

## Advances in Brain Tumor Segmentation and Skull Stripping: A 3D Residual Attention U-Net Approach



Tamara A. Dawood<sup>\*</sup>, Ashwaq T. Hashim, Ahmed R. Nasser

Control and System Engineering Department, University of Technology-Iraq, Baghdad 10001, Iraq

Corresponding Author Email: [cse.21.12@grad.uotechnology.edu.iq](mailto:cse.21.12@grad.uotechnology.edu.iq)

<https://doi.org/10.18280/ts.400510>

### ABSTRACT

**Received:** 15 March 2023

**Revised:** 26 June 2023

**Accepted:** 12 July 2023

**Available online:** 30 October 2023

#### Keywords:

*attention, bias field correction, brain tumor, deep learning, skull stripping, residual block, segmentation, U-Net*

The timely diagnosis of brain tumors plays a critical role in enhancing patient prognosis and survival rates. Despite its superior accuracy, manual tumor segmentation is known to be a labor-intensive process. Over the years, a collection of automated tumor segmentation methodologies has been devised and investigated. However, a universally applicable resolution that consistently delivers reliable outcomes across diverse datasets continues to be elusive. Additionally, skull stripping remains a crucial prerequisite to the tumor segmentation procedure. This paper introduces an integrated 3D Attention Residual U-Net (3D\_Att\_Res\_U-Net) model that seamlessly merges attention mechanisms and residual units within the U-Net architecture to augment the performance of brain tumor segmentation and skull stripping in Magnetic Resonance Imaging (MRI). An initial preprocessing stage is implemented, incorporating bias field correction and intensity normalization to optimize performance. The proposed model is trained using the Brain Tumor Segmentation (BraTS) 2020 dataset, along with the Neurofeedback Skull Stripping (NFBS) dataset. The proposed methodology achieved Dice Similarity Coefficients (DSC) of 0.9961 for skull stripping, and 0.9985, 0.9982, and 0.9980 for whole tumor, enhanced tumor, and tumor core segmentation, respectively. Experimental results underscore the applicability and superiority of the proposed approach compared to existing methods in this research domain.

## 1. INTRODUCTION

Brain tumors are characterized by the uncontrolled proliferation of abnormal cells within the brain tissue. They can be broadly classified into benign and malignant types. While benign brain tumors do not impinge upon surrounding healthy tissues, malignant tumors pose a significant health risk due to their invasive nature. Gliomas, the most common form of malignant brain tumors, can be further stratified into high-grade gliomas (HGG) and low-grade gliomas (LGG) [1, 2]. Early detection of brain tumors is critical for enhancing patient survival rates. Manual segmentation, despite being the industry standard, is labor-intensive, costly, and prone to inter-observer variability [3]. Therefore, automatic tumor segmentation, particularly for large datasets, is recommended, given its utility in continuous tumor surveillance and adaptive treatment planning in clinical practice [4].

Magnetic resonance imaging (MRI) is one of the most frequently utilized imaging modalities for brain tumor detection in clinical settings due to its superior soft tissue resolution [5]. Additionally, MRI poses no known health risks. Quantitative analysis of brain tumors can be performed using multimodal brain scans, allowing physicians to devise the most accurate diagnostic and treatment strategies for patients.

However, MRI images are often compromised by field bias, a low-frequency, highly smooth signal, particularly those generated by older MRI machines. Consequently, image processing techniques such as skull stripping, segmentation, or classification, which rely on the gray-level values of image pixels, may not yield accurate results. Therefore, when

applying these methods to distorted MRI images, a preprocessing step is needed to correct the bias field signal [6].

In most brain MRI examinations, skull stripping is usually the first step, especially prior to brain tumor segmentation. However, automatic skull stripping poses a challenge due to lack of intensity uniformity, low contrast MRIs, and indistinct brain boundaries [7]. The task becomes even more complex when dealing with MRI datasets associated with clinical conditions [8]. On T1-weighted images, both the skull and the cerebrospinal fluid (CSF) space appear black, providing clear boundaries between the brain and the skull. However, even these sharp edges can become distorted during MRI acquisition due to low resolution or the presence of other anatomical partial structures in the brain (e.g., connections between the brain and optic nerves or brainstem). Therefore, developing a fully automated method for skull removal from MRI images is crucial prior to tumor segmentation.

The variability in size, shape, and structure of brain tumors, coupled with the effects of surrounding tissues and imaging device noise, presents significant challenges in accurately detecting and segmenting tumors from brain MRI images [9]. Thankfully, advancements in deep learning technology have led to significant progress in automatic image segmentation techniques [10].

Semantic segmentation, a fundamental task in computer vision that involves assigning a semantic label to each pixel in an image, is important for enabling machines to understand and interpret image content more meaningfully. This approach has numerous applications, including self-driving cars, medical imaging, and object recognition [11]. Convolutional

Neural Networks (CNNs), a type of deep learning model, are employed in the semantic segmentation process to analyze an image and predict the class to which each pixel belongs. In the context of brain MRI images, the semantic segmentation model labels each pixel of the entire brain image for skull stripping and brain tumor segmentation.

The U-Net architecture, proposed by Ronneberger et al. [12], forms the basis of most semantic segmentation algorithms for brain tumor cell segmentation. Created for Biomedical Image Segmentation in 2015, U-Net is one of the most widely used methods for semantic segmentation tasks. It is a fully convolutional neural network designed with smaller training sample sizes in mind. By integrating an encoder path for gathering context information and a decoder path for ensuring precise localization, U-Net significantly improves the performance of the medical image segmentation task [13].

Moreover, the residual block technique has been shown to aid in the convergence of the network to the optimal solution, thereby improving performance and reducing training time [14]. A residual block is a construction block of a Convolutional Neural Network (CNN) that aids in addressing the vanishing gradients problem in deep networks.

Additionally, attention mechanisms have been demonstrated to be successful in the field of computer vision for capturing long-range dependencies and significant responses, where the segmentation of medical images has successfully incorporated attention mechanisms [15]. Because U-Net must gradually recover the down-sampled image generated by the inherent pooling and stride convolution, attention approaches can more effectively relate the information flow from deep layers to shallow layers and control the learning of upsampling.

The main contributions of this work are as follows:

- 1) Create a 3D attention residual U-Net (3D\_Att\_ResU-Net) to handle the segmenting of brain tumor and skull stripping tasks that successfully integrates the attention and residual blocks into the U-Net network design.

- 2) Instead of learning the direct mapping between inputs and outputs, the model can quickly learn the residual mapping according to a residual block technique. Over-fitting and vanishing gradients were problems for the U-Net design, especially when dealing with huge and intricate images. To solve these problems, residual blocks were developed, which let the network learn the residual mapping between the inputs and outputs rather than the direct mapping.

- 3) To investigate the implications of local responses for the segmentation of brain tumors, the current U-Net model on the brain attention mechanism is embedded. To increase performance even more, a 3D segmenting network for brain tumors is shown by concurrently incorporating attention mechanisms and residual units into U-Net.

- 4) The network's performance was greatly enhanced by preprocessing processes. The N4 bias field correction is applied to MRI image data to rectify low-frequency intensity non-uniformity. Moreover, the multimodal scans in the BraTS 2020 and NFBS datasets were obtained using various scanners from numerous institutions and a variety of clinical methods, leading to non-standard intensity distribution. Hence, in order to execute multi-mode scan by a single method, the normalizing stage is required.

The rest of the argument is structured as follows: A list of related works is provided in Section 2. The details of the suggested 3D\_Att\_ResU-Net technique are introduced in Section 3. Part 4 describes the implementation of the proposed

method, Section 5 describes the experiments for skull stripping and segmenting brain tumors, and Section 6 provides conclusions.

## 2. RELATED WORK

In this paper, two fields related to the segmentation process are concerned. The first one is skull-stripping and the other tumor localization. Throughout the past two decades, numerous solutions to the issue of skull stripping in brain MRI images are put forth, and new designs are still being created to address these issues and constraints. The vast variation in brain MRI datasets and standards, however, places limitations on every technique. In recent times, deep learning-based methods like CNNs have produced excellent results in the segmentation of biomedical 3D images, with accuracy that is comparable to that of humans. In 2017, Milletari et al. [16] presented CNN with the Hough voting method to effectively integrate the location and segments of an area of interest for 3D deep brain segmentation. In 2016, Kleesiek et al. [17] proposed a method for brain segmentation and skull removing based on 3D CNN and demonstrated to perform well. Their design can accommodate a variety of formats, involving contrast enhanced images. Even though Kleesiek's technique was the first CNN to address the issue and demonstrated precision, the structure is not deep. It should be noted that while the network's depth is not a concern for the current task, it might be if the network were to be used to subsequent tasks. Because subsurface layers cannot combine features from different levels of abstraction, the short depth typically has a restricted capacity for learning. In 2019, The usage of 3D U-Net for the stripping of the skull in brain MRI was suggested by Hwang et al. [7]. They used it on actual, freely obtainable brain MRI database and had successful results.

