**International Journal of Intelligent Communication and Computer Science**
**Vol. 3, No. 3, 2025, Pages 32-50.**

# A Deep U-Net Framework for Accurate Brain Tumour Segmentation from MRI Images

**Snigdha Singh[1] and Atul Mathur[1]**

[1]Naraina College of Engineering & Technology, Kanpur, U.P. INDIA

Email: Snigdha3884@gmail.com

**Research Paper**

Received: 23 Feb 2026, Revised: 29 March. 2026, Accepted: 17 May, 2026

## Abstract:

Brain tumour segmentation is a crucial step in medical image analysis, as it assists clinicians in diagnosis, treatment planning, and disease monitoring. Accurate delineation of tumour regions enables the assessment of tumour size, shape, location, and progression, which are essential for developing effective therapeutic strategies. However, automatic segmentation of brain tumours from Computed Tomography (CT) images remains a challenging task due to variations in tumour appearance, irregular boundaries, intensity inhomogeneity, image noise, and low contrast between healthy and abnormal tissues. In recent years, deep learning has significantly advanced the field of medical image analysis, achieving remarkable improvements in image segmentation tasks that were previously difficult to address using conventional methods. Among various deep learning architectures, encoder–decoder networks based on convolutional neural networks (CNNs) have demonstrated outstanding capability in learning complex spatial and contextual features from medical images. In particular, the U-Net architecture has emerged as one of the most widely adopted and effective models for biomedical image segmentation due to its ability to capture both local and global image information while preserving fine structural details. In this study, a U-Net-based deep learning framework is employed for brain tumour segmentation using the BraTS (Brain Tumour Segmentation) dataset. The performance of the proposed model is evaluated using standard segmentation metrics, including Dice Similarity Coefficient (DSC), precision, recall, and accuracy. Experimental results demonstrate that the U-Net architecture provides accurate tumour segmentation while effectively preserving boundary information and enhancing feature representation, making it a reliable approach for automated brain tumour analysis.

**Keywords:** Brain Tumour Segmentation, Deep Learning, U-Net Architecture, Medical Image Analysis, Computed Tomography (CT).

## 1. Introduction

Brain tumours represent a serious category of neurological diseases that can adversely influence both life expectancy and overall patient well-being when detection is delayed [1]. Identifying and separating tumor regions from medical images is an essential component of modern healthcare, supporting diagnosis, treatment design, surgical procedures, radiation therapy, and follow-up assessment of disease progression [2]. Advanced imaging techniques, particularly Magnetic Resonance Imaging (MRI) and Computed Tomography (CT), offer valuable insights into the structural and spatial properties of tumours, such as their dimensions, anatomical position, and shape. Nevertheless, the conventional practice of manually outlining tumour regions requires considerable expertise and effort, often leading to inconsistencies between specialists. These challenges have encouraged extensive research into automated and intelligent tumour segmentation approaches [3]. Automatic brain tumour segmentation remains a challenging task due to the heterogeneous nature of tumours, irregular boundaries, variations in shape and size, intensity inhomogeneity, and the presence of noise and artifacts in

medical images [4]. Conventional image processing and machine learning methods often depend on handcrafted features, which may not adequately capture the complex spatial and contextual information present in tumour regions. As a result, their performance is often limited when applied to diverse clinical datasets [5].

The emergence of deep learning techniques has significantly advanced medical image processing by facilitating automated representation learning and direct optimization from extensive imaging datasets [6]. Convolutional Neural Networks (CNNs) have proven highly effective across numerous healthcare imaging tasks, such as disease identification, lesion localization, object recognition, and image segmentation [7]. Among the various segmentation frameworks, U-Net introduced by Ronneberger et al. [8] is widely recognized as a landmark architecture in biomedical image analysis. The model employs a symmetric contracting and expanding pathway that captures semantic information at multiple scales. Additionally, shortcut connections between corresponding layers enable the retention of detailed spatial features, thereby improving the precision of tissue and tumour boundary segmentation. The encoder path of U-Net progressively extracts hierarchical feature representations through convolution and pooling operations, whereas the decoder path reconstructs the segmentation map through upsampling and feature fusion. The skip connections between corresponding encoder and decoder layers facilitate efficient information transfer and improve localization accuracy [8]. Due to its effectiveness, U-Net and its variants have been successfully applied to numerous medical imaging tasks, including brain tumour segmentation, organ delineation, lesion detection, and retinal vessel extraction [9,10].

In this study, a U-Net-based deep learning framework is proposed for automated brain tumour segmentation using the BraTS (Brain Tumour Segmentation) dataset [11]. The performance of the model is evaluated using widely adopted segmentation metrics, including Dice Similarity Coefficient (DSC), precision, recall, and accuracy. Experimental results demonstrate the capability of the proposed architecture to accurately identify tumour regions while preserving boundary information and capturing discriminative features. The obtained findings highlight the effectiveness of U-Net for robust brain tumour segmentation and its potential to assist clinicians in diagnosis and treatment planning.

## 2. Literature Survey

Recent advances in deep learning have significantly improved the accuracy and reliability of brain tumour detection, classification, and segmentation from magnetic resonance imaging (MRI). Various convolutional neural network (CNN)-based architectures and hybrid deep learning frameworks have been proposed to enhance diagnostic performance.

Paul et al. [12] investigated the application of deep learning techniques for brain tumour classification using MRI images. Their CNN-based framework demonstrated the potential of deep feature extraction in distinguishing different tumour categories and achieved an accuracy of 91.43%. Sultan et al. [13] proposed a deep neural network-based multiclass brain tumour classification system capable of identifying multiple tumour types. Their approach achieved classification accuracies of 96.03% and 98.70% for different model configurations, highlighting the effectiveness of deep neural networks in medical image analysis.

Badža and Barjaktarović [14] developed a CNN-based model for classifying brain tumours from MRI scans. The proposed architecture achieved an accuracy of 95.40%, demonstrating the capability of CNNs to automatically learn discriminative tumour features. Noreen et al. [15] introduced a deep learning framework based on feature concatenation using Inception-V3 and DenseNet-201 architectures. Their method achieved accuracies of 99.34% and 99.51%, respectively, indicating that feature fusion can substantially improve classification performance.

Khan et al. [16] combined K-means clustering with deep learning and synthetic data augmentation techniques for brain tumour segmentation and classification. Their framework achieved an accuracy of 94.06%, showing that data augmentation can enhance model generalization. Díaz-Pernas et al. [17] proposed a multiscale CNN architecture for simultaneous brain tumour classification and segmentation. The multiscale feature extraction strategy enabled the model to achieve an accuracy of 97.30%, demonstrating its effectiveness in capturing tumour characteristics at different spatial scales.

Masood et al. [18] developed a novel deep learning framework utilizing Mask R-CNN with a DenseNet-41 backbone. The model achieved segmentation and classification accuracies of 96.30% and 98.34%, respectively, proving its effectiveness for both localization and diagnosis tasks. Dipu et al. [19] evaluated CNN and YOLOv5 architectures for brain tumour detection and classification. Their CNN model achieved 95.78% accuracy, while

YOLOv5 obtained 85.95%, indicating that CNN-based approaches remained more effective for the considered dataset.

Ramesh et al. [20] proposed a framework incorporating a modified median filter for noise removal followed by deep learning-based segmentation and classification. Their method achieved 96% accuracy, demonstrating the importance of preprocessing in improving classification performance. Banerjee et al. [21] introduced a deep radiomics approach employing VGGNet, ResNet, and ConvNet architectures for brain tumor detection and classification from multi-sequence MRI data. The proposed framework achieved an accuracy of 95%, illustrating the utility of deep radiomic features for tumour characterization.

Khan et al. [22] presented an intelligent brain tumour identification model based on a hierarchical deep learning framework and CNN architecture. The proposed system achieved a precision of 92.13%, indicating reliable tumour identification capabilities. Anaya-Isaza and Mera-Jiménez [23] explored transfer learning and data augmentation techniques using ResNet50 and principal component analysis (PCA)-based augmentation. Their approach achieved an F1-score of 92.34%, demonstrating the effectiveness of transfer learning in scenarios with limited medical imaging data.

Raza et al. [24] proposed a hybrid deep learning framework integrating GoogleNet and Leaky ReLU activation functions for brain tumour classification. Their method achieved an accuracy of 99.67%, making it one of the highest-performing classification systems reported in the literature. Similarly, Zahoor et al. [25] developed a deep hybrid boosted and ensemble learning model incorporating BRAIN-RENet and ensemble classifiers. The proposed system achieved an accuracy of 99.56%, demonstrating the effectiveness of combining multiple learning strategies for robust tumour classification.

