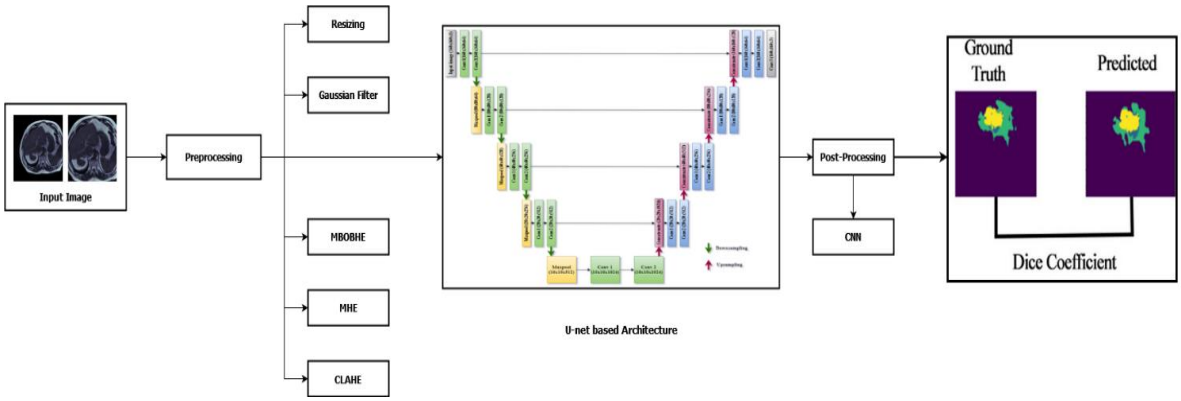
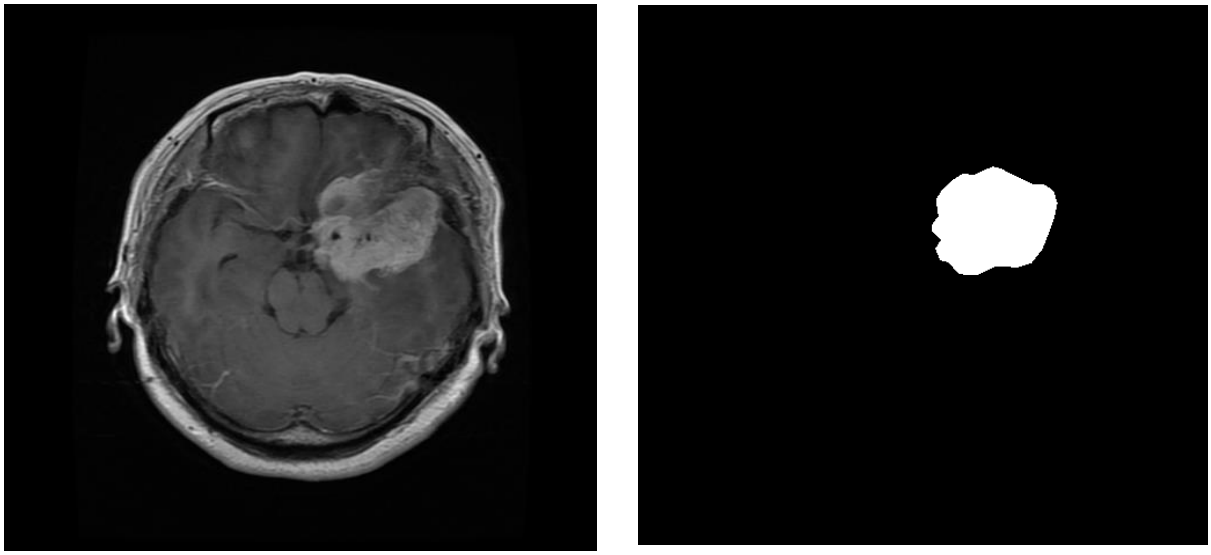


FIGURE 1. The architecture diagram for the proposed model

3.1 Dataset Description

The Kaggle dataset ("Brain MRI Segmentation") contains 3,064 brain MRI images, and each is carefully matched to a mask that indicates the exact location of any currently present tumors. This vast dataset is priceless for medical imaging researchers and practitioners because it offers a solid basis for creating, refining, and testing sophisticated ML models for tumor identification, segmentation, and analysis. Because of its thorough annotations, this dataset has a great deal of promise to enhance diagnostic accuracy and expand the capabilities of automated medical imaging systems. The sample of the dataset is shown in Figure 2.



(a) Original Image (b) Corresponding mask indicating tumor location

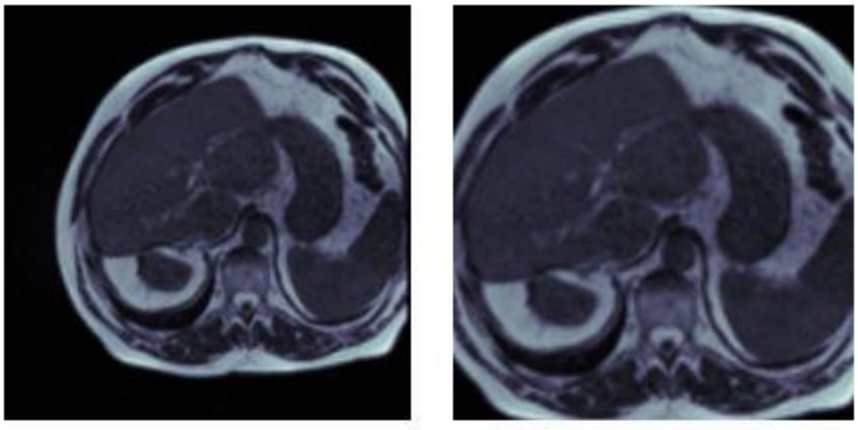
FIGURE 2. Sample of the dataset

Figure (a) illustrates the original brain MRI image, while Figure (b) displays the corresponding mask that delineates the tumor location within the same image.