There have been more articles on brain tumor segmentation throughout the last few decades. As a result, research is being done in this area to develop an automation system for segmenting brain tumors. The three types of brain tumor segmentation techniques are manual, semi-automatic, and fully automated procedures. Deep learning based and classification approaches using neural networks are both fully auto-mated. Researchers have started using CNN to segment biological images in recent years due to the effectiveness of deep learning techniques [18, 19].

Deep networks, which significantly outperform conventional techniques, have recently been developed in the area of the segmentation of brain tumors. Among these, patch-wise based brain tumor segmentation networks are learned on small patches with labels to accurately differentiate tumor tissues from natural tissues as representative studies proposed early. Researchers have created a variety of modules to add more contextual contact information amongst distinct network slices in order to achieve beneficial performance. In 2016, A cascaded Convolutional Neural Network architecture with two paths was proposed by Havaei et al. [20]. One pathway concentrates on the finer details of gliomas, and the other pathway considers the bigger context. additionally, suggest a two-phase patch-wise training method. The (BRATS'13, BRATS'15) data sets were used to evaluate the model. In 2019, Derikvand and Khotanlou [21] used a combination of a used a cascading structure constructed using a convolutional and convolutional neural network in an automated brain tumor segmentation algorithm. Using a patch-based method, the

input images are firstly divided to patches before being fed into the neural network. In order to enhance the segmentation outcomes, it lastly adds a label to the mid voxel of every patch based on both local and global information. the BraTS 2017 database used to validate the model. For the brain tumor segmentation challenge. In 2017, Kamnitsas et al. [22] used a 3D CNN model with two pathways and a dense structure. Additionally, this model used conditional random fields to implement multi-scale analysis on input pictures and post-processing on the resulting images (CRF). Finally, this piece of work won the BraTS 2015 challenge. In 2018, Zhao et al. [23] employed a fully convolutional neural network with CRF to perform the brain tumor segmentation. They trained three 2D patch-wise models from the axial view, sagittal view, and coronal view using a voting-based fusion technique.

The semantic segmentation system describes each pixel of the entire brain image into a set of assigned labels to complete segmenting the brain tumors. In 2015, The U-Net architecture suggested by Ronneberger et al. [12] is the foundation of the majority of the semantic segmentation methods for the segmenting of brain tumors challenge one of the most widely used methods for any semantic segmentation task in use today is the UNET architecture, which was created for Biomedical Picture Segmentation in 2015. It is a fully convolutional neural network built to learn from a smaller number of training data. The performance of the medical image segmentation task is significantly enhanced by U-Net, which has an encoder path to gather context information and a decoder path to assure precise positioning [24]. In order to increase the generalization ability of their 2D U-Net based segmentation of brain tumors network, Dong et al. [25] used real time data augmentation in 2017. By including Two-pathway residual blocks to the u net architecture. In 2021 Aghalari et al. [26] were able to improve the U-Net assessment criteria like DSC, sensitivity, and the number of parameters. which concurrently take advantage of both local and more global qualities. The BRATS'2018 database was used to evaluate the proposed models. Path aggregation U-Net (PAU-Net) model for brain tumor segmentation was proposed by Lin et al. [27] in 2021. The bottom-up path aggregation encoder (PA) specifically reduced the entry of sounds by shortening the distance between output layers and deep features. More intact information was stored using the enhanced decoder (ED). Furthermore, the efficient feature pyramid (EFP), which uses fewer resources to achieve the feature pyramid effect, was applied to enhance mask prediction further. In 2022, Munir et al. [13] suggest a framework for segmenting brain tumor based on the U-Net architecture and Inception modules. Each Inception module in the model contains a number of convolutional filters of various sizes to collect contextual data at various scales. The proposed framework also uses a new losing function based on the improved Dice similarity coefficient (DSC) to enhance the segmentation accuracy

Attention mechanisms are increasingly being used in computer vision tasks for two main purposes. The first one is to emphasize based on distant dependencies. For example, in 2019, Fu et al. [28] introduced dual attention modules consisting of spatial and channel attention for semantic segmentation, where the spatial attention is similar to the non-local (NL) operation in NL-Net and the channel attention follows a similar idea. To further enhance segmentation performance, in 2019, Zhang et al. [29] developed NL using a prior distribution and constructed an ensemble of NLs with

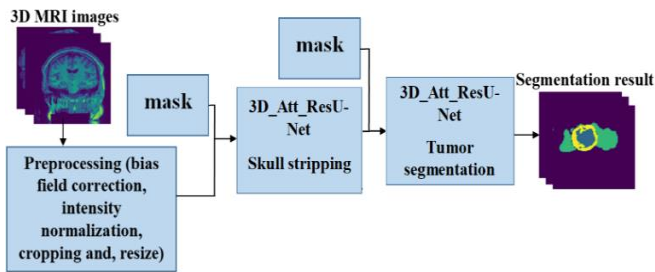
weights in other works. The learning of each channel's scaling factors for feature maps is the second goal of attention processes. This method, which emphasizes the channel relationship and implements dynamic channel-wise feature recalculation to improve feature expression, is typically illustrated by SENet [30]. Researchers are investigating how to employ attention to learn channel-specific variables to selectively improve channel responses based on the success of attention.

Residual blocks are being increasingly utilized in computer vision tasks to combat the problem of vanishing gradients, which can impede the learning process by making the gradients too small during backpropagation. To address this issue, in 2019, Abd-Allah et al. [31] introduced the TPUAR-Net which employed a deep residual based CNN model with two parallel U-Nets for BTS. By adding skip connections and residual blocks, this model was able to extract global and local feature responses at different levels, and the model was effective in terms of execution speed. To take advantage of the contextual information in 3 diminution MRI image, in 2021, Ghaffari et al. [32] introduced a 3D CNN design for segmenting different parts of brain tumors. The decoder path employed the self ensembling strategy to reduce the number of feature maps at each level, while the encoder path used residual blocks to learn non-linear residual and improve the learning process. However, this suggested model requires greater computational and hardware power. In 2023, Raza et al. [33] proposed a deep learning architecture called dResU-Net, which aims to dominate the problem of vanishing gradient in deep neural networks. The suggested architecture combines the deep residual network and the U-Net model, two well-known deep learning models. The suggested architecture uses the deep residual network as an encoder and the U-Net model as a decoder to tackle the vanishing gradient problem.

Based on the previous works, the U-Net architecture has proven superior to other segmentation methods in medical image analysis, making it a popular choice in this field. Plus, it allows for efficient training with relatively small data sets, which is often the case in medical imaging studies. Therefore, in this study, the U-Net architecture was employed with some improvements by merging the attention units and residual units, which can be a powerful combination for identifying brain tumors in medical images. The attention units allow the model to selectively focus on the most relevant features for tumor detection, while the residual units can improve training stability and prevent overfitting. This combination of the model can more accurately localize the tumor and help guide treatment planning. Image preprocessing techniques were employed in this study to improve the model's performance by removing unwanted elements and highlighting important features in the images. Additionally, skull stripping has been given particular importance in this study as it can significantly affect the accuracy of brain tumor diagnosis. By removing the skull from the image, the model can focus better on the brain tissue, which can lead to more accurate tumor localization and identification.

### 3. MATERIAL AND METHOD

Figure 1 provides an overview of the workflow of the proposed method to better explain the procedure.



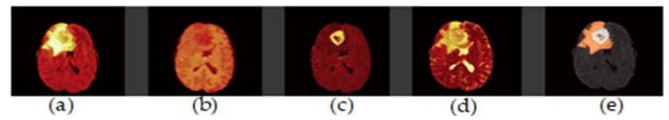
**Figure 1.** The overall architecture of the proposed system

### 3.1 Magnetic resonance imaging

Magnetic Resonance Imaging (MRI) is an imaging technique that provides high contrast and detailed images of the spinal cord, brain, and vascular anatomy. Compared to CT scans, MRI does not involve radiation, making it a safer option. Coronal, sagittal, and axial planes of the brain can all be seen with MRI technology. T1-weighted (T1w), T2-weighted (T2w), and Fluid Attenuated Inversion Recovery (FLAIR) are the three most popular MRI sequences [34].

T1w MRI is mainly utilized to differentiate between healthy and diseased tissues, establishing a strong difference between gray and white matter T2w MRI is particularly suitable for brain disorders where water accumulates in brain tissues, due to its sensitivity to water content as shown in Figure 2. This modality helps to determine the location of edema, resulting in a bright signal on the image. Cerebrospinal fluid (CSF), which is a colorless fluid found in the spinal cord and brain, can be effectively distinguished using T1w and T2w images. In T2w images, the CSF appears bright, whereas in T1w images, it appears dim. T1w MRI with gadolinium contrast enhancement (T1-Gd) is another MRI sequence used for imaging. In this modality, a contrast agent, such as gadolinium ions, is accumulated in the active cell area of tumor tissues to produce a bright signal, facilitating the demarcation of the tumor boundaries. Necrotic cells, which are segmented as a hypo-intense region of the tumor's core and do not correlate with contrast agents, help to segment the hypo-intense region of the active cell zone. Fluid Attenuated Inversion Recovery (FLAIR) is similar to T2-weighted images except for its acquisition protocol. It achieves the suppression of the water molecule, which helps to distinguish between edema and the CSF. FLAIR has the capability to suppress water signals, making hypertensive periventricular lesions easily visible.