Anjum et al. [26] employed transfer learning using GoogleNet and ResNet101 architectures for brain tumour detection from MRI images. Their CNN-based transfer learning framework achieved an accuracy of 99.33%, confirming the advantages of pretrained deep networks for medical image classification tasks.

Overall, the literature indicates that CNN-based and hybrid deep learning models have achieved remarkable performance in brain tumour analysis. However, challenges remain in accurately segmenting tumour boundaries, reducing computational complexity, and maintaining robust performance across heterogeneous MRI datasets. These limitations motivate the development of advanced segmentation architectures such as MultiResUNet, which can effectively capture multiscale contextual information while preserving fine-grained tumour details.

## 3. Variants of U-Net

Among the numerous deep learning models proposed for image segmentation, U-Net has emerged as a highly effective and widely utilized architecture in the field of medical image analysis. Its popularity stems from a distinctive encoder–decoder design that enables the network to capture both high-level semantic features and fine-grained structural details from medical images. The inclusion of skip connections between corresponding encoding and decoding layers facilitates the transfer of spatial information, allowing the model to generate accurate segmentation maps while preserving important anatomical boundaries.

To address the increasing complexity of medical imaging tasks, researchers have introduced several improved versions of the original U-Net framework. Notable extensions include 3D U-Net, which is designed to process volumetric imaging data; ResU-Net, which incorporates residual learning to enhance feature propagation; Attention U-Net, which selectively focuses on clinically relevant regions; Dense U-Net, which improves information flow through dense connectivity; U-Net++, which utilizes redesigned skip pathways for better feature fusion; R2U-Net, which combines recurrent and residual mechanisms; and MultiResUNet, which captures multi-scale contextual information more effectively. These advanced architectures have demonstrated superior performance in extracting discriminative features, modeling complex tissue structures, and improving segmentation precision across diverse biomedical imaging applications.

Figure 1 illustrates the fundamental U-Net architecture together with its prominent variants. These models have been extensively investigated and successfully applied in brain tumour segmentation, organ delineation, lesion detection, and other medical image analysis tasks, contributing significantly to the advancement of automated healthcare diagnostics and treatment planning.

### 3.1 3D U-Net

Unlike the standard U-Net model that operates on two-dimensional image slices, 3D U-Net employs three-dimensional convolutional layers to analyze complete volumetric datasets [27]. This modification allows the network to capture spatial relationships across multiple image planes and effectively learn structural information from full MRI volumes. The 3D convolution operation is expressed as:

$$F_{l+1} = f(W_{3D} * F_l + b_l) \quad (1)$$

where $W_{3D}$ denotes the three-dimensional convolution kernel and $F_l$ represents the input feature volume. The architecture effectively captures spatial dependencies along all three dimensions of the MRI scan.

### 3.2 Residual U-Net (ResU-Net)

ResU-Net [28] enhances the segmentation framework by embedding shortcut pathways within the network architecture, enabling efficient transmission of feature information across layers. These pathways mitigate the degradation problem encountered in very deep models, support faster convergence, and improve the network's ability to learn complex image characteristics. The residual mapping can be represented as:

$$H(x) = F(x) + x \quad (2)$$

where $x$ is the input feature map and $F(x)$ denotes the learned residual function. This formulation enables efficient gradient propagation and improved feature learning.

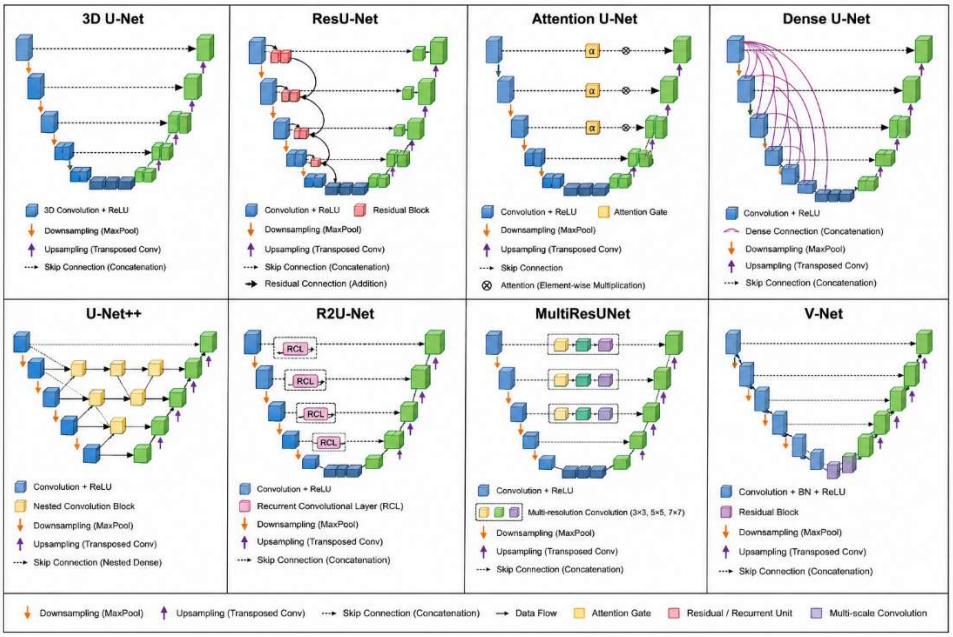



**Figure 1: Evolution of U-Net and Its Advanced Variants for Medical Image Segmentation**

### 3.3 Attention U-Net

The Attention U-Net architecture [29] employs adaptive attention modules to identify and highlight the most meaningful feature responses during learning. By filtering out less informative areas and concentrating computational resources on important anatomical structures, the network achieves greater accuracy in detecting and segmenting regions of interest. The attention coefficient is computed as:

$$\alpha = \sigma(W^T x + b) \quad (3)$$

where $\alpha$represents the attention weight, $W$denotes trainable parameters, and $\sigma(\cdot)$is the sigmoid activation function. The attended feature map is then obtained as:

$$F_{att} = \alpha \odot F \quad (4)$$

where $\odot$denotes element-wise multiplication.

### 3.4 Dense U-Net

Dense U-Net [30] utilizes a connectivity strategy in which each convolutional layer is linked to every earlier layer within the same block. This dense pattern of connections encourages continuous feature reuse, strengthens gradient propagation during training, and improves the richness of learned representations for segmentation tasks. The dense connection can be expressed as:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (5)$$

Where [] denotes feature concatenation and $H_l(\cdot)$ represents the transformation function at layer $l$. This design encourages feature reuse and improved information flow.

### 3.5 U-Net++

U-Net++ [31] refines the information exchange process between the contracting and expanding sections of the network through a series of densely connected intermediate layers. This redesigned connectivity structure improves feature compatibility across different network depths, enabling more effective feature fusion and enhanced segmentation performance. The dense skip connection can be represented as:

$$X^{i,j} = H([X^{i,0}, X^{i,1}, \dots, X^{i,j-1}]) \quad (6)$$

where $X^{i,j}$denotes the feature map at depth $i$and stage $j$. The nested architecture facilitates more effective feature fusion.

### 3.6 R2U-Net

R2U-Net [32] integrates recurrent convolutional operations together with residual connections to improve feature refinement within the network. The recurrent component enables iterative enhancement of feature maps, while the residual links support stable gradient flow, resulting in more robust learning and improved segmentation accuracy. The recurrent operation is defined as:

$$h_t = f(W_x x_t + W_h h_{t-1} + b) \quad (7)$$

where $h_t$is the hidden state at iteration $t$, $W_x$and $W_h$are trainable weights, and $x_t$is the input feature map. The recurrent mechanism enhances contextual feature extraction.

### 3.7 MultiResUNet

MultiResUNet [33] is designed to capture information at different spatial resolutions by applying convolutional operations with multiple receptive field sizes within the same architecture. This multi-scale feature extraction strategy allows the model to represent both fine details and broader contextual patterns more effectively, leading to improved segmentation performance in complex medical images. The multi-resolution feature representation is given by:

$$F_{MR} = [F_{3\times3}, F_{5\times5}, F_{7\times7}] \quad (8)$$

where the outputs from different convolution kernels are concatenated to form a comprehensive feature representation. This enables the network to capture both local and global image characteristics.

### 3.8 V-Net

V-Net [34] is a deep learning architecture developed specifically for volumetric medical image segmentation tasks. It incorporates residual connections to support efficient training of deep networks and is trained using an objective function based on the Dice similarity measure. This direct optimization of overlap between predicted and ground truth regions makes it particularly effective for handling imbalanced medical datasets and producing accurate 3D segmentations. The Dice Similarity Coefficient (DSC) used in V-Net is expressed as:

$$DSC = \frac{2\sum_i p_i g_i}{\sum_i p_i^2 + \sum_i g_i^2} \quad (9)$$

where $p_i$ and $g_i$ denote the predicted and ground truth voxel labels, respectively. This objective function improves overlap between predicted and actual segmentation regions.