Data Preprocessing:

Pre-processing steps for the dataset are resizing, filtering, normalization, and histogram equalization; a detailed description of each step is provided below. Pre-processing is done to improve the data and make it more compatible for subsequent stages.

Resizing: The method of resizing involves changing each image in the collection to a specific size. We must adjust the dataset because the neural network receives images of the same size. After executing the resizing step, the image measures 160 × 160 × 1. An input image is shown in Figure 3 (a), and its scaled image is shown in Figure 3 (b).



(a) Input Image (b) Scaled Image

FIGURE 3. Dataset Resizing (First column: Input Image, Second column: Resized Image)

Gaussian Filter: To reduce noise (high-frequency components) and blur specific areas of an image, a low-pass filter known as a Gaussian filter is employed. The weights within the kernel, which are used to determine the weighted average of the closest points (pixels) in an image, have the form of a Gaussian distribution, as the function's name suggests. A randomly selected input image is shown in Figure 4 (a), and the filtered image is shown in Figure 4 (b).

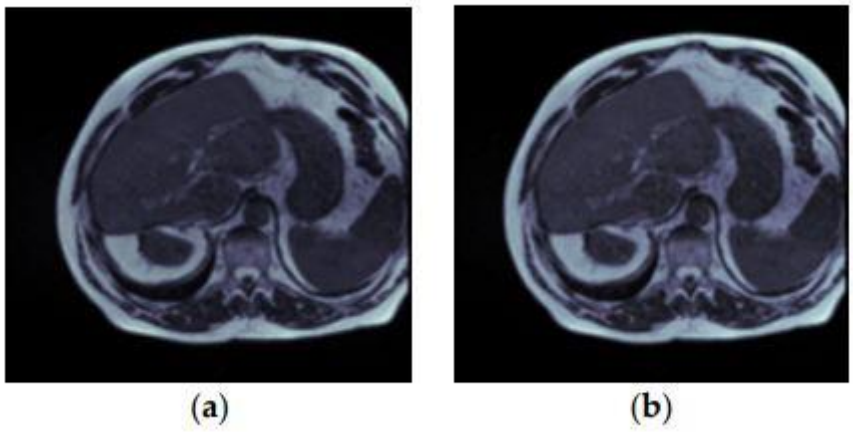


FIGURE 4. The filtered image is shown in Figures, (a) input image and (b) after filtering the input image.

Furthermore, before additional analysis or processing, preprocessing with CLAHE, MHE, and MBOBHE includes utilizing these image enhancement techniques to increase an input image's contrast and visibility. Every technique has advantages and qualities of its own.

Multi-Purpose Beta Optimized Bi-HE (MBOBHE): Though famous, the conventional Histogram Equalization (HE) method has trouble producing an enhancement that is evenly distributed. By taking into account three crucial characteristics—brightness preservation, detail preservation, and contrast enhancement—MBOBHE seeks to address this problem. Two sub-histograms that are produced after segmenting the original histogram using an optimal separating point are subjected to independent histogram optimization via MBOBHE. A weighted-sum aggregated objective function (AOF), which considers the three performance criteria, directs this optimization process. According to (Hum, 2014), MBOBHEO provides a more complete picture-enhancing performance than current bi-HE techniques. Its efficacy is confirmed by both quantitative and qualitative data, which show that MBOBH delivers a comprehensive perspective and successfully strikes a balance between contrast, brightness, and feature preservation in the enhanced images. This innovative method constitutes a noteworthy development in the field of image contrast enhancement.

Multipeak Histogram Equalization (MHE): Using this image enhancement technique, photographs having many prominent intensity peaks in their histogram can have better contrast and clarity (Shi, 2004). With pixel coordinates *(x, y)* and intensity values *I (x, y)* ranging from *0* to *L-1*, let's say we have an input image *I*. *L* is the number of intensity levels (usually 256 for 8-bit images).

Determine the input image's histogram:

$$X(i),\ i = 0, 1\ldots..L\text{-}1$$

The number of pixels in the image with an intensity value of *i* is denoted by *X(i).*

Determine the histogram's peaks:

Prominent intensity levels in the image are represented by peaks. To ascertain the locations of peaks, we can employ a variety of techniques, such as locating local maxima.

Divide the histogram into areas, often known as peaks:

Divide the histogram into various regions or sub-histograms based on the peaks that have been detected. Every sub-histogram depicts a local peak and the intensity levels around it.

For every sub-histogram, carry out histogram equalization:

Use the histogram equalization procedure for each sub-histogram to adjust the intensity values and enhance contrast.

The following represents the classic histogram equalization function:

$$A(x)=\text{round}\left((L-1)*\sum (H(y)/N) \text{ for } j = 0 \text{ to } x\right)$$

N is the total number of pixels in the sub-histogram (sum of *H(y)* for each *j* in the sub-histogram), and *A(x)* is the new intensity value for the input intensity *x*.

The final improved image is formed by combining the equalized sub-histograms obtained from the individual sub-histogram equalization based on the segmentation.

CLAHE stands for contrast-limited adaptive histogram equalization: CLAHE, a variation of Adaptive Histogram Equalization (AHE), includes a height parameter to control local contrast and limit noise amplification (Reza, 2004). The steps involved are:

I. The image is divided into $M \times N$ non-overlapping sub-regions, with the size dependent on the desired local enhancement strength.