Although brain tumors are not as common as liver, esophageal, and breast tumors, they have a significant impact on global mortality rates. Brain cancer is one of the deadliest forms of cancer, affecting both adults and children. As a result, research has focused on developing diagnostic techniques to improve early detection and increase survival rates. Accurate tumor diagnosis helps doctors determine the best treatment options, including chemotherapy and surgery. The World Health Organization has categorized tumors into four types [35] Gliomas and metastases are the two most common types of malignant tumors, accounting for about 80% of cases [36] Gliomas are classified as either low-grade (LGG) or high-grade (HGG) based on their aggressiveness. Magnetic Resonance Imaging (MRI) is a noninvasive diagnostic tool that provides precise information about the tumor and surrounding healthy tissues [2].

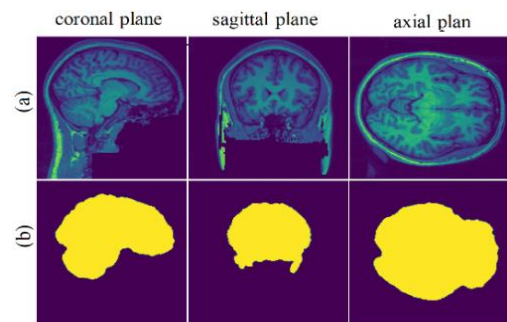


**Figure 2.** Brain tumor MRI modalities. (a) FLAIR, (b) T1-weighted, (c) T1ce respectively, (d) T2-weighted (e) Ground truth

### 3.2 Datasets

#### 3.2.1 Skull stripping dataset

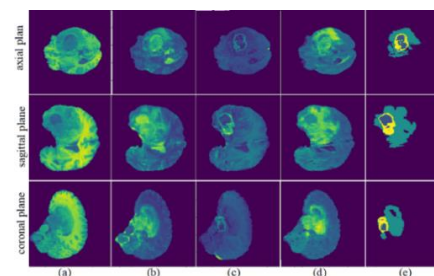
This study's skull stripping MRI dataset was made available through NFBS challenge [37]. The data collection consists of a T1-weighted (T1w) model brain MRI scan for 124 participants. Ground truth labels that are manually segregated for each subject in the training dataset are given. A sample from the NFBS dataset is shown in Figure 3.



**Figure 3.** 3 samples from the NFBS data set. (a) the t1w MRI images (b) the ground truth

#### 3.2.2 Brain tumor segmentation dataset

The MRI dataset utilized in this research is obtained from the BraTS'2020 Challenge [38] which contains multimodal 3D brain MRI scans of 369 individuals in the training dataset. Each subject underwent four scans, including native T1-weighted (T1w), contrast-enhanced T1-weighted (T1ce), T2-weighted (T2w), and T2 Fluid Attenuated Inversion Recovery (FLAIR). Ground truth labels were manually segmented by highly skilled and certified neuroradiologists, and include annotations for the GD-enhancing tumor (ET-label 4), peritumoral edema (ED-label 2), and necrotic and non-enhancing tumor core (NCR/NET-label 1), background (Label 0). Validation datasets of 125 individuals with similar scans but lacking expert segmentation annotations and grading information were also included. Figure 4 displays a sample of the Brats2020 dataset.



**Figure 4.** 3 samples from the Brats2020 data set. (a) T1w (b) T2w (c) T1ce (d) FLAIR (e) the ground truth

### 3.3 Preprocessing

The presence of variations in intensity within the same tissues is known as intensity inhomogeneity, which is mainly caused by RF coil imperfections. This issue can adversely affect the results of segmentation, but several methods are available to address it, such as N4 bias field correction. This popular approach uses multiscale optimization to correct low-frequency intensity non-uniformity in MRI image data. In addition to specifying the "real" pixel, the "MaskImage" can be used to prevent excessive processing [9]. The N4 Bias Correction method [39] can be used to correct the MRI scan's intensity inhomogeneity. The N4 bias field correction algorithm is initially applied to T2w, T1ce, and flair scans. The BraTS 2020 multimodal scans contain an unstandardized intensity distribution because they were obtained utilizing various clinical methods and scanners. Normalization is therefore a crucial step in the processing of multi-mode scanning by a single algorithm. To improve calculation efficiency and preserve as much of the original image data as feasible, the original Brats 2020 images were cropped and resized to 128×128×128 pixels, and the original NFBS images were reduced from 256×256×192 pixels to 128×128×128 pixels by eliminating as many zero backgrounds as possible. Finally, the data is normalized to have a zero mean and unit variance.

### 3.4 Proposed structure

The proposed 3D\_Att\_ResU-Net used for brain tumor segmentation is developed from the U-Net architecture with some modifications.

#### 3.4.1 U-Net architecture

U-Net, proposed by Ronneberger et al. [12], is a well-known method for semantic medical image segmentation, that offers superior efficiency and accuracy [40]. U-Net has been proven to be a successful method for medical image

segmentation, even when dealing with limited amounts of training data [41]. It has a symmetrical "U" shape structure and is comprised of two main components. The first component, called the contracting encoder, is located on the left side of the network and is responsible for extracting global features by employing convolution layers and max pooling. The second component, called the decoder, is located on the right side of the network and is responsible for accurate localization, which leads to significant improvement in the performance of medical image segmentation. The decoder uses up-sampling, concatenation with the corresponding cropped feature map from the encoder, and convolution layers at each step. Figure 5 shows the U-Net diagram. In this paper, we demonstrate that replacing the plain unit with a residual unit and attention mechanism can further enhance the performance of U-Net.

#### 3.4.2 Attention mechanism

"Attention" refers to an intentional action that directs focus toward a specific object or goal, giving it priority over other sensory inputs. This ability to selectively assign importance to different stimuli is sometimes described as "giving need". Concentration reinforces the focus on the object of attention while reducing attention to other stimuli [42].

The attention mechanism, a popular technique in deep learning, is inspired by a human vision which can rapidly focus on objects of interest while filtering out extraneous information. It is widely used in many fields including natural language processing [43] and computer vision [44]. Typically, the attention mechanism is used to improve the performance of the encoder-decoder architecture. This paper introduces a type of attention mechanism called "channel-wise attention" which is applied on the channel-wise level to make the U-Net network focus on key feature regions. The channel-wise attention mechanism uses a gating mechanism to weigh the importance of each channel in the input feature map for the final output. The block diagram of the channel-wise attention mechanism is shown in Figure 6.

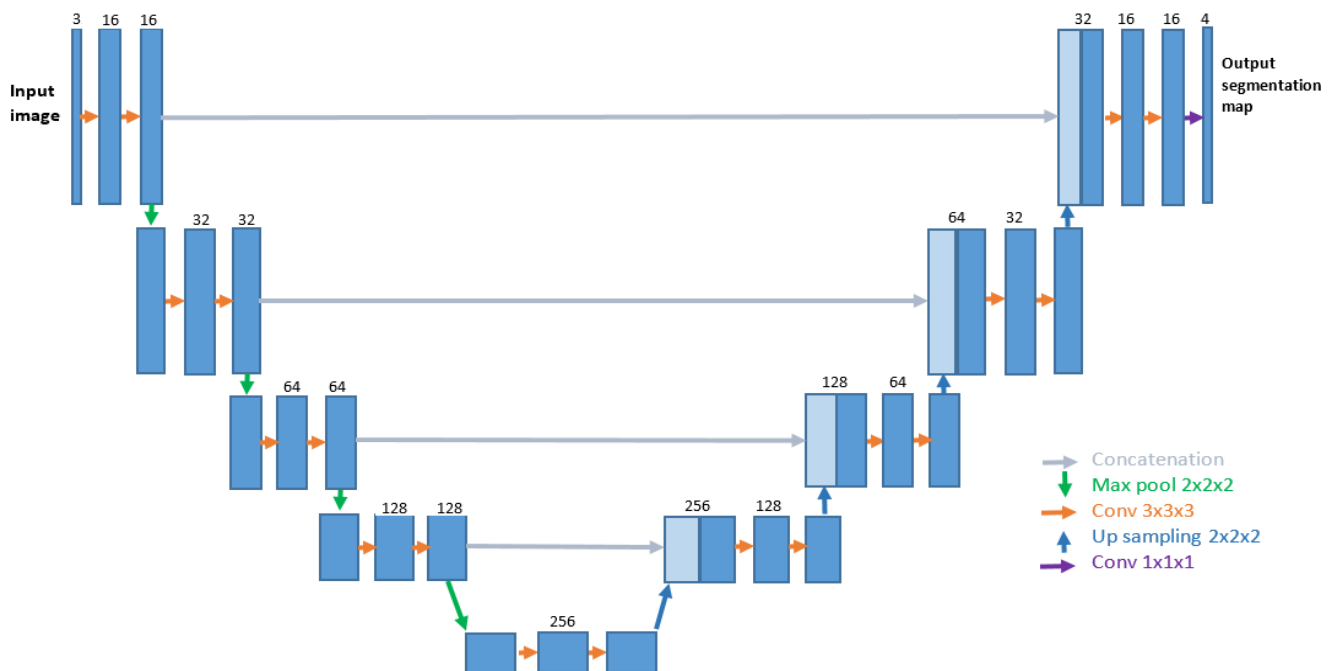


Figure 5. A network diagram of U-Net. The number of channels in each stratum is indicated by numbers above it