## 4. Proposed U-Net Based Brain Tumour Segmentation Framework

The proposed framework employs a U-Net-based deep learning architecture for accurate brain tumour segmentation from multimodal MRI images. The overall workflow consists of dataset preparation, image pre-processing, feature extraction using the encoder path, feature reconstruction using the decoder path, and final tumour segmentation. The model utilizes four MRI modalities, namely T1, T1ce, T2, and FLAIR, to capture complementary tumour information and improve segmentation performance.

Figure 2 illustrates the proposed U-Net-based brain tumour segmentation framework. The framework begins with MRI image acquisition and preprocessing, followed by feature extraction and tumour segmentation using a U-Net architecture. To evaluate the effectiveness of different segmentation models, the conventional U-Net can be replaced with its advanced variants, including 3D U-Net, ResU-Net, Attention U-Net, Dense U-Net, U-Net++, R2U-Net, and MultiResUNet.

### 4.1 Input Representation

The proposed segmentation framework utilizes multimodal MRI data to exploit the complementary information provided by different imaging modalities. Let the input MRI volume be represented as:

$$X = \{X_{T1}, X_{T1ce}, X_{T2}, X_{FLAIR}\} \quad (1)$$

where $X_{T1}$, $X_{T1ce}$, $X_{T2}$, and $X_{FLAIR}$ denote the T1-weighted, T1-contrast enhanced, T2-weighted, and FLAIR MRI modalities, respectively. Each modality is represented as a three-dimensional volume of dimensions $H \times W \times D$, where $H$, $W$, and $D$ correspond to the height, width, and depth of the MRI scan. The use of multiple modalities enables the model to capture diverse anatomical and pathological characteristics of brain tumors. For instance, T1ce images highlight actively enhancing tumor regions, while FLAIR images are highly effective in identifying edema and infiltrative tumor tissues.

The corresponding ground truth segmentation mask is denoted as:

$$Y \in \{0,1,2,3\}^{H \times W \times D} \quad (2)$$

where each voxel is assigned one of four class labels. Label 0 represents the background region, label 1 corresponds to oedema, label 2 indicates the tumour core, and label 3 represents the enhancing tumour region. These annotations are provided by expert radiologists and serve as the reference standard during model training.

The objective of the segmentation network is to learn a mapping function:

$$f: X \rightarrow Y \quad (3)$$

that accurately predicts the tumour label for every voxel in the input MRI volume. By leveraging multimodal information and voxel-wise annotations, the proposed U-Net model learns both local texture features and

global contextual information, enabling precise delineation of tumour boundaries and sub-regions. This voxel-level classification approach facilitates accurate identification of the Whole Tumour (WT), Tumour Core (TC), and Enhancing Tumour (ET), which are critical for diagnosis, treatment planning, and monitoring disease progression.

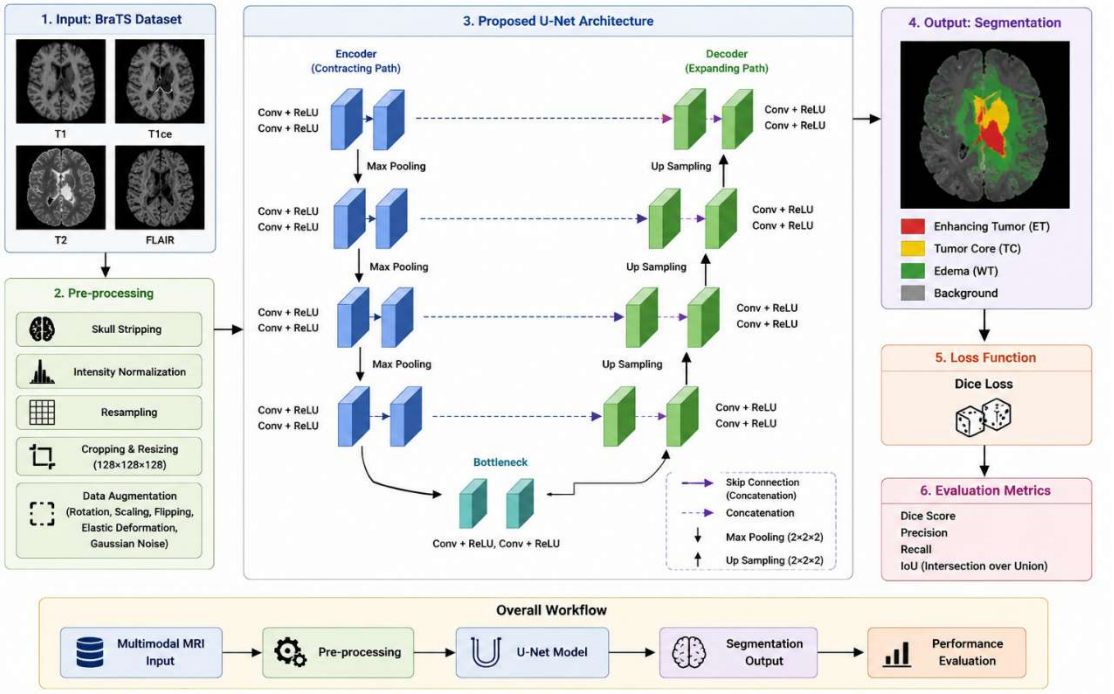



**Figure 2: Proposed U Net based Brain Tumour Segmentation Framework**

### 4.2 Pre-processing Stage

Pre-processing plays a crucial role in improving the quality and consistency of MRI scans before they are provided to the segmentation network. Since MRI images are acquired using different scanners and imaging protocols, variations in intensity distribution, image resolution, and anatomical coverage may exist. Therefore, several pre-processing operations are performed to standardize the data and enhance model performance.

#### 4.2.1 Skull Stripping

The first step involves skull stripping, which removes non-brain tissues such as the skull, scalp, and surrounding structures from the MRI scans. Eliminating these irrelevant regions allows the model to focus solely on brain tissues and tumour structures, thereby reducing computational complexity and improving segmentation accuracy. If $I(x, y, z)$represents the original MRI volume and $M(x, y, z)$denotes the binary brain mask, the skull-stripped image is obtained as:

$$I_s(x, y, z) = I(x, y, z) \times M(x, y, z) \quad (4)$$

where the mask retains brain voxels while suppressing non-brain regions.

#### 4.2.2 Intensity Normalization

MRI scans often exhibit intensity variations due to differences in acquisition devices and scanning parameters. To minimize these variations and achieve a consistent intensity distribution across all samples, Z-score normalization is applied. The normalized image is computed as:

$$X_{norm} = \frac{X-\mu}{\sigma} \quad (5)$$

where $\mu$and $\sigma$represent the mean and standard deviation of voxel intensities, respectively.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i \quad (6)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2} \quad (7)$$

where $N$denotes the total number of voxels in the image volume. This normalization process transforms the intensity values to have zero mean and unit variance, facilitating stable and efficient network training.

### 4.2.3 Image Resizing

To ensure uniformity across all MRI volumes and enable batch processing during network training, the pre-processed images are resized to a fixed spatial dimension of $128 \times 128 \times 128$. Standardizing the image dimensions reduces memory requirements and allows the U-Net architecture to process all samples consistently. Furthermore, resizing helps maintain a balance between computational efficiency and preservation of anatomical details necessary for accurate tumor segmentation.
Through skull stripping, intensity normalization, and image resizing, the MRI data become standardized and suitable for subsequent feature extraction and segmentation using the proposed U-Net model.

## 4.3 Encoder Network

The encoder constitutes the contracting path of the U-Net architecture and is responsible for extracting hierarchical feature representations from the input MRI volumes. It progressively captures both low-level and high-level features through a sequence of convolution and max-pooling operations. The initial layers learn basic image characteristics such as edges, textures, and intensity variations, while deeper layers extract more complex semantic information associated with tumour regions and surrounding tissues.