II. A grayscale histogram *H(i*) is calculated for each sub-region.

CNN image post-processing: The suggested CNN-based architecture for MRI segmentation in the second phase of the model makes use of the segmented images from the first phase of the fusion model. The segmented image from the earlier stage is provided as input to increase the segmentation accuracy of the suggested CNN model. An image without segmentation will display all background elements, such as borders and textures. This results in the removal of unwanted features from low-priority regions. The convolution and maxpooling layers of the suggested CNN model, each of which employs a different CB, are laid out as follows. For this activity, the regular stride length is used and no cushioning is used.

The first CB consists of one MP layer and one convolution layer. The MP layer is 2-by-2, and the first convolution layer has 32 3-by-3 filters. The second CB consists of one MP layer and two convolution layers. While the MP layer has 32 filters in a (2 * 2) configuration, the two convolution layers each contain 16 filters in a (3 * 3) configuration. The third CB consists of one MP layer and two convolution layers once more. The filters in the MP layer are of size (2 * 2). whereas there are a total of 16 filters in the third convolution layer, each with a size of (3 * 3). Following the third CB, the flattening layer is applied. It "flattens" the features by condensing the feature space into a single feature vector.

Segmentation Using Proposed U-Net Model: DL has significantly transformed in different fields concerning analyzing large-scale images, audio, text, video, and tabular data. One of the immediate challenges that initially hampered the success of convolutional neural networks (CNNs) in medical image segmentation was the necessity for adequate medical images for training DL models. The U-Net architecture, developed for segmenting medical images in smaller datasets, was presented to address this issue. The segmented region of interest is highlighted in a pixel-by-pixel annotated image that U-Net generates. U-Net maintains "what" (the content) and "where" (the location) information, in contrast to conventional CNNs used for image classification, which frequently lose spatial information essential for segmentation tasks. U-Net emerges from how closely it resembles the letter *U* because it only has convolutional layers and no dense layers. It can process images of any size because it is an end-to-end fully convolutional network (FCN).

In this study, a U-Net model has been developed from scratch for the automatic segmentation of the stomach, large bowel, and small bowel in the gastrointestinal (GI) tract. The optimal number and arrangement of layers were selected through comprehensive experimentation to perform the best segmentation results. The presented U-Net model consists of a combination of convolution and max-pooling layers. As illustrated in Figure 5, the model has a U-shaped structure with an encoder on the left and a decoder on the right, following the standard U-Net layout. Interestingly, this model only includes convolution, max-pooling, and transpose convolution layers but no dense layers are included. This version of U-Net is intended for input images with a resolution of 160 by 160 pixels, in contrast to the standard version. This U-Net model's specific layers have been adjusted to fit the dimensions of the input image.

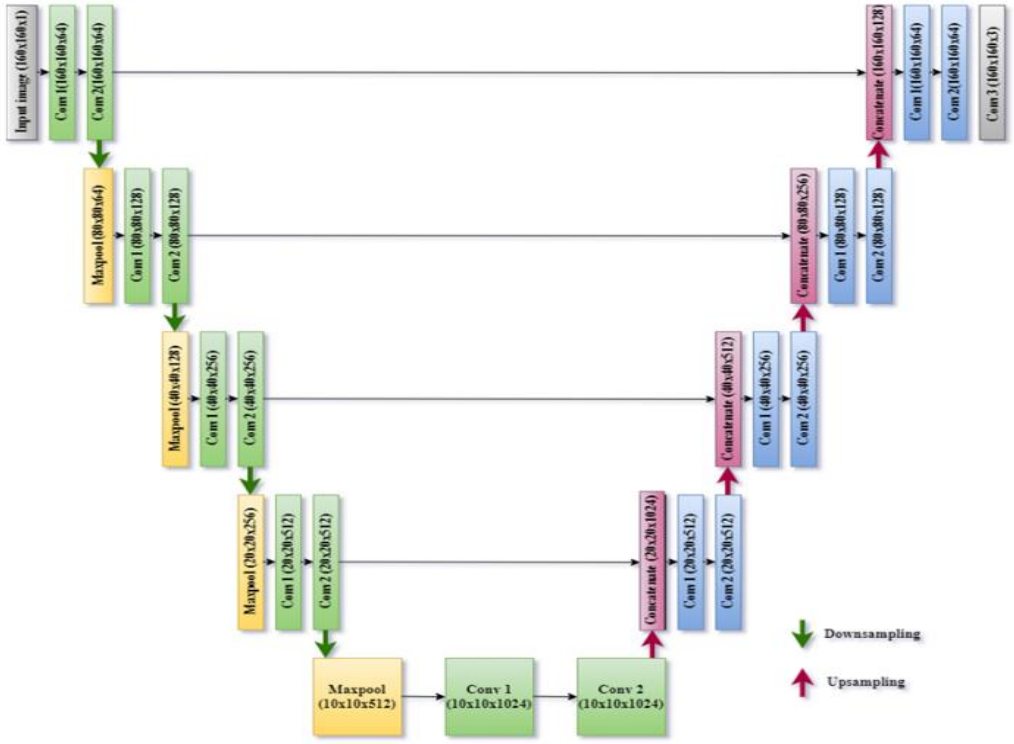


FIGURE 5. Graphical Representation of U-Net Architecture

The architecture is accessed in two ways. The context of the image is extracted utilizing the first path, also referred to as the contraction path or encoder. The encoder is a conventional stack of max pooling and convolutional layers. The second path is the decoder, which is the symmetric extending path used to obtain exact localization by transposed convolutions This is where the encoder does downsampling, and the decoder does upsampling. Precise segmentation tasks require the architecture to collect high-resolution features and the context of the input image.

A. Encoder (Contracting Path)

The encoder path comprises repeated application of two 3x3 convolutions, individually followed by a ReLU activation and a 2×2 max pooling operation with stride 2 for downsampling. At every downsampling step, the number of feature channels is doubled.