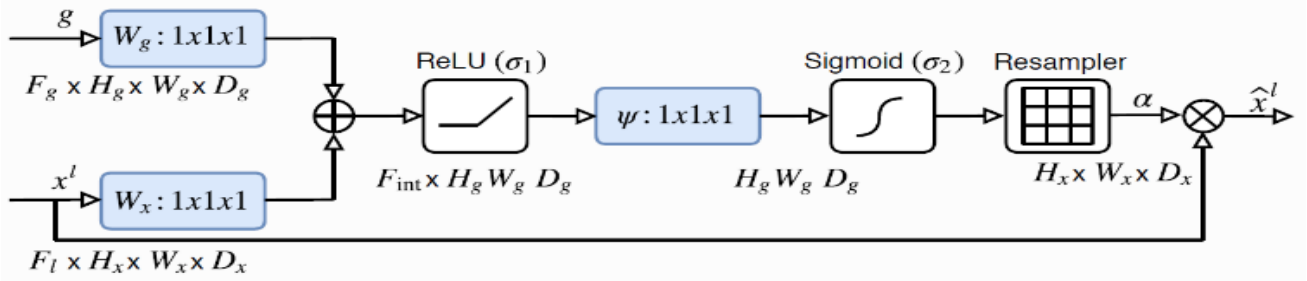


Figure 6. The block diagram of channel-wise attention mechanism

### 3.4.3 Residual unit

The issue of degradation can occur when a multilayer neural network is made deeper, as it could hinder the training process [16]. To address this issue and improve training He et al. [45] suggested the residual neural network, In 2016. This network uses adaptive skip connections in residual blocks to preserve low-level features and overcome the problem of vanishing gradients. The use of identity mapping connections facilitates direct transfer of activations from earlier to later layers, and the concatenation operator joins the output with the input of the residual convolutional block. Moreover, the residual network enhances information transfer and hastens the convergence of the model to the global minima. The residual neural network is composed of a collection of residual elements, each of which can be represented by Eq. (1). and Eq. (2).

$$Y_i = h(x_i) + F(x_i, w_i) \quad (1)$$

$$X_{i+1} = f(y_i) \quad (2)$$

where,  $h(x_i)$  is an identity mapping function, a typical example being  $h(x_i)=x_i$ , and  $x_i$  and  $X_{i+1}$  are the input and output of the  $i$ th residual unit,  $F(\cdot)$  is the residual function,  $f(y_i)$  is the activation function, and  $h(x_i)$  is the mapping function's identity. In Figure 7, a plain unit is distinguished from a residual unit. There are numerous combinations of batch normalization (BN), rectified linear unit activation (ReLU), and convolutional layers in a residual unit. Xu et al. [44] provided a thorough analysis of the effects of various combinations and proposed a complete reactivation scheme, as shown in Figure 7 (b) where (a) represent the basic neural unit of U-Net.

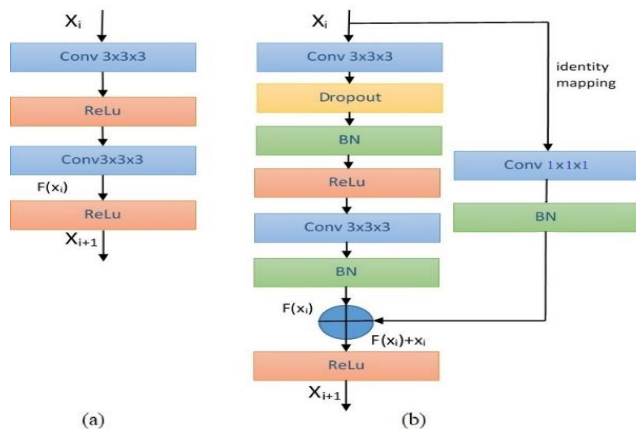


Figure 7. Neural network building blocks. (a) The basic neural unit of U-Net (b) The proposed 3D\_Att\_ResU-Net's residual unit with identity mapping

### 3.4.4 3D\_Att\_ResU-Net

The semantic segmentation neural network is presented. The 3D\_Att\_ResU-Net incorporates the advantages of the UNet, residual neural network, and Attention mechanism. To start, the residual connections are added to the left and right branches' various layers. Next, channel-wise consideration is adding to the skip connection. Three advantages result from this combination: 1) The residual unit will make network training easier. 2) When producing predictions, the UNet network may concentrate on the region of interest in the input image thanks to an attention mechanism, which also lessens the impact of noise. 3) The skip connections within the residual unit between the low levels and high levels of the network will enable information propagation without degradation, allowing the design of a neural network with a great deal fewer parameters while still achieving an ever-improving performance on semantic segmentation. As illustrated in Figure 8, we use a 9-level 3D\_Att\_ResU-Net architecture in this letter to segment brain tumors and strip away the skull. Three components make up the network: encoding, a bridge, and decoding. The incoming image is first encoded into small representations. The final section, or semantic segmentation, recovers the representations to a pixel-by-pixel categorization. The middle section functions as a link between the encoding and decoding paths. The three components are constructed using residual units, which are two  $3 \times 3 \times 3$  convolution blocks and an identity mapping. The identity mapping links the unit's input and output. A BN layer, a ReLU activation layer, a dropout layer, an addition layer, and a convolutional layer are all included in each residual block. The first residual block in the suggested system is shown in Figure 9. For quicker convergence, batch normalization (BN) is applied. The activation function value or convolution output value is normalized to become BN. When BN is used, the weight propagation process is unaffected by a parameter scale. Consequently, the learning rate can have boosted that determines how much alter the weights are altered, allowing for quick learning. A regularization method known as dropout eliminates overfitting [46]. This is because the model is overly tuned and fitted to the training dataset. As a result, it is limited to the training dataset. In order to avoid overfitting during training, the dropout strategy deliberately removes some network units. Dropout enhances performance by essentially building numerous models and making predictions. Overfitting may inevitably result from learning with just one model. The risk of overfitting can be decreased, though, if numerous models are trained and predictions are made using each model. There are four residual units in the encoding path. The maximum pooling used for down sampling in each unit has voxel sizes of  $2 \times 2 \times 2$  and a stride of 2 in each dimension to minimize the feature map in half. The attention is added as the horizontal link in order to obtain richer low-level and high-

level information, strengthening feature information representations between the two types of data. The decoding path also includes four residual units in this manner. Before to each unit, lower level feature maps from the previous level are up sampled, and they are concatenated with the feature maps from the appropriate attention mechanisms path. Figure 10 shows the first attention block between the output of the fourth Res. block and the output of the bridge Res. block and then the

concatenation that linked the output of the attention block and the output of the first up sampling after the bridge Res. block. The multichannel feature maps are projected into the appropriate segmentation using Softmax for brain tumor segmentation after the final level of decoding path, a  $1 \times 1 \times 1$  convolution, and a sigmoid activation function for skull stripping. Table 1 lists the input and output specifications for each phase.