For an input feature map $F_l$, the convolution operation at the $l^{th}$layer is expressed as:

$$F_{l+1} = f(W_l * F_l + b_l) \quad (8)$$

where $W_l$ represents the convolution kernel, $b_l$denotes the bias term, $*$indicates the convolution operation, and $f(\cdot)$ is the activation function. In this work, the Rectified Linear Unit (ReLU) activation function is employed due to its computational efficiency and ability to alleviate the vanishing gradient problem. The ReLU function is defined as:

$$ReLU(x) = \max(0, x) \quad (9)$$

Following each convolutional block, max-pooling is applied to reduce the spatial dimensions of the feature maps while retaining the most discriminative information. This operation increases the receptive field of the network and enables the extraction of contextual information from larger image regions. The max-pooling operation is given by:

$$P(i,j,k) = \max_{(m,n,p)\in\Omega} F(m,n,p) \quad (10)$$

where Ω denotes the pooling window and $P(i,j,k)$represents the resulting pooled feature map. As the network progresses deeper into the encoder path, the spatial resolution decreases while the number of feature channels increases. This hierarchical representation enables the network to learn increasingly abstract and meaningful features related to tumour morphology, shape, texture, and intensity characteristics. Consequently, the encoder generates rich feature maps that serve as the foundation for accurate tumour localization and segmentation in the subsequent decoder stage.

### 4.4 Decoder Network

The decoder forms the expanding path of the U-Net architecture and is responsible for reconstructing the segmentation map from the high-level feature representations extracted by the encoder. While the encoder gradually reduces the spatial dimensions to capture semantic information, the decoder progressively restores the original image resolution to achieve precise localization of tumour regions. This reconstruction process enables the network to generate pixel-wise segmentation maps while preserving important anatomical details. The decoder begins by up sampling the feature maps obtained from the bottleneck layer. The up-sampling operation increases the spatial dimensions of the feature maps and can be performed using transposed convolution or interpolation-based techniques. Mathematically, the up-sampling process is represented as:

$$U_l = Up(F_l) \tag{11}$$

where $F_l$ denotes the input feature map at the $l^{th}$ decoder level and $Up(\cdot)$represents the upsampling operation. To recover the spatial information lost during the down sampling process, the up sampled feature map is combined with the corresponding feature map from the encoder through skip connections. This operation allows the decoder to utilize both high-level semantic information and low-level spatial features. The concatenation process is expressed as:

$$C_l = [U_l, E_l] \tag{12}$$

where $U_l$represents the upsampled decoder feature map, $E_l$denotes the corresponding encoder feature map, and [ ]indicates channel-wise concatenation.
The concatenated feature map is then refined through convolutional operations to learn meaningful representations and improve segmentation accuracy. The convolution operation in the decoder is given by:

$$F_l' = f(W_l' * C_l + b_l') \tag{13}$$

where $W_l'$and $b_l'$represent the convolution kernel and bias, respectively, and $f(\cdot)$denotes the activation function, typically ReLU.
Through successive up sampling and convolution operations, the decoder gradually reconstructs the spatial structure of the input image while retaining the contextual information learned by the encoder. The integration of encoder and decoder features through skip connections helps preserve fine-grained details such as tumour boundaries and tissue transitions. Consequently, the decoder generates high-resolution feature maps that facilitate accurate delineation of tumour sub-regions, including oedema, tumour core, and enhancing tumour tissues.

### 4.5 Skip Connections

A key feature that distinguishes U-Net from conventional CNNs is the incorporation of skip connections. During the encoding process, repeated pooling operations reduce the spatial resolution of feature maps, which may lead to the loss of fine-grained structural details and boundary information. Such information is particularly important in medical image segmentation, where accurate delineation of tumour boundaries is essential for reliable diagnosis and treatment planning.

To address this issue, feature maps from the encoder are directly transferred to the corresponding decoder layers through skip connections. These transferred feature maps contain detailed spatial information that complements the high-level semantic features learned by the decoder. The skip connection operation can be expressed as:

$$S_l = Concat(E_l, U_l) \tag{14}$$

where $E_l$ denotes the feature map obtained from the encoder at level $l$, $U_l$ represents the corresponding upsampled decoder feature map, and $Concat(\cdot)$refers to channel-wise concatenation.

The integration of encoder and decoder features enables the network to simultaneously utilize local spatial details and global contextual information. As a result, the segmentation model can accurately identify tumour boundaries, preserve anatomical structures, and improve the localization of small tumour regions. The use of skip connections significantly enhances segmentation performance and contributes to the superior accuracy of the U-Net architecture in medical image analysis.

### 4.6 Output Segmentation Layer

The final stage of the U-Net architecture generates the voxel-wise segmentation map. After the decoder reconstructs the high-resolution feature maps, a $1 \times 1 \times 1$convolution layer is applied to transform the extracted features into class-specific scores. This operation reduces the number of feature channels to the total number of segmentation classes and produces a probability distribution for each voxel.

The voxel-wise class probabilities are computed using the Softmax activation function:

$$P_c(x) = \frac{e^{z_c(x)}}{\sum_{k=1}^{C} e^{z_k(x)}} \tag{15}$$

where $P_c(x)$represents the probability of voxel $x$belonging to class $c$, $z_c(x)$denotes the output score (logit) associated with class $c$, and $C$is the total number of segmentation classes.
The Softmax function converts the raw output scores into normalized probability values ranging between 0 and 1, ensuring that the sum of probabilities across all classes equals one. This probabilistic representation allows the network to estimate the likelihood of each voxel belonging to different tissue categories.
The final segmentation mask is obtained by assigning each voxel to the class with the highest probability:

$$\hat{Y}(x) = \arg\max_{c} P_c(x) \tag{16}$$

where $\hat{Y}(x)$denotes the predicted class label of voxel $x$.

For the BraTS dataset, the output layer classifies each voxel into one of the predefined categories, including background, oedema, tumour core, and enhancing tumour regions. The resulting segmentation map provides a detailed delineation of tumour sub-regions, enabling accurate assessment of tumour extent and supporting clinical decision-making. By combining hierarchical feature extraction, skip connections, and voxel-wise classification, the proposed U-Net model achieves precise and reliable brain tumour segmentation.

### 4.7 Loss Function

To optimize segmentation performance, the Dice Loss is employed. The Dice Similarity Coefficient (DSC) between the predicted segmentation $\hat{Y}$and ground truth mask $Y$is given by:

$$DSC = \frac{2\sum_{i=1}^{N} Y_i \hat{Y}_i + \epsilon}{\sum_{i=1}^{N} Y_i + \sum_{i=1}^{N} \hat{Y}_i + \epsilon} \tag{17}$$

where $\epsilon$is a small constant added to prevent division by zero.

The Dice Loss is defined as:

$$L_{Dice} = 1 - DSC \tag{18}$$

Minimizing Dice Loss encourages greater overlap between predicted and actual tumour regions.

## 5. Results and Discussion

The BraTS 2021 dataset [35] consists of multi-institutional, multi-parametric MRI scans collected from patients diagnosed with gliomas. Each subject contains four MRI modalities: T1, T1ce, T2, and FLAIR, along with expert-annotated segmentation masks. The dataset includes a total of 2,040 subjects**,** divided into training, validation, and testing subsets. The effectiveness of the proposed U-Net-based brain tumour segmentation model is evaluated using five widely accepted performance metrics: Accuracy, Dice Score, Precision, Recall, and Intersection over Union (IoU). These metrics provide a comprehensive assessment of the segmentation quality by measuring the agreement between the predicted tumour region and the ground-truth annotation.

**Accuracy** represents the overall proportion of correctly classified pixels, including both tumour and background regions. It is calculated as:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (19)$$

where TP denotes True Positives, TN denotes True Negatives, FP denotes False Positives, and FN denotes False Negatives.

**Dice Score**, also known as the Dice Similarity Coefficient (DSC), measures the overlap between the predicted segmentation and the ground-truth mask. A higher Dice Score indicates better segmentation performance and is computed as:

$$\text{Dice} = (2 \times TP) / (2 \times TP + FP + FN) \quad (20)$$

**Precision** measures the proportion of correctly identified tumour pixels among all pixels predicted as tumour. It evaluates the reliability of positive predictions and is defined as:

$$\text{Precision} = TP / (TP + FP) \quad (21)$$

**Recall**, also referred to as Sensitivity, measures the ability of the model to correctly detect actual tumour pixels. A higher recall value indicates fewer missed tumour regions and is calculated as:

$$\text{Recall} = TP / (TP + FN) \quad (22)$$

**Intersection over Union (IoU)** quantifies the overlap between the predicted tumour region and the ground-truth region relative to their combined area. It is one of the most commonly used metrics for evaluating segmentation performance and is given by:

$$\text{IoU} = TP / (TP + FP + FN) \quad (23)$$

These evaluation metrics collectively provide a detailed understanding of model performance. Accuracy measures overall classification effectiveness, Dice Score and IoU evaluate segmentation overlap quality, Precision assesses the correctness of tumour predictions, and Recall determines the model's ability to identify all tumour pixels. Therefore, the combined use of these metrics ensures a robust and reliable evaluation of the proposed brain tumour segmentation framework.