Let y be the input image, and l be the layer index in the encoder:

$m_l = \text{ReLU}(\text{Conv2D}(m_{l-1}, f = 3, t = 1, q = \text{'same'}))$

$m_l = \text{ReLU}(\text{Conv2D}(m_l, f = 3, t = 1, q = \text{'same'}))$

$m_{l+1} = \text{MaxPool}(m_l, f = 2, t = 2)$

Here,

- Conv2D (m_l, f, t, q) denotes a 2D convolutional layer involved in feature map m_l with kernel size f, stride t, and padding q.
- ReLU(y) describes the ReLU activation function applied to y.
- MaxPool (y, f, t) represents the max pooling operation with kernel size f and stride t.

B. Bottleneck

The bottleneck is the deepest portion of the network, with the greatest number of feature channels but the smallest spatial dimensions.

It includes two convolutional layers with ReLU activation:

$b = \text{ReLU}(\text{Conv2D}(m_L, f = 3, t = 1, q = \text{'same'}))$

$b = \text{ReLU}(\text{Conv2D}(b, f = 3, t = 1, q = \text{'same'}))$

C. Decoder (Expansive Path)

After upsampling the feature map, the decoder path consists of two 3×3 convolutions with ReLU activation. After every upsampling step, the corresponding feature map from the encoder path is concatenated (skip connections).

Let, v_l be the upsampled feature map, d_l be the concatenated feature map, and m be the decoder's layer index:

$v_l = \text{UpConv2D}(b, f = 2, t = 2)$

$d_l = \text{concat}(v_l, m_{L-l-1})$

$d_l = \text{ReLU}(\text{Conv2D}(d_l, f = 3, t = 1, q = \text{'same'}))$

$d_l = \text{ReLU}(\text{Conv2D}(d_l, f = 3, t = 1, q = \text{'same'}))$

Here:

- UpConv2D (y, f, t) denotes a 2D transposed convolution used to y with kernel size f and stride t.
- Concat (a,b) defines the concatenation of feature maps a and b along the channel dimension.

D. Output Layer

The final output layer involves a 1×1 convolution to map each 64-component feature vector to the expected number of classes (generally 1 for binary segmentation), followed by a sigmoid activation to make the final segmentation map.

$z = \text{sigmoid}(\text{Conv2D}(d_0, f = 1, t = 1, q = \text{'same'}))$

Here,

- Conv2D (y, f, t, q) describes a 2D convolutional layer with kernel size f, stride t, and padding q.
- Sigmoid (y) denotes the sigmoid activation function applied to y.

Up sampling: Boosting2D by repeating the row values twice, the Upsampling2D layer up samples the layer output dimension. This is where the contracting path and upsampling2D concatenate to form the expanding path. Similar to how maxpooling is present in the encoder section, the up sampling layer follows two conv2D layers in the decoder part.

Skip Connection (Residual connection): Skip connection Following each of the two Conv2D layers in the encoder comes the Skip Connection (Residual connection), also known as the identity mapping, which connects to the matching same-shaped dimension layer in the decoder section. The copy and crop in the architecture above handle the residual connection task. The task of combining those two layers is completed by the Concatenate layer. Even in the worst situation, this skip connection does not negatively impact the model; on the contrary, it positively impacts the model's output.

The output segmentation map with filter 2 marks the conclusion of the decoder section. Once more, this layer passes through filters 1 and a Conv2D with Relu activation. The output categorizes each pixel according to whether a tumor is present or not.

Hyper parameter Tuning: Twenty epochs were used to train the models with a batch size of 32. While the epochs parameter specifies the number of runs over the entire training data, the batch size hyper parameter specifies the number of samples to proceed before changing the model's internal parameters. The learning rate, which regulates the model's pace of learning, is the essential hyper-parameter. It must not be unnecessarily high or low. The network may overshoot the low-loss regions if the learning rate is set too high, or it may take an

excessive amount of time to reach the minimal loss if it is set too low. In this work, the learning rate is set at 0.0001. The Adam (Zhang, 2018) optimization method was applied for model compilation. In addition, the ReLU (Agarap, 2018) activation function has been used to activate all convolutional layers. Table 1 shows the parameters of the proposed model.

TABLE 1. Parameters of the proposed model

Parameter	Value/Explanation
Model	'unet((H, W, 3))' - A U-Net model with input shape (Height, Width, 3 channels)
Loss Function	'dice_loss'
Optimizer	'Adam(lr)'
Metrics	'dice_coef',' accuracy'
Callbacks	Various callbacks used during training
ModelCheckpoint	Saves the best model only ('save_best_only=True'); verbose output enabled ('verbose=1')
ReduceLROnPlateau	Reduces learning rate by a factor of 0.1 if no improvement in validation loss for 5 epochs ('patience=5'); minimum learning rate set to '1e-7'; verbose output enabled ('verbose=1')
CSVLogger	Logs training data to CSV file specified by 'csv_path'
EarlyStopping	Stops training if no improvement in validation loss for 20 epochs ('patience=20'); does not restore best weights ('restore_best_weights=False')
Training Data	'train_dataset'
Epochs	'num_epochs'
Validation Data	'valid_dataset'
Verbose	'0' (silent mode for training output)