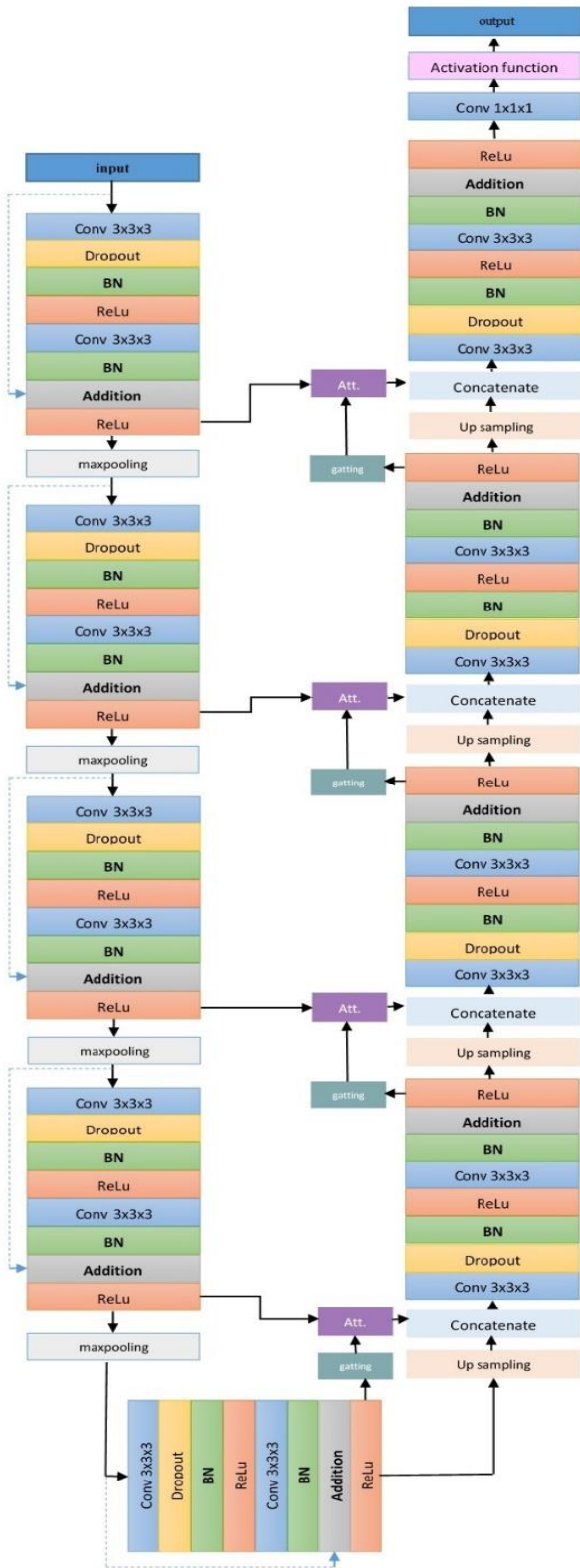


Figure 8. The proposed 3D\_Att\_ResU-Net architecture

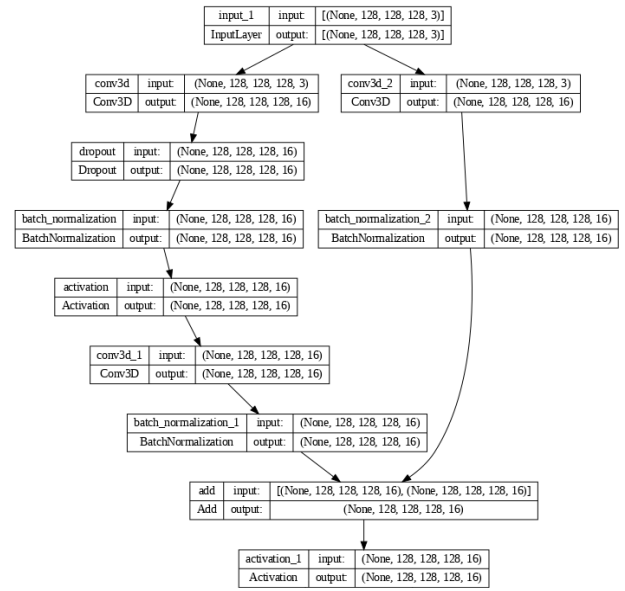


Figure 9. The first residual block in the proposed system

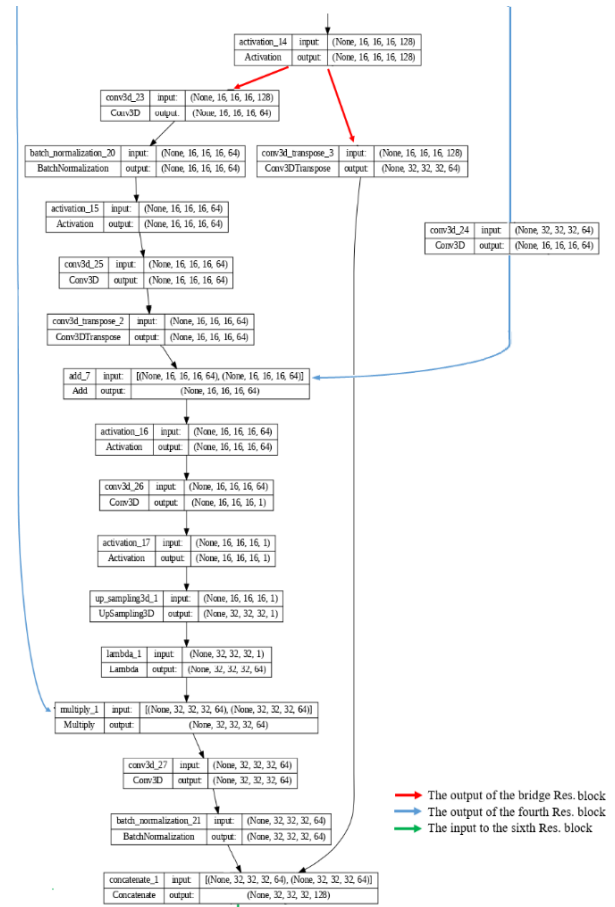


Figure 10. The first attention block

**Table 1.** Network structure of 3D\_Att\_ResU-Net

	Unit Level	Conv Layer	Filter Size	Output Size	No of Parameters	
Input				128,128,128,3	0	
Encoding	Level 1	Res block	16	128,128,128,16	8,496	
		maxpooling		64,64,64,16	0	
	Level 2	Res block	32	64,64,64,32	42,464	
		pooling		32,32,32,32	0	
Bridge	Level 3	Res block	64	32,32,32,64	168,896	
		maxpooling		16,16,16,64	0	
	Level 4	Res block	128	16,16,16,128	673,664	
		maxpooling		8,8,8,128	0	
Decoding	Level 5	Res block	256	8,8,8,256	2,690,816	
		Attention		16,16,16,128	640,769	
	Level 6	Down sampling	—	16,16,16,128	884,864	
		concatenate		16,16,16,256	0	
	Level 7	Res block	128	16,16,16,128	1,361,792	
		Attention		32,32,32,64	160,641	
		Down sampling		—	32,32,32,64	221,248
		concatenate		—	32,32,32,128	0
	Level 8	Res block	64	32,32,32,64	340,928	
		Attention		—	64,64,64,32	40,385
Down sampling		—		64,64,64,32	55,328	
concatenate		—		64,64,64,64	0	
Level 9	Res block	32	64,64,64,32	85,472		
	Attention		—	128,128,128,16	10,209	
	Down sampling		—	128,128,128,16	13,840	
	concatenate		—	128,128,128,32	0	
Output	Conv	Res block	16	128,128,128,16	21,488	
		2	128,128,128,2 skull stripping	68		
			4	128,128,128,4 brain tumor segmentation		
Total parameters: 7,421,368						
Trainable parameters: 7,415,992						
Non-trainable parameters: 5,376						

**3.5 Evaluation metrics**

The intersection over union (IoU) evaluation metric was one of several used in the work to objectively assess the efficacy of the suggested strategy Eq. (3). The IoU (Intersection over Union) is a segmentation performance metric that determines the overlap between the expected and ground-truth segments.

$$IoU = \frac{TP}{TP + FP + FN} \tag{3}$$

Dice similarity coefficient (DSC) Eq. (4), precision Eq. (5), specificity Eq. (6), sensitivity Eq. (7), and accuracy are the other five frequent evaluation metrics employed in image segmentation tasks, particularly in the field of medicine Eq. (8).

$$DSC = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

where,

True Positives (TP) are the proportion of pixels that are correctly identified as belonging to an item.

False Positives (FP) are the quantity of pixels that are mistakenly identified as belonging to an item.

False Negatives (FN) are the number of pixels that should be considered to be a part of the item but are instead not.

True Negatives (TN): the quantity of pixels that are correctly identified as not being a part of the object.

**4. MODEL IMPLEMENTATION**

Python programming, the Keras library, and TensorFlow as the backend were used to create the suggested model (3D\_Att\_ResU-Net). The T1w channel of the MRI images was used to create predicted brain masks using the 3D\_Att\_ResU-Net model. These masks were utilized to extract the images of the brain tissue. The 124 images of T1w MRI that were chosen, with an image size of 256×256×192 were used. The T2, T1ce and Flair channels of the MRI images were used to create predicted brain tumor masks using the 3D\_Att\_ResU-Net model. These masks were utilized to extract the images of the brain tumor tissue. The 369 images of 3 modalities MRI that were chosen, with an image size of 240×240×155×1 were used. preprocessing was performed on them, including bias field correction, intensity normalization, cropping, and resizing, as the images' size becomes 128×128×128×1 for skull stripping and 128×128×128×3 for brain tumor segmentation. Batch normalization were used typically improves model stability whereas layer-by-layer normalizing the network. Due to the limited computing power, a batch size of 2 was used to train the model. 90% of the da-ta were taken for training and 10%



for validation, served as the basis for the studies. The Google Colab Pro Tesla T4 GPU and 25 GB RAM were used for the model training, because Google Colab is an exceptional cloud service that includes a comprehensive Keras library and enables users to interact with a server using a Jupyter notebook environment [47]. To determine the ideal combination of the hyperparameters, a number of tests were run using the suggested method. Starting with smaller filters to gather as much local information as possible, the trials gradually raised the filter width to lower the resulting feature space width and obtain more representative data. The dropout layer is utilized during training. The dropout rate was initially set to 0.3, however by empirically adjusting the dropout value, it was discovered that a 0.2 dropout rate was ideal for the test's outcome. Also, the learning rate was empirically adjusted in a variety of tests to determine the ideal learning rate value. By starting with a higher learning rate and gradually lowering it, this method enables us to reach the global minima much faster. After training, the brain mask and brain tumor mask were predicted for the selected MRI images for each patient using the 3D\_Att\_ResU-Net. Each image's forecast time ranges from 12 to 80 ms. Table 2 provides more information on the hyperparameter settings that were made during model training.