Figure 3, illustrates the training loss curves of the eight U-Net variants employed for brain tumour segmentation. It can be observed that the loss values consistently decrease with increasing training epochs for all models, indicating effective learning and convergence during the training process. The conventional U-Net exhibits the slowest convergence and retains a relatively higher final loss value compared to its advanced counterparts. Variants such as 3D U-Net, ResU-Net, Attention U-Net, Dense U-Net, U-Net++, and R2U-Net

demonstrate improved convergence behaviour due to their enhanced feature extraction and representation capabilities. Among all the models,

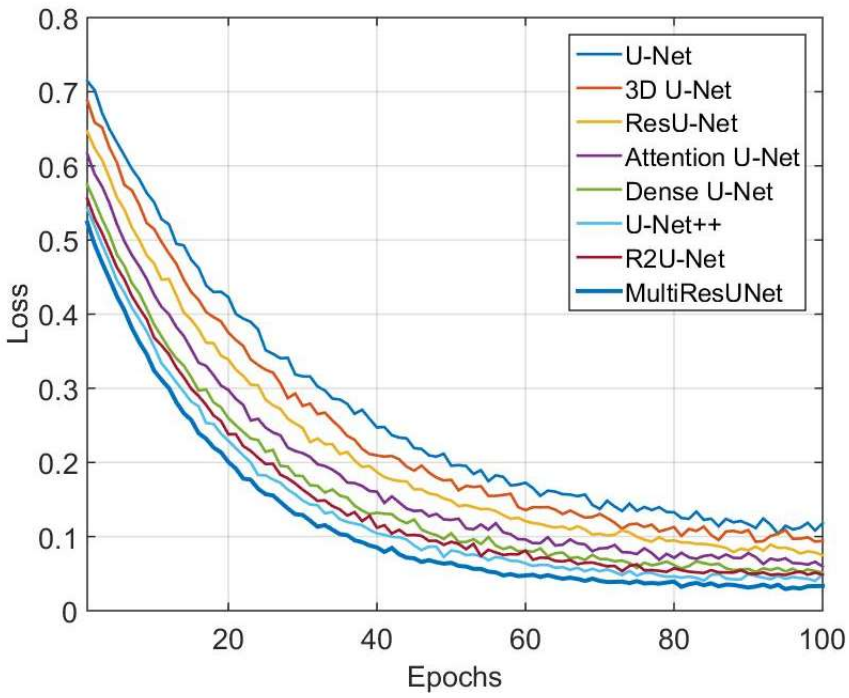
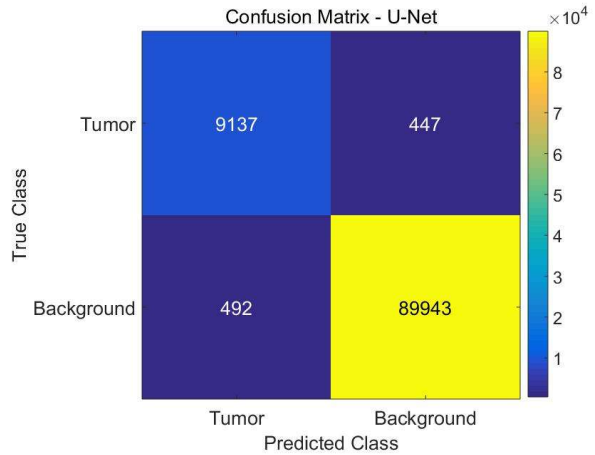
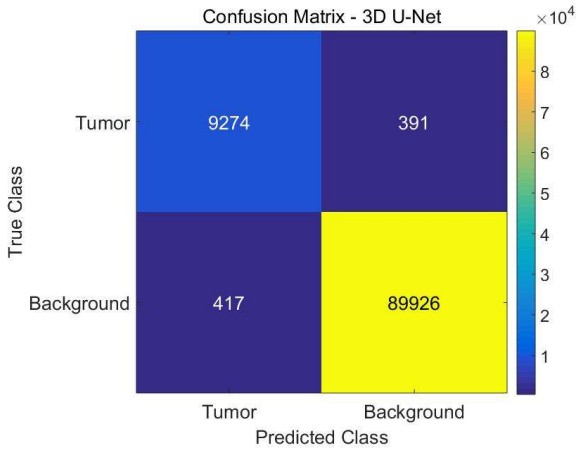



**Figure 3: Loss vs. Epochs**

MultiResUNet achieves the fastest convergence and the lowest final loss, highlighting its superior ability to capture multiscale contextual information and accurately segment tumour regions. The reduced loss values obtained by MultiResUNet indicate better optimization performance and improved generalization capability, which is further reflected in its higher Accuracy, Dice Score, Precision, Recall, and IoU values. Overall, the loss curves confirm that advanced U-Net variants outperform the conventional U-Net architecture, with MultiResUNet providing the most effective learning performance for brain tumour segmentation.




**(a)**




**(b)**

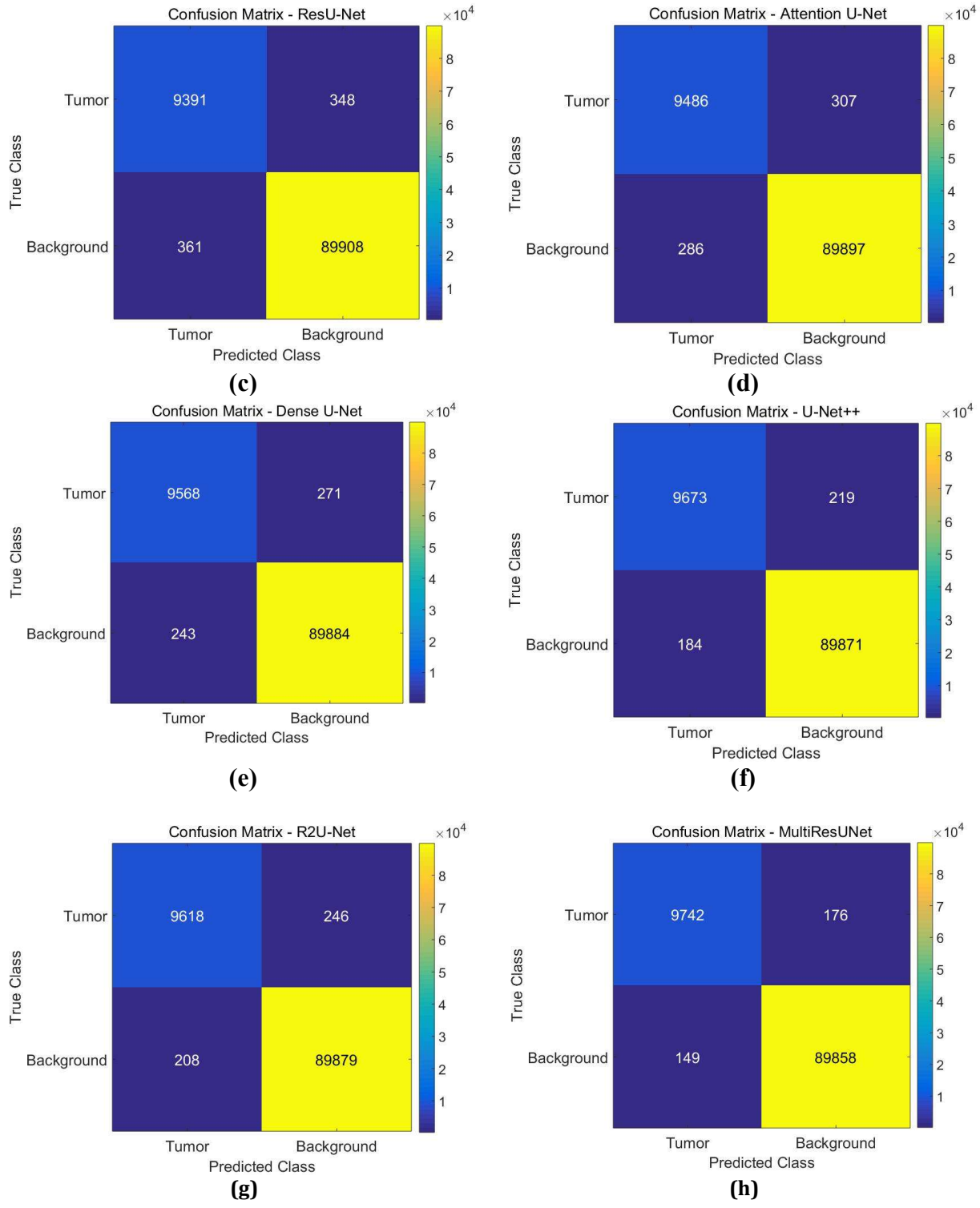



**Figure 4: Confusion Matrices of U-Net and Its Variants for Brain Tumour Segmentation**

Figure 4(a)–4(h) present the confusion matrices obtained for the eight U-Net-based brain tumour segmentation models, namely U-Net, 3D U-Net, ResU-Net, Attention U-Net, Dense U-Net, U-Net++, R2U-Net, and MultiResUNet, respectively. The confusion matrices provide a comprehensive evaluation of the segmentation

performance by illustrating the distribution of TP, FP, FN, and TN for each model. Higher TP and TN values, together with lower FP and FN values, indicate superior segmentation accuracy and reliability.