4. EXPERIMENTAL ANALYSIS & RESULTS

Our brain tumor detection and segmentation experimental setup makes use of several potent tools and libraries. To implement and assess machine learning models, including feature extraction, data preprocessing, and model performance measures, we utilize Scikit-learn, also known as sklearn. To visualize the MRI images, masks, and segmentation results, Matplotlib is used, which makes the findings easy to understand and display. In addition to enabling smooth code execution and sharing, Google Colab offers a scalable and collaborative computing environment that makes use of high-performance computing resources essential for training intricate models on sizable datasets. This integrated design increases the reliability and validity of our experimental results by guaranteeing a productive, collaborative, and repeatable research process. To achieve robust model evaluation and generalization, we have divided the dataset into training, validation, and test sets in our experimental setup. The dataset is divided into three sets: 612 images and masks from the validation set, 612 images and masks from the test set, and 1,840 images and masks from the training set. Our machine learning models can be trained, tuned, and tested more effectively thanks to this stratified split, which also ensures accurate performance evaluation and optimization. A set of evaluation metrics were used to evaluate the model. To assess the performance of our tumor detection and segmentation models, we use the following evaluation metrics:

- **Precision:** Precision measures the proportion of true positive predictions among all positive predictions. It indicates how many of the predicted positive cases are positive.

$$\text{Precision} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)}$$

- **Recall (Sensitivity):** Recall measures the proportion of true positive predictions among all actual positive cases. It indicates how well the model identifies positive cases.

$$\text{Recall} = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negative\ (FN)}$$

- **F-score:** The F1-Score is the harmonic mean of precision and recall, providing a single metric that balances both aspects. It is particularly useful when the class distribution is imbalanced.

$$\text{F1-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

- **Accuracy:** Accuracy measures the proportion of true positive and true negative predictions among all predictions. It indicates the overall correctness of the model.

$$\text{Accuracy} = \frac{True\ Positives\ (TP) + Ture\ Negative\ (TN)}{Total\ Predictions}$$

Analysis of Dice Coefficient: A popular statistic for assessing how comparable two samples are—in this case, the anticipated and ground truth masks of brain tumors in MRI images—is the dice coefficient. It is computed as the ratio of the total areas of both masks to twice the intersection of the ground truth and anticipated masks. In essence, more overlap and agreement between the estimated and actual tumor locations are indicated by a higher Dice coefficient.

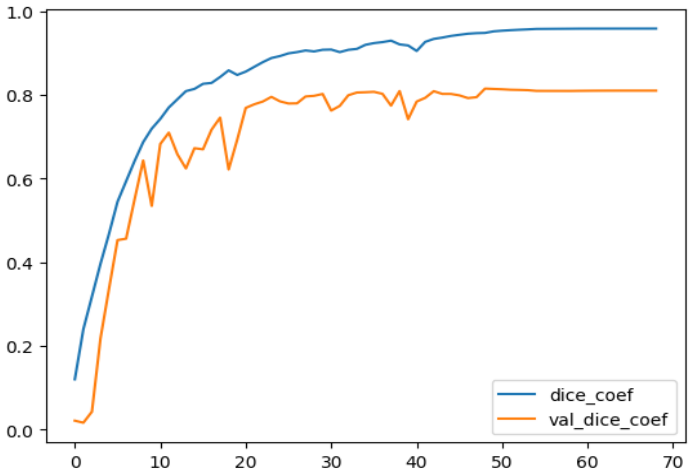


FIGURE 6. Dice Coefficient of the proposed model

The curve is displayed in Figure 6. In the case of our model, we find that the validation Dice coefficient is approximately 84% and the training Dice coefficient is approaching 93%, indicating that the model performs reasonably well in generalizing to unknown data but with a minor decline in performance from the training set. Since both curves have a similar tendency, the consistency and small divergence between the training and validation curves suggest that the model is not appreciably overfitting. The model successfully learns the patterns and features required to precisely forecast tumor locations from the training data, as shown by the training curve's high Dice coefficient. However, the marginally lower validation Dice value raises the possibility that the distribution of data between the training and validation sets differs in some way. This may be the result of variables the model encounters during validation, such as variances in image quality, anatomical variability, or other features found in real-world data.

Overall, the validation set yielded a Dice coefficient of about 84%, which is a good result and shows that the model can correctly identify and segment tumors in unseen MRI images. The model's capacity for generalization appears to be strong based on the consistency of its performance across training and validation sets; nonetheless, changes and additional research should be directed toward minimizing any disparity in these performance metrics.

Generalization analysis of the proposed model: In this work, we evaluate a crucial component of model generalization for MRI image-based brain tumor identification and segmentation. The term "generalization" describes a model's capacity to extrapolate taught patterns and characteristics into real-world applications by applying its newly acquired knowledge to previously viewed data that falls outside of its training set. The training and validation accuracy also training and validation loss in Figures 7 and 8.

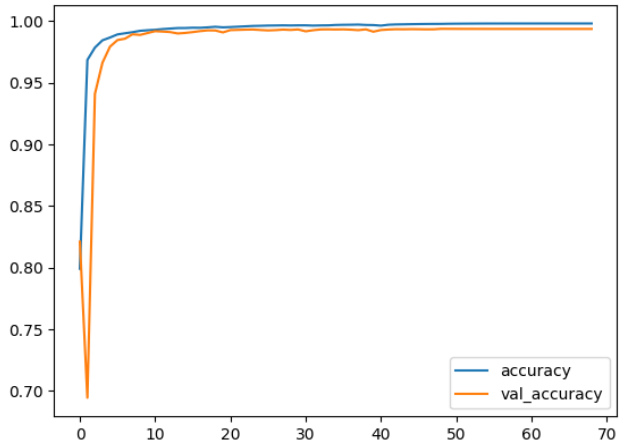


FIGURE 7. Training and validation accuracy

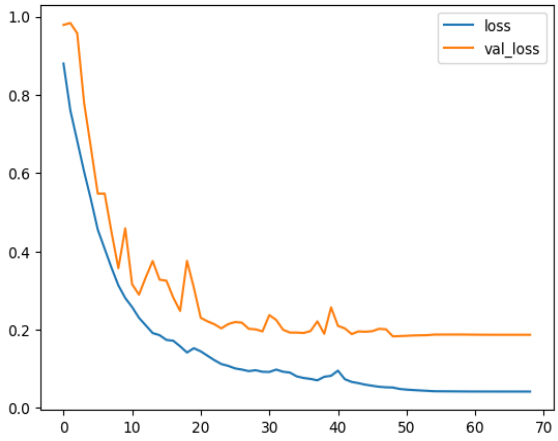


FIGURE 8. Training and validation loss

The figure 8 shows that, when it comes to the task of identifying and segmenting brain tumors from MRI images, our model shows strong signs of generalization. The term "generalization" describes the model's ability to extrapolate learnt patterns and characteristics into real-world applications by applying its learned knowledge to new, unseen data outside of the training set.