**Table 2.** The 3D\_Att\_ResU-Net's hyperparameter

Hyperparameters	Skull Stripping	Brain Tumor Segmentation
Input size	128×128×128×1	128×128×128×3
Learning rate	0.001	0.001
Batch size	2	2
The hidden layer activation function	ReLU	ReLU
Optimizer	ADAM	ADAM
epochs	25	190
Dropout	0.2	0.2
Output layer activation function	Sigmoid	SoftMax
Output size	128×128×128×2	128×128×128×4

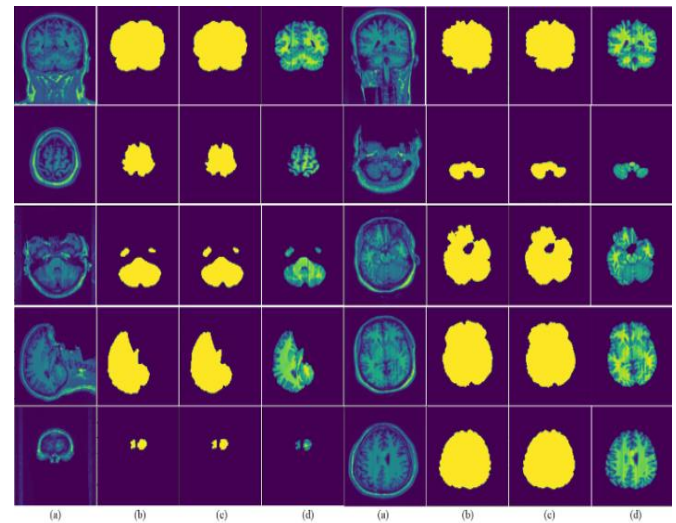
## 5. RESULTS AND DISCUSSION

This section covers the evaluation measures used to assess the proposed model's performance, implementation details, obtained results, and a comparison with state-of-the-art techniques. Evaluation measures such as DSC, Precision, Sensitivity, Specificity, and IOU were employed to evaluate

the model's accuracy and detection capabilities. Implementation details include architecture, hyperparameters, and preprocessing or post-processing techniques. the evaluation process divided into two section skull stripping evaluation and brain tumor segmentation evaluation as shown below.

### 5.1 Skull stripping evaluation

The proposed method of skull stripping is trained and tested on the NFBS data set. The method is implemented once without images preprocessing and once after preprocessing to find out the extent of the effect of image preprocessing on the desired results. The efficacy of the method was tested by measuring IoU, DSC, precision, specificity, Sensitivity and, accuracy. Table 3 shows the comparison of the proposed skull stripping method with and without preprocessing. Table 4 shows a comparison of mean deviation for MVU-Net method, 3D U-net, MHF method and, the proposed 3D\_Att\_ResU-Net method. Figure 11 shows Examples of skull stripping results on the validation set of NBFS dataset. The visual representation demonstrates how closely the outcomes correspond to the values of the ground truth.



**Figure 11.** Examples of skull stripping outcomes on the validation set of NBFS dataset. (a) the input t1w MRI image, (b) the ground truth, (c) the predicted brain mask and, (d) the extracted brain

**Table 3.** Comparison of the proposed skull stripping method with and without preprocessing

Preprocessing	Accuracy	IoU	DSC	Precision	Specificity	Sensitivity
Without	0.9814	0.9196	0.8969	0.9861	0.9865	0.9318
With	<b>0.9942</b>	<b>0.9644</b>	<b>0.9961</b>	<b>0.9916</b>	<b>0.9916</b>	<b>0.9943</b>

**Table 4.** Comparison of mean deviation for MVU-Net method, 3D U-Net and the proposed method. The strongest values are boldened

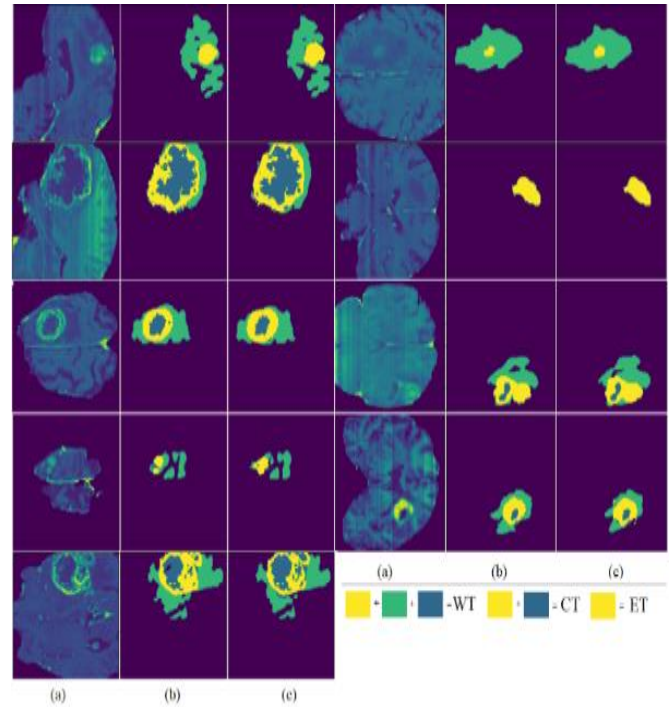
Method	DSC	Sensitivity	Specificity
[7] 3D U-Net	0.9903	0.9853	0.9953
[48] MVU-Net	0.9681	0.9763	<b>0.9954</b>
[49] MHF	0.9416	-	-
Proposed: 3D_Att_ResU-Net	<b>0.9961</b>	<b>0.9943</b>	0.9916

## 5.2 Brain tumor segmentation evaluation

The proposed method of brain tumor segmentation was trained and tested on the BraTS 2020 database. The method was implemented once without images preprocessing and once after preprocessing to find out the extent of the effect of image preprocessing on the desired results. The efficacy of the method was tested by measuring IoU, DSC, precision, specificity, Sensitivity and, accuracy. Table 5 shows the comparison of the proposed brain tumor segmentation method with and without preprocessing. Table 6 shows a comparison for related works and the proposed 3D\_Att\_ResU-Net method. Table 6 shows a Comparison of Accuracy, IoU, Mean DSC, precision, specificity and, Sensitivity of proposed 3D\_Att\_ResU-Net model with the most recent techniques. Figure 12 shows Examples of brain tumor segmentation results on the validation set of BraTS dataset. The visual representation demonstrates how closely the results match the ground truth values for WT, CT, and ET.

The 3D\_Att\_ResU-Net model is evaluated against existing models for brain tumor semantic segmentation. The results show that the recommended strategy outperformed existing methods in terms of segmentation results for certain tumor locations (WT, TC, and ET). However, the most challenging regions to segment were the enhancing tumor and its dispersion with necrosis, which many existing models struggled with. On the other hand, the 3D Att ResU-Net model was able to segment these areas. The 3D Att ResU-Net model outperformed cutting-edge techniques for tumor core and enhancing tumor classes, as shown in Table 7. The suggested strategy outperformed previous approaches in terms of complete tumor segmentation, producing results that were

superior. Overall, it seems that the 3D Att ResU-Net model is a better method for producing segmented images that are more accurate.



**Figure 12.** Examples of brain tumor segmentation results on the validation set of BraTS 2020 dataset. (a) the input MRI image, (b) the ground truth, (c) the model prediction

**Table 5.** Comparison of the proposed brain tumor segmentation method with and without a preprocessing

Preprocessing	Accuracy	Mean IoU	DSC	Precision	Specificity	Sensitivity
Without	0.9490	0.8744	0.9381	0.9501	0.9844	0.9590
With	<b>0.9793</b>	<b>0.9595</b>	<b>0.9983</b>	<b>0.9691</b>	<b>0.9897</b>	<b>0.9691</b>

**Table 6.** Comparison for related works and the proposed method. The best values are emboldened

Method	Mean DSC	Precision	Specificity	Sensitivity
[23] FCNN	0.8267	0.8167	-	0.9033
[5] SK-TPCNN	0.8540	0.8167	-	0.9033
[50] OM-Net	0.8878	-	0.9942	0.9012
[51] Modified UNet	0.8187	-	0.995	0.843
[26] TPRE-UNet	0.894	0.9088	<b>0.9987</b>	87.77
TPRD-UNet	0.8954	0.9049	<b>0.9987</b>	88.31
TPRED-UNet	0.8976	0.9065	<b>0.9987</b>	0.8919
[13] Depth-wise separable Hybrid model	0.8775	-	0.9942	0.9026
Proposed 3D_Att_ResU-Net	<b>0.9983</b>	<b>0.9691</b>	0.9897	<b>0.9691</b>