Figure 4(a) corresponds to the conventional U-Net architecture, which achieved 9,137 true positives and 89,943 true negatives, with 447 false positives and 492 false negatives. Although the model demonstrates satisfactory segmentation performance, the relatively higher number of misclassified pixels suggests limitations in accurately identifying complex tumour boundaries. Figure 4(b) shows the confusion matrix of the 3D U-Net model, which improves segmentation performance by increasing the number of true positives to 9,274 and reducing false positives and false negatives to 391 and 417, respectively. This improvement is primarily due to the ability of 3D convolutions to exploit volumetric spatial information present in MRI data.

The confusion matrix of ResU-Net shown in Figure 4(c) further improves the segmentation outcome by incorporating residual connections, resulting in 9,391 true positives, 348 false positives, and 361 false negatives. Figure 4(d) presents the results of Attention U-Net, which utilizes attention mechanisms to focus on relevant tumour regions. Consequently, the model achieves 9,486 true positives while reducing false positives and false negatives to 307 and 286, respectively. These results demonstrate the effectiveness of attention-guided feature learning in enhancing tumour localization.

Figure 4(e) illustrates the confusion matrix of Dense U-Net, which employs dense connectivity to facilitate feature reuse and efficient information propagation across layers. The model records 9,568 true positives, 271 false positives, and 243 false negatives, indicating improved segmentation accuracy compared to the previous architectures. Figure 4(f) corresponds to U-Net++, which introduces nested skip connections for enhanced feature fusion between encoder and decoder pathways. As a result, the model achieves 9,673 true positives while reducing false positives and false negatives to 219 and 184, respectively.

Similarly, Figure 4(g) presents the confusion matrix of R2U-Net, which integrates recurrent residual units to strengthen contextual feature extraction. The model achieves 9,618 true positives with only 246 false positives and 208 false negatives, demonstrating robust segmentation performance. Finally, Figure 4(h) illustrates the confusion matrix of the proposed MultiResUNet architecture. Among all evaluated models, MultiResUNet achieves the highest number of true positives (9,742) and maintains a high number of true negatives (89,858), while significantly reducing false positives and false negatives to only 176 and 149, respectively. These results indicate that MultiResUNet is highly effective in capturing multiscale features and preserving fine tumour details, leading to more accurate segmentation outcomes.

Overall, the confusion matrices shown in Figures 4(a)–4(h) reveal a progressive improvement in segmentation performance from the conventional U-Net to its advanced variants. The reduction in false positives and false negatives across successive architectures demonstrates enhanced discrimination between tumour and healthy tissue. Among all the models, MultiResUNet exhibits the most favourable confusion matrix characteristics, confirming its superior capability for brain tumour segmentation and supporting the high Accuracy, Dice Score, Precision, Recall, and IoU values reported in this study.

Figure 5 illustrates the Receiver Operating Characteristic (ROC) curves and the corresponding Area Under the Curve (AUC) values for the U-Net architecture and its variants employed for brain tumour segmentation. The ROC curve provides a graphical representation of the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR), while the AUC value quantifies the overall discriminative capability of a model. A higher AUC value indicates better segmentation performance and a stronger ability to distinguish tumour pixels from non-tumour regions.

As observed from Figure 5, all U-Net-based architectures exhibit ROC curves that are significantly above the diagonal reference line, indicating effective classification and segmentation performance. The conventional U-Net achieves the lowest AUC value among the evaluated models, reflecting its relatively limited capability in capturing complex tumour features. The performance gradually improves with the introduction of advanced architectural modifications such as volumetric feature extraction in 3D U-Net, residual learning in ResU-Net, attention mechanisms in Attention U-Net, and dense connectivity in Dense U-Net. Similarly, U-Net++ and R2U-Net demonstrate enhanced discriminative performance due to improved feature fusion and recurrent residual learning strategies.

Among all the evaluated models, MultiResUNet achieves the highest ROC curve and the largest AUC value, approaching 0.95. The superior performance of MultiResUNet can be attributed to its multiresolution feature extraction capability, which effectively captures both local and global contextual information while preserving fine tumour boundary details. The larger AUC value indicates that MultiResUNet provides the best balance between sensitivity and specificity, resulting in fewer false positive and false negative predictions. These findings are consistent with the Accuracy, Dice Score, Precision, Recall, and IoU results reported in this study,

where MultiResUNet achieved the highest segmentation performance. Overall, the ROC and AUC analysis confirms that MultiResUNet is the most effective architecture among the considered U-Net variants for accurate and reliable brain tumour segmentation.

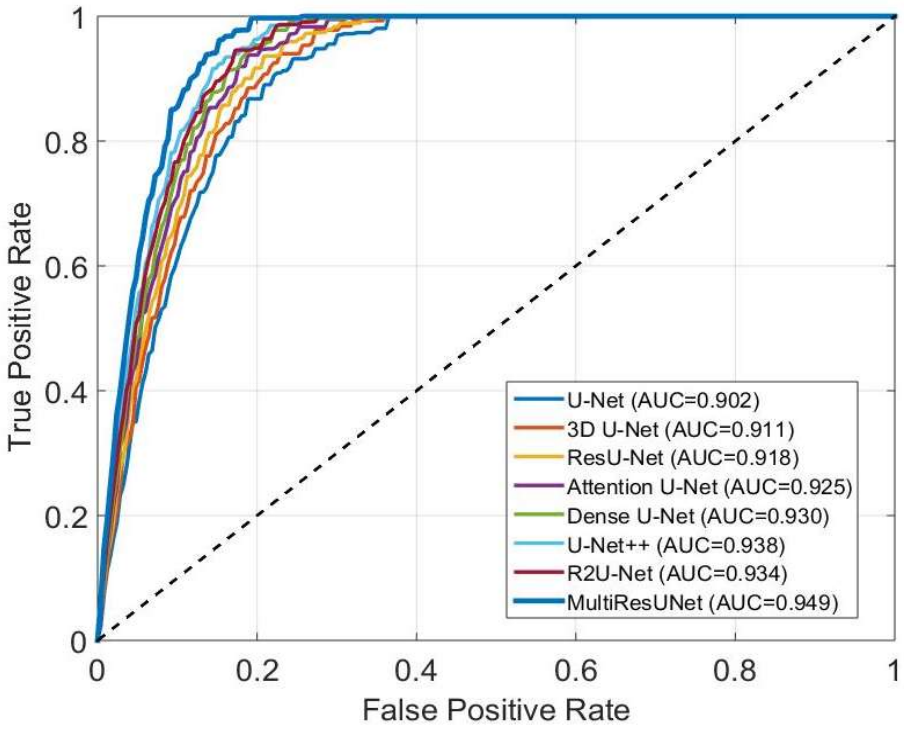



**Figure 5: RoC and AUC for U-Net and its Variants for Brain Tumour Segmentation**

The Table 1, presents a comparative performance analysis of different U-Net variants for brain tumour segmentation. It can be observed that all advanced U-Net architectures outperform the conventional U-Net in terms of Accuracy, Dice Score, Precision, Recall, and IoU. Among the evaluated models, MultiResUNet achieves the highest performance with an Accuracy of 99.67%, Dice Score of 98.36%, Precision of 98.23%, Recall of 98.49%, and IoU of 96.77%, demonstrating its superior ability to capture multiscale contextual information and accurately delineate tumour regions. U-Net++ and R2U-Net also exhibit strong segmentation performance, whereas the conventional U-Net records the lowest values across all evaluation metrics. These results confirm the effectiveness of incorporating advanced architectural modifications into the original U-Net framework for improved brain tumour segmentation.