First off, the model's ability to accurately categorize tumor and non-tumor regions within the training set of MRI images is demonstrated by its training accuracy of roughly 99.65%. The strong learning process, which allows the model to successfully capture complex features and patterns unique to the dataset, is shown by this high accuracy. Simultaneously, the validation accuracy of approximately 99.32% suggests that the model continues to function well when used with unobserved data. This validation accuracy is important because it shows that the model can generalize well, applying its gained information to new instances outside of the training samples, even though it is somewhat lower than the training accuracy. Both training and validation losses were continuously minimal throughout the training procedure. The model's inaccuracy during training was measured by the training loss, which dropped over time from an initial value of 0.880 in the first epoch to 0.092 in the last epoch. In a similar vein, over that time the validation loss dropped from 0.979 to 0.196. These low loss values show that, for both the training and validation datasets, the model successfully reduces mistakes and differences between expected and actual outputs. The model's capacity to generalize well is further supported by this consistency in loss reduction, which shows that the model learns to produce precise predictions while retaining resilience against overfitting or underfitting. The model's consistent and dependable performance across several datasets is supported by the small difference between training and validation accuracies as well as the convergence of training and validation losses at the end of the training process. The fact that the model has successfully acquired the important characteristics of brain tumor images without retaining extraneous information or background noise from the training set is indicated by this convergence. In summary, our proposed U-Net model shows good generalization capabilities in the crucial task of brain tumor detection and segmentation, based on the thorough examination of training and validation metrics from images. These results are crucial for increasing the precision and reliability of tumor detection in clinical settings, where it is necessary to improve patient outcomes and treatment plans.

Performance of the proposed model: Lastly, we have used test accuracy, F-score, precision, and recall to assess our model's performance. The performance of the proposed U-Net model is shown in Table 2. Precision, a metric used to gauge how accurate the positive forecasts were, came in at 92.66%. With a high precision value, the model minimizes false positives—regions that are wrongly labeled as tumors—and is highly effective at correctly identifying actual positive cases of brain tumors. The recall measure, which assesses how well the model captures all pertinent cases of brain tumors, came in at 93.65%. This illustrates how well the model detects most true positive cases, meaning that very few tumors are overlooked. In medical diagnostics, where missing a tumor could have serious repercussions for patient outcomes, a high recall rate is essential. The harmonic mean of recall and precision, or the F-score, is 93.15%. This well-balanced indicator captures the total efficacy of the model by taking into account both its performance in identifying true positives and its ability to prevent false positives. The model performs well in both areas, maintaining a good trade-off between precision and recall, as indicated by the high F-score. Furthermore, the model's overall accuracy on the test set is a remarkable 97.79%. The percentage of accurate outcomes (true positives and true negatives) among all instances investigated is known as accuracy. This

high accuracy rate demonstrates the model's dependable performance in a variety of scenarios by demonstrating its strong ability to accurately categorize both tumor and non-tumor regions.

TABLE 2. Performance of the proposed U-Net model

Evaluation Metrics	Performance
Precision	92.66 %
Recall	93.65 %
F-score	93.15 %
Accuracy	97.79 %

Performance Comparison with Trending Method: Our presented feature-enhanced U-Net models exhibited outstanding performance in brain tumour MRI image segmentation when compared to traditional methods with (Zhang Y. a., 2024), (Huang, 2020) , and (Ishfaq, 2023) as shown in Table 3. The model designed by (Huang, 2020) employed lightweight feature extraction modules and attention mechanisms to improve performance, allowing accurate diagnosis and adequate treatment planning. (Ishfaq, 2023) employed segmentation, clustering, and multi-class SVM to extract features and classify tumours. However, the multi-class SVM method needs to be enhanced in separating complicated datasets, resulting in lower performance. Additionally, both (Huang, 2020), (Ishfaq, 2023) struggled with semantic understanding of complex data, further hindering their performance.

TABLE 3. Compare our proposed model with the existing method

Reference	Accuracy
(Zhang Y. a., 2024)	0.806
(Huang, 2020)	0.816
(Ishfaq, 2023)	0.846
Proposed U-net	0. 97

In contrast, our U-NET models prioritized feature advancement during image preprocessing by employing techniques like CLAHE to enhance the visibility of complicated structures and indistinct features in medical images. We also tailored the U-Net architecture with a personalized layered design explicitly optimized for the challenges of brain tumour segmentation. This holistic approach addresses the complexities of medical image segmentation, resulting in excellent outcomes. Our proposed U-Net method performed an accuracy of 0.97, greatly exceeding traditional methods, which achieved accuracies of 0.806 (Huang, 2020), 0.816 (Ishfaq, 2023), and 0.846 (Zhang Y. a., 2024). This combination of innovative strategies highlighted the efficacy of our models in increasing the field of healthcare through enhanced medical image segmentation.