**Table 7.** Comparison of DSC TC, WT and ET of proposed 3D\_Att\_ResU-Net model with the present techniques Brain Tumor Segmentation methods

Method	Dataset	Dice Score (DSC)		
		Tumor Core (TC)	Whole Tumor (WT)	Enhancing Tumor (ET)
[51] Modified U-Net	Brats 2018	0.805	0.868	0.783
[52] 3D-CNNs	Brats 2020	0.7526	0.8463	0.6215
[32] cascaded 3D densely-connected U-Net	Brats 2020	0.82	0.90	0.78
[53] Transformer	Brats 2020	0.8173	0.9009	0.7873
[25] 3D U-Net	Brats 2015	0.86	0.86	0.65
[33] dResU-Net	Brats 2020	0.8357	0.8660	0.8004
[54] 3D Attention U-Net	Brats2019	0.7927	0.898	0.7047
Proposed 3D_Att_ResU-Net	Brats 2020	<b>0.9985</b>	<b>0.9982</b>	<b>0.9980</b>

The preprocessing steps applied in the proposed system resulted in notable performance enhancements, as depicted in Table 3 and Table 5. The improvements were observed across various evaluation metrics. Specifically, the preprocessing steps led to increased accuracy, IOU (Intersection over Union), DSC (Dice Similarity Coefficient), precision, specificity, and sensitivity, with improvement rates ranging between 0.51% and 9.92%. These enhancements highlight the effectiveness of the preprocessing steps in improving the overall performance of the system.

The proposed 3D\_Att\_ResU-Net method for brain tumor segmentation and skull stripping has demonstrated superior performance compared to previous approaches, as indicated in Table 4 and Table 6. It has outperformed previous works in terms of mean DSC, Precision, and Sensitivity. However, the previous work does outperform the proposed system in terms of Specificity, by a relatively small margin ranging between 0.35% and 0.9%.

When examining medical images to detect specific diseases like cancer, the measure of Sensitivity becomes more crucial than Specificity. Sensitivity represents the ability to accurately identify positive cases, meaning the capability to detect true positive cases correctly. In this context, it is more important to achieve a higher sensitivity to ensure the correct identification of as many true positive cases as possible. While high specificity is desirable to avoid false positives, it may be prioritized slightly lower compared to sensitivity in the context of disease detection from medical images.

## 6. ABLATION STUDY

The ablation study compared the performance of the proposed 3D\_Att\_Res\_U-Net model with U-Net, Residual U-Net, and Attention U-Net. The integration of attention and residual units in the proposed model resulted in improved feature representation, better preservation of low-level details, and enhanced training stability, leading to higher segmentation accuracy in brain tumor segmentation. By comparing the proposed method with the ablation studies referenced as [24, 32, 53] in Table 7, it becomes evident that the performance of the proposed method is significantly superior to the respective ablated models, particularly for the challenging ET class, which historically posed difficulties in accurate segmentation. The inclusion of attention and residual mechanisms proved crucial in achieving accurate segmentation, validating their importance in skull stripping and brain tumor identification. In summary, the 3D\_Att\_Res\_U-Net model outperformed the other models, highlighting its effectiveness in addressing the complexities of brain tumor segmentation.

## 7. CONCLUSION

This paper introduces the 3D\_Att\_ResU-Net model for MRI skull stripping and brain tumor segmentation tasks. The superior performance of the architecture is achieved due to three main things. First reason is that the segmentation performance of brain tumors is improved by the model's use of residual units and attention mechanisms. The inclusion of attention units during the down-sampling and up-sampling processes allows for adaptive feature rescaling, increasing the local responses of the residual down-sampling features and the up-sampling process' recovery effects. The second reason

residual blocks were developed, which let the network learn the residual mapping between the inputs and outputs rather than the direct mapping between inputs and outputs, the model can quickly learn the residual mapping according to a residual block technique. The third reason a different preprocessing technique, i.e., bias field correction and intensity normalization was employed. To confirm the confidence of bias field correction (N4ITK) as a preprocessing technique, normalization was employed after bias field correction. Based on the results of bias field correction as a preprocessing method, it was concluded that segmentation results may be improved. As a result, bias field correction should be used as a preprocessing method because it enhances segmentation outcomes. Experimental results show that the proposed method outperforms state-of-the-art models. For skull stripping, the model achieved a dice score of 0.9961, and for brain tumor segmentation, the model achieved dice scores of 0.9985, 0.9982, and 0.9980, for TC, WT, and ET respectively, without any data augmentation or extensive post-processing. In the future, we intend to apply the 3D\_Att\_ResU-Net model for other diseases that require the properties provided by this method.

## REFERENCES

- [1] Thillaikkarasi, R., Saravanan, S. (2019). An enhancement of deep learning algorithm for brain tumor segmentation using kernel based CNN with M-SVM. *Journal of Medical Systems*, 43: 1-7. <https://doi.org/10.1007/s10916-019-1223-7>
- [2] Maddalena, L., Granata, I., Manipur, I., Manzo, M., Guarracino, M.R. (2020). Glioma grade classification via omics imaging. In *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020)*. In *Bioimaging*, pp. 82-92. <https://doi.org/10.5220/0009167700820092>
- [3] Naser, M.A., Deen, M.J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Computers in Biology and Medicine*, 121: 103758. <https://doi.org/10.1016/j.compbimed.2020.103758>
- [4] Cui, S., Mao, L., Jiang, J., Liu, C., Xiong, S. (2018). Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. *Journal of Healthcare Engineering*. <https://doi.org/10.1155/2018/4940593>
- [5] Yang, T., Song, J., Li, L. (2019). A deep learning model integrating SK-TPCNN and random forests for brain tumor segmentation in MRI. *Biocybernetics and Biomedical Engineering*, 39(3): 613-623. <https://doi.org/10.1016/j.bbe.2019.06.003>
- [6] Juntu, J., Sijbers, J., Van Dyck, D., Gielen, J. (2005). Bias field correction for MRI images. In *Computer Recognition Systems: Proceedings of the 4th International Conference on Computer Recognition Systems CORES'05*. Springer Berlin Heidelberg, pp. 543-551. [https://doi.org/10.1007/3-540-32390-2\\_64](https://doi.org/10.1007/3-540-32390-2_64)
- [7] Hwang, H., Rehman, H.Z.U., Lee, S. (2019). 3D U-Net for skull stripping in brain MRI. *Applied Sciences*, 9(3): 569. <https://doi.org/10.3390/app9030569>
- [8] Soltani, M., Bonakdar, A., Shakourifar, N., Babaei, R., Raahemifar, K. (2021). Efficacy of location-based features for survival prediction of patients with

- glioblastoma depending on resection status. *Frontiers in Oncology*, 11: 661123. <https://doi.org/10.3389/fonc.2021.661123>
- [9] Salman, L.A., Hashim, A.T., Hasan, A.M. (2022). Automated brain tumor detection of MRI image based on hybrid image processing techniques. *Telkomnika Telecommunication Computing Electronics and Control*, 20(4): 762-771. <http://doi.org/10.12928/telkomnika.v20i4.22760>
- [10] Dawood, A.S., Faris, Z.M. (2023). Architecture of deep learning and its applications. *Architecture of Deep Learning and Its Applications*, 23(1): 3556. <https://doi.org/10.33103/uot.ijccce.23.1.4>
- [11] Abdullah, B.W., Ahmed, H.M. (2022). A convolutional neural network for detecting covid-19 from chest x-ray images. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, 22(3). <https://doi.org/10.33103/uot.ijccce.22.3.1>
- [12] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III*. Springer International Publishing, 18: 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [13] Munir, K., Frezza, F., Rizzi, A. (2022). Deep learning hybrid techniques for brain tumor segmentation. *Sensors*, 22(21): 8201. <https://doi.org/10.3390/s22218201>
- [14] Zhang, Z., Liu, Q., Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5): 749-753. <https://doi.org/10.1109/LGRS.2018.2802944>
- [15] Zhou, C., Chen, S., Ding, C., Tao, D. (2019). Learning contextual and attentive information for brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II*. Springer International Publishing, 4: 497-507. [https://doi.org/10.1007/978-3-030-11726-9\\_44](https://doi.org/10.1007/978-3-030-11726-9_44)
- [16] Milletari, F., Ahmadi, S.A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., Navab, N. (2017). Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Computer Vision and Image Understanding*, 164: 92-102. <https://doi.org/10.1016/j.cviu.2017.04.002>
- [17] Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., Biller, A. (2016). Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. *NeuroImage*, 129: 460-469. <https://doi.org/10.1016/j.neuroimage.2016.01.024>
- [18] Iqbal, S., Ghani, M.U., Saba, T., Rehman, A. (2018). Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN). *Microscopy Research and Technique*, 81(4): 419-427. <https://doi.org/10.1002/jemt.22994>
- [19] Ahmed, S., Ameen, S.H. (2021). Detection and classification of leaf disease using deep learning for a greenhouses' robot. *Iraqi Journal of Computers, Communications, Control and Systems Engineering*, 21(4): 15-28. <https://doi.org/10.33103/uot.ijccce.21.4.2>
- [20] Havaei, M., Dutil, F., Pal, C., Larochelle, H., Jodoin, P.M. (2016). A convolutional neural network approach to brain tumor segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: First International Workshop, BrainLes 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 5, 2015, Revised Selected Papers*. Springer International Publishing, 1: 195-208. [https://doi.org/10.1007/978-3-319-30858-6\\_17](https://doi.org/10.1007/978-3-319-30858-6_17)
- [21] Derikvand, F., Khotanlou, H. (2019). Patch and pixel based brain tumor segmentation in MRI images using convolutional neural networks. In *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. IEEE, pp. 1-5. <https://doi.org/10.1109/ICSPIS48872.2019.9066097>
- [22] Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36: 61-78. <https://doi.org/10.1016/j.media.2016.10.004>
- [23] Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., Fan, Y. (2018). A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Medical Image Analysis*, 43: 98-111. <https://doi.org/10.1016/j.media.2017.10.002>
- [24] Muthana, M., Nasser, A.R. (2022). Using dynamic pruning technique for efficient depth estimation for autonomous vehicles. *Mathematical Modelling of Engineering Problems*, 9(2): 451-457. <https://doi.org/10.18280/mmep.090221>
- [25] Dong, H., Yang, G., Liu, F., Mo, Y., Guo, Y. (2017). Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In *Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11-13, 2017, Springer International Publishing. Proceedings*, 21: 506-517. [https://doi.org/10.1007/978-3-319-60964-5\\_44](https://doi.org/10.1007/978-3-319-60964-5_44)
- [26] Aghalari, M., Aghagolzadeh, A., Ezoji, M. (2021). Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks. *Biomedical Signal Processing and Control*, 69: 102841. <https://doi.org/10.1016/j.bspc.2021.102841>
- [27] Lin, F., Wu, Q., Liu, J., Wang, D., Kong, X. (2021). Path aggregation U-Net model for brain tumor segmentation. *Multimedia Tools and Applications*, 80: 22951-22964. <https://doi.org/10.1007/s11042-020-08795-9>
- [28] Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., Lu, H. (2019). Dual attention network for scene segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3146-3154.
- [29] Zhang, H., Zhang, H., Wang, C., Xie, J. (2019). Co-occurrent features in semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 548-557. <https://doi.org/10.1109/CVPR.2019.00064>
- [30] Hu, J., Shen, L., Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [31] Abd-Allah, M.K., Khalaf, A.A., Awad, A.I., Hamed, H.F. (2019). TPUAR-Net: Two parallel U-Net with asymmetric residual-based deep convolutional neural