**Table 1: Performance Comparison of Different U-Net Variants for Brain Tumour Segmentation**

| Method | Accuracy (%) | Dice Score (%) | Precision (%) | Recall (%) | IoU (%) |
|---|---|---|---|---|---|
| U-Net | 99.06 | 95.11 | 95.34 | 94.89 | 90.68 |
| 3D U-Net | 99.19 | 95.83 | 95.95 | 95.70 | 91.99 |
| ResU-Net | 99.29 | 96.36 | 96.43 | 96.30 | 92.98 |
| Attention U-Net | 99.41 | 96.97 | 96.87 | 97.07 | 94.12 |
| Dense U-Net | 99.49 | 97.38 | 97.25 | 97.52 | 94.90 |
| U-Net++ | 99.60 | 97.96 | 97.79 | 98.13 | 96.00 |
| R2U-Net | 99.55 | 97.69 | 97.51 | 97.88 | 95.49 |
| **MultiResUNet** | **99.67** | **98.36** | **98.23** | **98.49** | **96.77** |

Table 2 presents a comparative analysis of state-of-the-art deep learning methods developed for brain tumour segmentation and classification using MRI images. The table highlights the diversity of approaches proposed in the literature, ranging from conventional CNNs to advanced transfer learning, ensemble learning, and hybrid deep learning frameworks. The reported results demonstrate significant progress in the application of deep learning techniques for automated brain tumour analysis.

**Table 2: Comparison of the state-of-the-art methods for Brain Tumour Segmentation**

| Authors | Method | Performance Results |
|---|---|---|
| Justin S. Paula et al. [12] | CNN | Accuracy = 91.43% |
| Hossam et al. [13] | CNN, Deep Neural Network | First Model Accuracy = 96.03%; Second Model Accuracy = 98.70% |
| Milica et al. [14] | CNN with Adam Optimization | Accuracy (10-Fold) = 95.40% |
| Neelum et al. [15] | Inception-V3, DenseNet-201 | Accuracy = 99.34%, 99.51% |
| Khan et al. [16] | K-means Clustering, VGG19 | Accuracy = 94.06% |
| Pernas et al. [17] | Multiscale CNN | Accuracy = 97.30% |
| Masood et al. [18] | Mask R-CNN with DenseNet-41 Backbone | Segmentation Accuracy = 96.30%; Classification Accuracy = 98.34% |
| Dipu et al. [19] | CNN, YOLOv5 | CNN Accuracy = 95.78%; YOLOv5 Accuracy = 85.95% |
| Ramesh et al. [20] | VGG16 | Accuracy = 96.00% |
| Subhashis et al. [21] | VGGNet, ResNet, ConvNet | Accuracy = 95.00% |
| Khan et al. [22] | HDL2BT, CNN | Precision = 92.13% |
| Isaza et al. [23] | Transfer Learning, ResNet50, PCA-Based Augmentation | F1-Score = 92.34% |
| Raza et al. [24] | GoogleNet with Leaky ReLU | Accuracy = 99.67% |
| Mirza Mumtaz Zahoor et al. [25] | Deep-Boosted Features, Ensemble Learning, BRAIN-RENet | Accuracy = 99.56% |
| Sadia Anjum et al. [26] | GoogleNet and ResNet101 with Transfer Learning | Accuracy = 99.33% |
| Proposed Method | MultiResUNet | Accuracy = 99.67%; Dice = 98.36%; Precision = 98.23%; Recall = 98.49%; IoU = 96.77% |

Among the earlier studies, Justin S. Paul et al. [12] employed a CNN-based framework and achieved an accuracy of 91.43%, demonstrating the feasibility of deep learning for brain tumour classification. Hossam et al. [13] further improved performance using CNN and deep neural network architectures, reporting accuracies of 96.03% and 98.70% for two different models. Similarly, Milica et al. [14] utilized a CNN optimized using the Adam optimization algorithm and achieved a 10-fold cross-validation accuracy of 95.40%. These studies established the effectiveness of CNN-based approaches in extracting discriminative features from MRI images. Subsequent research focused on more sophisticated deep learning architectures. Neelum et al. [15] employed Inception-V3 and DenseNet-201 models and reported accuracies of 99.34% and 99.51%, respectively, demonstrating the effectiveness of deep feature extraction and transfer learning. Khan et al. [16] combined K-means clustering with the VGG19 architecture to perform segmentation and classification, achieving an accuracy of 94.06%. Díaz-Pernas et al. [17] proposed a multiscale CNN framework capable of capturing tumour information at different spatial resolutions and obtained an accuracy of 97.30%.

Further improvements were achieved through advanced segmentation networks. Masood et al. [18] introduced a Mask R-CNN model with a DenseNet-41 backbone, achieving segmentation and classification accuracies of

96.30% and 98.34%, respectively. Dipu et al. [19] compared CNN and YOLOv5 architectures, reporting accuracies of 95.78% and 85.95%, respectively, indicating that CNN-based methods were more suitable for the considered brain tumour dataset. Ramesh et al. [20] utilized a VGG16-based framework and achieved an accuracy of 96.00%, while Banerjee et al. [21] employed deep radiomics using VGGNet, ResNet, and ConvNet architectures to obtain an accuracy of 95.00%.

Recent studies have explored hybrid and transfer learning approaches to further enhance performance. Abdul Hannan Khan et al. [22] proposed a hierarchical deep learning-based brain tumour identification framework (HDL2BT) and achieved a precision of 92.13%. Anaya-Isaza et al. [23] integrated transfer learning with ResNet50 and PCA-based data augmentation, resulting in an F1-score of 92.34%. Raza et al. [24] developed a hybrid deep learning model based on GoogleNet and Leaky ReLU activation functions, achieving an impressive accuracy of 99.67%, which represents one of the highest performances reported in the literature. Similarly, Zahoor et al. [25] proposed a deep hybrid boosted and ensemble learning framework incorporating BRAIN-RENet and ensemble classifiers, obtaining an accuracy of 99.56%. Anjum et al. [26] employed transfer learning using GoogleNet and ResNet101 architectures and achieved an accuracy of 99.33%, confirming the effectiveness of pretrained deep networks for brain tumour analysis.

The proposed MultiResUNet-based framework demonstrates highly competitive performance when compared with existing state-of-the-art methods. The proposed model achieves an accuracy of 99.67%, matching the highest accuracy reported by Raza et al. [24]. In addition, the proposed approach provides a comprehensive segmentation evaluation with a Dice Score of 98.36%, Precision of 98.23%, Recall of 98.49%, and IoU of 96.77%. Unlike many existing studies that primarily report classification accuracy, the proposed framework emphasizes precise tumour segmentation by effectively capturing multiscale contextual information and preserving fine-grained tumour boundaries. The superior Dice Score and IoU values indicate excellent overlap between the predicted segmentation masks and ground-truth annotations, while the high precision and recall values demonstrate reliable tumour detection with minimal false positives and false negatives.

Overall, the results presented in Table 2 indicate that deep learning architectures have significantly advanced the field of brain tumour analysis. Although several existing methods report high classification accuracy, the proposed MultiResUNet framework provides a balanced and robust segmentation performance across multiple evaluation metrics. These findings demonstrate the effectiveness of the proposed architecture for accurate brain tumour segmentation and highlight its potential for supporting clinical diagnosis and treatment planning.

## 7. Conclusion

This paper presented a comprehensive comparative study of U-Net and its advanced variants, including 3D U-Net, ResU-Net, Attention U-Net, Dense U-Net, U-Net++, R2U-Net, and MultiResUNet, for brain tumour segmentation using MRI images. The objective was to evaluate the effectiveness of these architectures in accurately identifying and delineating tumour regions while preserving fine structural details. A U-Net-based segmentation framework was developed, and the performance of each model was assessed using Accuracy, Dice Score, Precision, Recall, and Intersection over Union (IoU). The experimental results demonstrated that all advanced U-Net variants outperformed the conventional U-Net architecture, highlighting the benefits of enhanced feature extraction, multiscale learning, residual connections, attention mechanisms, and dense connectivity. Among the evaluated models, MultiResUNet achieved the best overall performance with an Accuracy of 99.67%, Dice Score of 98.36%, Precision of 98.23%, Recall of 98.49%, and IoU of 96.77%. In comparison, the conventional U-Net achieved an Accuracy of 99.06%, Dice Score of 95.11%, Precision of 95.34%, Recall of 94.89%, and IoU of 90.68%. The confusion matrix analysis further confirmed the superiority of MultiResUNet, which produced the highest number of true positive detections and the lowest number of false positives and false negatives among all considered architectures.

A comparison with recently published state-of-the-art methods revealed that the proposed MultiResUNet framework achieved performance comparable to or better than existing deep learning approaches reported in the literature. The high Dice Score and IoU values indicate excellent agreement between the predicted segmentation masks and the ground-truth annotations, while the high precision and recall values demonstrate reliable tumour detection and boundary localization.