Visual analysis of the segmentation: We have presented the performance of our model using a set of photos in the results part of our brain MRI segmentation analysis, showing both accurate and inaccurate segmentations. In Figure 9, the visulization is displayed. The real part of the mask image is clearly superimposed over the MRI scans in the first two figures. The capacity of the model to accurately identify and segment the tumor regions is seen in these figures. The model has successfully trained to separate tumor tissues from the surrounding brain structures based on how well the segmented areas in these images match the ground truth masks. Because it guarantees that the tumor zones are accurately identified without encompassing non-tumor areas, this precise segmentation is essential for accurate diagnosis and therapy planning. Unfortunately, there is a segmentation problem in the final image. The primary tumor site is properly segmented by the model, but an additional area is mistakenly identified as a tumor. This incorrect segment highlights a false positive in the model's predictions because it is located next to the correct segment area. Even though they are less common, these mistakes are important because they may result in MRI images being misinterpreted or needless therapeutic procedures. The intricacy and unpredictability of brain tumor appearances in MRI scans, the caliber of the training data, and possible overfitting to particular features within the training set are some of the reasons for this inaccuracy. In order to correct these flaws, the model must be improved by employing strategies including adding more diverse instances to the dataset, putting sophisticated regularization techniques into practice, and adjusting the model's hyperparameters to improve generalization. The focus of future research can be on error reduction strategies.

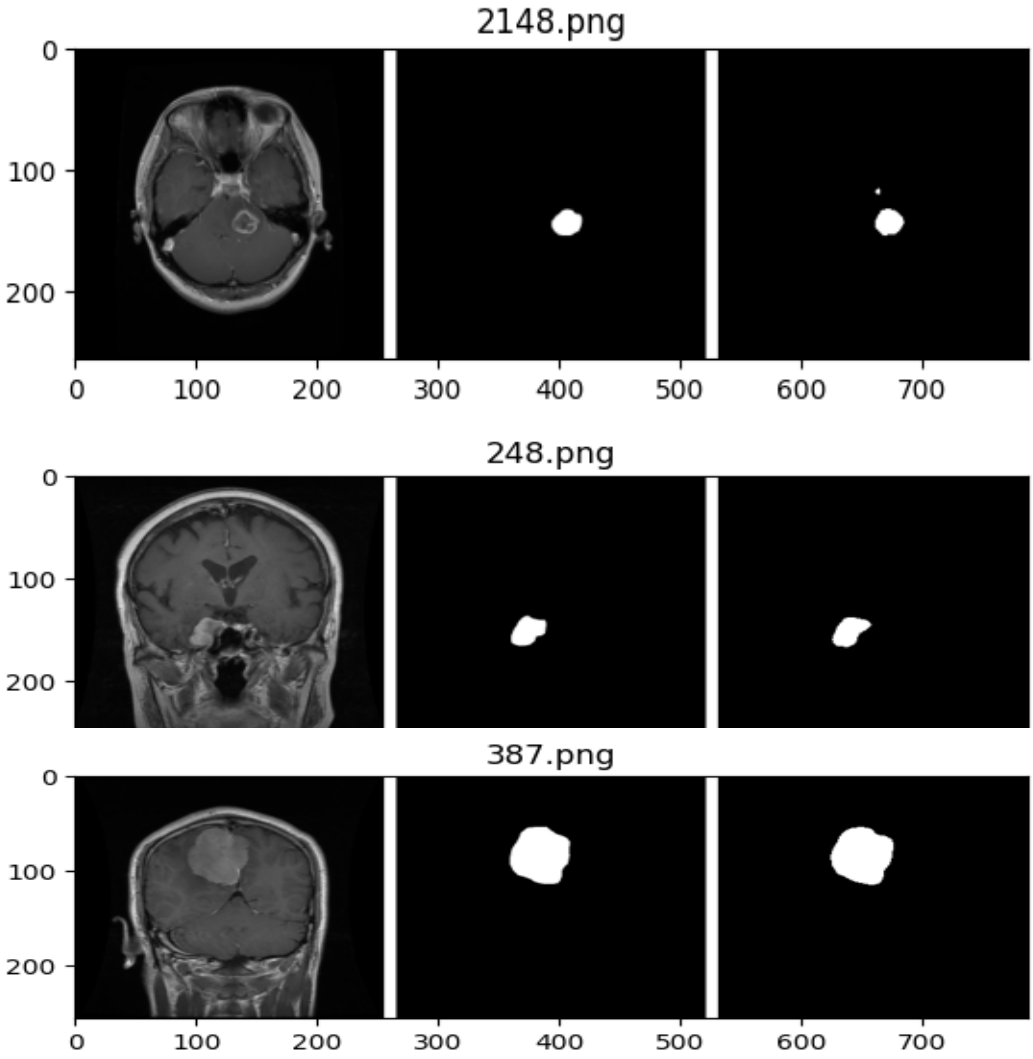


FIGURE 9. Visual analysis of segmented images

5. CONCLUSION

In this study, we present a unique approach to addressing the challenges of brain tumour MRI image segmentation in the realm of modern healthcare by using the power of DL algorithms. We introduce an enhanced U-Net model developed explicitly for brain tumour MRI image segmentation to improve precision. Our approach comprises a feature augmentation employing methods like CLAHE, MHE, and MBOBHE in the image preprocessing phase. Then, we modify the U-Net model's architecture with a customized layered design to enhance segmentation outcomes. Eventually, we utilize a CNN model for post-processing to optimize segmentation results further using additional convolutional layers. We performed remarkable recall 93.66 %, accuracy 97.79 %, F-score 93.15 %, and precision 92.66 %. The Dice coefficient's training and validation curves demonstrated slight variation, with training acquiring roughly 93% and validation 84 %, indicating good generalization ability. There are other possibilities where additional research can be done. This work laid the foundation for more precise and effective medical image segmentation techniques, which could significantly impact healthcare by helping radiologists with accurate diagnosis and treatment planning.