- network for brain tumor segmentation. In *Image Analysis and Recognition: 16th International Conference, ICIAR 2019*, Waterloo, ON, Canada, August 27-29, 2019, Proceedings, Part II. Springer International Publishing, 16: 106-116. [https://doi.org/10.1007/978-3-030-27272-2\\_9](https://doi.org/10.1007/978-3-030-27272-2_9)
- [32] Ghaffari, M., Sowmya, A., Oliver, R. (2021). Automated brain tumour segmentation using cascaded 3d densely-connected U-Net. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part I*. Springer International Publishing, 6: 481-491. [https://doi.org/10.1007/978-3-030-72084-1\\_43](https://doi.org/10.1007/978-3-030-72084-1_43)
- [33] Raza, R., Bajwa, U.I., Mehmood, Y., Anwar, M.W., Jamal, M.H. (2023). dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI. *Biomedical Signal Processing and Control*, 79: 103861. <https://doi.org/10.1016/j.bspc.2022.103861>
- [34] Dong, Q., Welsh, R.C., Chenevert, T.L., Carlos, R.C., Maly-Sundgren, P., Gomez-Hassan, D.M., Mukherji, S.K. (2004). Clinical applications of diffusion tensor imaging. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 19(1): 6-18. <https://doi.org/10.1002/jmri.10424>
- [35] Pereira, S., Meier, R., Alves, V., Reyes, M., Silva, C.A. (2018). Automatic brain tumor grading from MRI data using convolutional neural networks and quality assessment. In *Understanding and Interpreting Machine Learning in Medical Image Computing Applications: First International Workshops, MLCN 2018, DLF 2018, and iMIMIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018*. Springer International Publishing. Proceedings, 1: 106-114. [https://doi.org/10.1007/978-3-030-02628-8\\_12](https://doi.org/10.1007/978-3-030-02628-8_12)
- [36] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6): 394-424. <https://doi.org/10.3322/caac.21492>
- [37] Preprocessed Connectomes Project. NFBS skull-stripped repository. [http://preprocessed-connectomes-project.org/NFB\\_skullstripped/](http://preprocessed-connectomes-project.org/NFB_skullstripped/), accessed on Mar. 09, 2023.
- [38] CBICA. (2020). Multimodal brain tumor segmentation challenge 2020: Data. Perelman school of medicine at the university of pennsylvania. <https://www.med.upenn.edu/cbica/brats2020/data.html>, accessed on Mar. 09, 2023.
- [39] Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6): 1310-1320. <https://doi.org/10.1109/TMI.2010.2046908>
- [40] Ravishankar, K., Devaraj, P., Yeliyur Hanumathaiyah, S.K. (2023). Floor Segmentation Approach Using FCM and CNN. *Acadlore Transactions on AI and Machine Learning*, 2(1): 33-45. <https://doi.org/10.56578/ataiml020104>
- [41] Hssayeni, M.D., Croock, M.S., Salman, A.D., Al-khafaji, H.F., Yahya, Z.A., Ghoraani, B. (2020). Intracranial hemorrhage segmentation using a deep convolutional model. *Data*, 5(1): 14. <https://doi.org/10.3390/data5010014>
- [42] Seddik, S., Routaib, H., Elhaddadi, A. (2023). Multi-variable time series decoding with long short-term memory and mixture attention. *Acadlore Transactions on AI and Machine Learning*, 2(3): 154-169. <https://doi.org/10.56578/ataiml020304>
- [43] Ji, Z., Li, S., Pang, Y. (2018). Fusion-attention network for person search with free-form natural language. *Pattern Recognition Letters*, 116: 205-211. <https://doi.org/10.1016/j.patrec.2018.10.020>
- [44] Xu, J., Lu, K., Wang, H. (2021). Attention fusion network for multi-spectral semantic segmentation. *Pattern Recognition Letters*, 146: 179-184. <https://doi.org/10.1016/j.patrec.2021.03.015>
- [45] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [46] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1): 1929-1958.
- [47] Alwawi, B.K.O.C., Abood, L.H. (2021). Convolution neural network and histogram equalization for COVID-19 diagnosis system. *Indonesian Journal of Electrical Engineering and Computer Science*, 420-427.
- [48] Fatima, A., Madni, T.M., Anwar, F., Janjua, U.I., Sultana, N. (2022). Automated 2D slice-based skull stripping multi-view ensemble model on NFBS and IBSR datasets. *Journal of Digital Imaging*, 35(2): 374-384. <https://doi.org/10.1007/s10278-021-00560-0>
- [49] Paredes-Orta, C., Mendiola-Santibañez, J.D., Ibrahim, D., Rodríguez-Reséndiz, J., Díaz-Florez, G., Olvera-Olvera, C.A. (2022). Hyperconnected openings codified in a max tree structure: An application for skull-stripping in brain MRI T1. *Sensors*, 22(4): 1378. <https://doi.org/10.3390/s22041378>
- [50] Zhou, C., Ding, C., Wang, X., Lu, Z., Tao, D. (2020). One-pass multi-task networks with cross-task guided attention for brain tumor segmentation. *IEEE Transactions on Image Processing*, 29: 4516-4529. <https://doi.org/10.1109/TIP.2020.2973510>
- [51] Kermi, A., Mahmoudi, I., Khadir, M.T. (2019). Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II*. Springer International Publishing, 4: 37-48. [https://doi.org/10.1007/978-3-030-11726-9\\_4](https://doi.org/10.1007/978-3-030-11726-9_4)
- [52] Ballestar, L.M., Vilaplana, V. (2020). Brain tumor segmentation using 3D-CNNs with uncertainty estimation. *arXiv Preprint arXiv: 2009.12188*. <https://doi.org/10.48550/arXiv.2009.12188>
- [53] Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., Li, J. (2021). Transbts: Multimodal brain tumor segmentation using transformer. In *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part I*. Springer

International Publishing, 24: 109-119.  
[https://doi.org/10.1007/978-3-030-87193-2\\_11](https://doi.org/10.1007/978-3-030-87193-2_11)

[54] Islam, M., Vibashan, V.S., Jose, V.J.M., Wijethilake, N., Utkarsh, U., Ren, H. (2020). Brain tumor segmentation and survival prediction using 3D attention UNet. In Brainlesion: Glioma, Multiple Sclerosis, Stroke and

Traumatic Brain Injuries: 5th International Workshop, BrainLes 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Revised Selected Papers, Part I. Springer International Publishing, 5: 262-272. [https://doi.org/10.1007/978-3-030-46640-4\\_25](https://doi.org/10.1007/978-3-030-46640-4_25)