Overall, the findings confirm that MultiResUNet is a highly effective architecture for brain tumour segmentation due to its ability to capture multiscale contextual information and preserve fine-grained tumour features. The proposed framework can assist clinicians in accurate tumour assessment, treatment planning, and disease monitoring. Future work may focus on validating the model on larger multi-institutional datasets,

incorporating multimodal MRI sequences, and integrating transformer-based attention mechanisms to further enhance segmentation accuracy and generalization capability.

## References


1. Louis, David N., Arie Perry, Guido Reifenberger, Andreas Von Deimling, Dominique Figarella-Branger, Webster K. Cavenee, Hiroko Ohgaki, Otmar D. Wiestler, Paul Kleihues, and David W. Ellison. "The 2016 World Health Organization classification of tumors of the central nervous system: a summary." *Acta neuropathologica* 131, no. 6 (2016): 803-820.
2. Bauer, Stefan, Roland Wiest, Lutz-P. Nolte, and Mauricio Reyes. "A survey of MRI-based medical image analysis for brain tumor studies." *Physics in medicine and biology* 58, no. 13 (2013): R97-R129.
3. Menze, Bjoern H., Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren et al. "The multimodal brain tumor image segmentation benchmark (BRATS)." *IEEE transactions on medical imaging* 34, no. 10 (2014): 1993-2024.
4. Khan, Sajid Ullah. "A Comprehensive Review of Automatic Semantic Segmentation of Brain MRI: Techniques, Discussion, Challenges." *UMT Artificial Intelligence Review* 4, no. 1 (2025): 62-82.
5. Wang, Pin, En Fan, and Peng Wang. "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning." *Pattern recognition letters* 141 (2021): 61-67.
6. Chen, Min, Xiaobo Shi, Yin Zhang, Di Wu, and Mohsen Guizani. "Deep feature learning for medical image analysis with convolutional autoencoder neural network." *IEEE transactions on big data* 7, no. 4 (2017): 750-758.
7. Mahmood, Tariq, Amjad Rehman, Tanzila Saba, Lubna Nadeem, and Saeed Ali Omer Bahaj. "Recent advancements and future prospects in active deep learning for medical image segmentation and classification." *IEEE Access* 11 (2023): 113623-113652.
8. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241. Cham: Springer international publishing, 2015.
9. Isensee, Fabian, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H. Maier-Hein. "No new-net." In *International MICCAI brainlesion workshop*, pp. 234-244. Cham: Springer International Publishing, 2018.
10. Çiçek, Özgün, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. "3D U-Net: learning dense volumetric segmentation from sparse annotation." In *International conference on medical image computing and computer-assisted intervention*, pp. 424-432. Cham: Springer International Publishing, 2016.
11. Bakas, Spyridon, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S. Kirby, John B. Freymann, Keyvan Farahani, and Christos Davatzikos. "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features." *Scientific data* 4, no. 1 (2017): 170117.
12. Paul, Justin S., Andrew J. Plassard, Bennett A. Landman, and Daniel Fabbri. "Deep learning for brain tumor classification." In Medical imaging 2017: Biomedical applications in molecular, structural, and functional imaging, vol. 10137, pp. 253-268. SPIE, 2017.
13. Sultan, Hossam H., Nancy M. Salem, and Walid Al-Atabany. "Multi-classification of brain tumor images using deep neural network." *IEEE access* 7 (2019): 69215-69225.
14. Badža, Milica M., and Marko Č. Barjaktarović. "Classification of brain tumors from MRI images using a convolutional neural network." *Applied sciences* 10, no. 6 (2020): 1999.
15. Noreen, Neelum, Sellappan Palaniappan, Abdul Qayyum, Iftikhar Ahmad, Muhammad Imran, and Muhammad Shoaib. "A deep learning model based on concatenation approach for the diagnosis of brain tumor." *IEEE access* 8 (2020): 55135-55144.
16. Khan, Amjad Rehman, Siraj Khan, Majid Harouni, Rashid Abbasi, Sajid Iqbal, and Zahid Mehmood. "Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification." *Microscopy Research and Technique* 84, no. 7 (2021): 1389-1399.
17. Díaz-Pernas, Francisco Javier, Mario Martínez-Zarzuela, Míriam Antón-Rodríguez, and David González-Ortega. "A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network." In *Healthcare*, vol. 9, no. 2, p. 153. MDPI, 2021.

18. Masood, Momina, Tahira Nazir, Marriam Nawaz, Awais Mehmood, Junaid Rashid, Hyuk-Yoon Kwon, Toqeer Mahmood, and Amir Hussain. "A novel deep learning method for recognition and classification of brain tumors from MRI images." *Diagnostics* 11, no. 5 (2021): 744.
19. Dipu, Nadim Mahmud, Sifatul Alam Shohan, and K. M. A. Salam. "Deep learning based brain tumor detection and classification." In *2021 International conference on intelligent technologies (CONIT)*, pp. 1-6. IEEE, 2021.
20. Ramesh, S., S. Sasikala, and Nirmala Paramanandham. "Segmentation and classification of brain tumors using modified median noise filter and deep learning approaches." *Multimedia Tools and Applications* 80, no. 8 (2021): 11789-11813.
21. Banerjee, Subhashis, Sushmita Mitra, Francesco Masulli, and Stefano Rovetta. "Deep radiomics for brain tumor detection and classification from multi-sequence MRI." *arXiv preprint arXiv:1903.09240* (2019).
22. Khan, Abdul Hannan, Sagheer Abbas, Muhammad Adnan Khan, Umer Farooq, Wasim Ahmad Khan, Shahan Yamin Siddiqui, and Aiesha Ahmad. "Intelligent model for brain tumor identification using deep learning." *Applied Computational Intelligence and Soft Computing* 2022, no. 1 (2022): 8104054.
23. Anaya-Isaza, Andrés, and Leonel Mera-Jiménez. "Data augmentation and transfer learning for brain tumor detection in magnetic resonance imaging." *IEEE access* 10 (2022): 23217-23233.
24. Raza, Asaf, Huma Ayub, Javed Ali Khan, Ijaz Ahmad, Ahmed S. Salama, Yousef Ibrahim Daradkeh, Danish Javeed, Ateeq Ur Rehman, and Habib Hamam. "A hybrid deep learning-based approach for brain tumor classification." *Electronics* 11, no. 7 (2022): 1146.
25. Zahoor, Mirza Mumtaz, Shahzad Ahmad Qureshi, Sameena Bibi, Saddam Hussain Khan, Asifullah Khan, Usman Ghafoor, and Muhammad Raheel Bhutta. "A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI." *Sensors* 22, no. 7 (2022): 2726.
26. Anjum, Sadia, Lal Hussain, Mushtaq Ali, Monagi H. Alkinani, Wajid Aziz, Sabrina Gheller, Adeel Ahmed Abbasi, Ali Raza Marchal, Harshini Suresh, and Tim Q. Duong. "Detecting brain tumors using deep learning convolutional neural network with transfer learning approach." *International Journal of Imaging Systems and Technology* 32, no. 1 (2022): 307-323.
27. Çiçek, Özgün, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. "3D U-Net: learning dense volumetric segmentation from sparse annotation." In *International conference on medical image computing and computer-assisted intervention*, pp. 424-432. Cham: Springer International Publishing, 2016.
28. Zhang, Zhengxin, Qingjie Liu, and Yunhong Wang. "Road extraction by deep residual u-net." *IEEE Geoscience and Remote Sensing Letters* 15, no. 5 (2018): 749-753.
29. Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori et al. "Attention u-net: Learning where to look for the pancreas." *arXiv preprint arXiv:1804.03999* (2018).
30. Wang, Chang, Zongya Zhao, Qiongqiong Ren, Yongtao Xu, and Yi Yu. "Dense U-net based on patch-based learning for retinal vessel segmentation." *Entropy* 21, no. 2 (2019): 168.
31. Micallef, Neil, Dylan Seychell, and Claude J. Bajada. "Exploring the u-net++ model for automatic brain tumor segmentation." *ieee Access* 9 (2021): 125523-125539.
32. Alom, Md Zahangir, Mahmudul Hasan, Chris Yakopcic, Tarek M. Taha, and Vijayan K. Asari. "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation." *arXiv preprint arXiv:1802.06955* (2018).
33. Ibtehaz, Nabil, and M. Sohel Rahman. "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation." *Neural networks* 121 (2020): 74-87.
34. Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." In *2016 fourth international conference on 3D vision (3DV)*, pp. 565-571. Ieee, 2016.
35. Baid, Ujjwal, Satyam Ghodasara, Suyash Mohan, Michel Bilello, Evan Calabrese, Errol Colak, Keyvan Farahani et al. "The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification." *arXiv preprint arXiv:2107.02314* (2021).