REFERENCES

[1]. Abdelatty, M. A. (2024). Magnetic resonance imaging of pilonidal sinus disease: interobserver agreement and practical MRI reporting tips. *European Radiology*, 115--125.
[2]. Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.
[3]. Akter, A. a. (2024). Robust clinical applicable CNN and U-Net based algorithm for MRI classification and segmentation for brain tumor. *Expert Systems with Applications*, 122347.

[4]. Al Garea, S. a. (2024). Image Segmentation Methods: Overview, Challenges, and Future Directions. In *2024 Seventh International Women in Data Science Conference at Prince Sultan University (WiDS PSU)* (pp. 56--61).
[5]. Altmann, S. a. (2024). Ultrafast brain MRI with deep learning reconstruction for suspected acute ischemic stroke. *Radiology*, e231938.
[6]. Arora, J. a. (2024). Conditional spatial biased intuitionistic clustering technique for brain MRI image segmentation. *Frontiers in Computational Neuroscience*, 1425008.
[7]. Cai, Z. a.-M. (2024). Enhancing Generalized Fetal Brain MRI Segmentation using A Cascade Network with Depth-wise Separable Convolution and Attention Mechanism. *arXiv preprint arXiv:2405.15205*.
[8]. Chen, B. a. (2024). Adaptive cascaded transformer U-Net for MRI brain tumor segmentation. *Physics in Medicine \& Biology*, 115036.
[9]. Ciceri, T. a. (2024). Fetal brain mri atlases and datasets: a review. *NeuroImage*, 120603.
[10]. Desale, P. a. (2024). Navigating Neural Landscapes: A Comprehensive Review of Magnetic Resonance Imaging (MRI) and Magnetic Resonance Spectroscopy (MRS) Applications in Epilepsy. *Cureus*.
[11]. Fidon, L. a. (2024). A Dempster-Shafer approach to trustworthy AI with application to fetal brain MRI segmentation. *IEEE transactions on pattern analysis and machine intelligence*.
[12]. G{\"o}rgec, B. a. (2024). MRI in addition to CT in patients scheduled for local therapy of colorectal liver metastases (CAMINO): an international, multicentre, prospective, diagnostic accuracy trial. *The Lancet Oncology*, 137--146.
[13]. Huang, S. a. (2020). Medical image segmentation using deep learning with feature enhancement. *IET Image Processing*, 3324-3332.
[14]. Hum, Y. C. (2014). Multiobjectives bihistogram equalization for image contrast enhancement. *Complexity*, 22-36.
[15]. Ishfaq, M. A. (2023). Brain tumor classification utilizing deep features derived from high-quality regions in MRI images. *Biomedical Signal Processing and Control*, 104988.
[16]. Kiran, L. a. (2024). An enhanced pattern detection and segmentation of brain tumors in MRI images using deep learning technique. *Frontiers in Computational Neuroscience*, 1418280.
[17]. Kumar, P. R. (2024). Automated human brain tissue segmentation from clinical MRI images for improved neurological diagnosis and treatment planning. *Intelligent Medicine*.
[18]. Kumar, P. R. (2024). Automated human brain tissue segmentation from clinical MRI images for improved neurological diagnosis and treatment planning. *Intelligent Medicine*.
[19]. Mahajan, A. a. (2024). Neuroimaging: CT Scan and MRI. In *Principles and Practice of Neurocritical Care* (pp. 189--215). Springer.
[20]. Nehra, M. a. (2021). Nanobiotechnology-assisted therapies to manage brain cancer in personalized manner. *Journal of Controlled Release*, 224--243.
[21]. Reddy, S. a. (2024). Region based image segmentation to improve accuracy of currency images compared with edge based segmentatio. In *AIP Conference Proceedings.*
[22]. Reza, A. M. (2004). Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*.
[23]. Rutoh, E. K. (2024). GAIR-U-Net: 3D guided attention inception residual u-net for brain tumor segmentation using multimodal MRI images. *Journal of King Saud University-Computer and Information Sciences*, 102086.
[24]. Shahzad, O. A. (2022). Implementation of a Modified U-Net for Medical Image Segmentation. *IEEE*, 4593--4597.
[25]. Sharif, M. a. (2024). Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. *Journal of ambient intelligence and humanized computing*, 1--20.
[26]. Shi, H. C. (2004). A simple and effective histogram equalization approach to image enhancement. *Digital Signal Processing*, 158-170.
[27]. Shiny, K. (2024). Brain tumor segmentation and classification using optimized U-Net. *The Imaging Science Journal*, 204--219.
[28]. Solanki, S. a. (2023). Brain Tumor Detection and Classification Using Intelligence Techniques: An Overview. *IEEE Access*, 12870-12886.
[29]. Soppari, K. a. (2024). A survey on brain MRI segmentation. *World Journal of Advanced Research and Reviews*, 1702--1710.
[30]. Yellu, R. R. (2024). edical Image Analysis-Challenges and Innovations: Studying challenges and innovations in medical image analysis for applications such as diagnosis, treatment planning, and image-guided surgery. *Journal of Artificial Intelligence Research and Applications*, 93--100.
[31]. Zhang, A. S.-S. (2017). Complete prevalence of malignant primary brain tumors registry data in the United States compared with other common cancers, 2010. *Neuro-oncology*, 726--735.
[32]. Zhang, Y. a. (2024). Interactive medical image annotation using improved Attention U-net with compound geodesic distance. *Expert systems with applications*, 121282.
[33]. Zhang, Z. (2018). Improved Adam Optimizer for Deep Neural Networks. In *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)* (pp. 1-2)

