

Original Article

UNet and Transformer-Based Model for Multi-Modality Brain Tumor Segmentation

Jayashree Shedbalkar¹, K. Prabhushetty², R. H. Havaladar³

¹Department of Computer Science and Engineering, KLS VEDIT, Haliyal, VTU Belagavi, Karnataka, India.

²Department of Electronics Communication and Engineering, SGBIT Belagavi, VTU Belagavi, Karnataka, India.

³Department of Biomedical Engineering, KLE DR.M S Sheshgiri College of Engineering and Technology, Belagavi, Karnataka, India.

¹Corresponding Author : jayashri2512@gmail.com

Received: 04 June 2023

Revised: 09 July 2023

Accepted: 04 August 2023

Published: 31 August 2023

Abstract - Currently, the human race is facing several health-related issues where brain tumours are recognized as one of the leading causes of morbidity and mortality worldwide. Several researches and surveys have reported that the timely detection and prediction of brain tumours can help prevent their diverse impacts. Therefore, segmentation of these brain tumors is one of the prime tasks to proliferate the accuracy of diagnosis. Deep learning has become a promising technique to facilitate an automated Brain Tumor Segmentation (BTS) approach. Most current deep learning approaches rely on Convolutional Neural Networks (CNNs), which fail to contain long-term dependencies and global context information. Moreover, the performance of these systems is affected due to receptive field limitations during convolution operations. Currently, UNet-based architectures are adopted to perform medical image segmentation. Thus, in this work, we considered UNet as the base model for segmentation and incorporated transformer-based modules to improvise the segmentation accuracy; along with this, we present a hybrid attention mechanism that uses local and global context information. Based on this architecture, we evaluated the efficiency of the proposed approach for various Brats Datasets (2015, 2017, 2019, 2020 and 2021). The proposed approach achieves the average dice score of 0.94, 0.921, 0.83, and 0.94 for Brat's dataset.

Keywords - Unet architecture, Brain tumor segmentation, Transformer based module, Deep learning model, Swin transform.

1. Introduction

Significant technological advancements have recently been made in biomedical research and human intelligence, allowing us to overcome several diseases. Nevertheless, despite these impressive feats, cancer remains a formidable and persistent challenge for humanity due to its unpredictable nature.

The brain, the principal and intricate organ of the human body, comprises nerve cells and tissues responsible for controlling crucial functions such as breathing, muscle movement, and sensory perception. While each cell possesses unique abilities, some may experience changes that cause them to lose their normal functions, resist regulation, and grow abnormally. These accumulations of aberrant cells form a tissue mass known as a tumour.

A brain tumor is characterized by the abnormal and uncontrolled proliferation of cells within the brain. Because of the limited volume and rigidity of the human skull, such growth can impact various brain functions depending on the tumour's location. Additionally, the tumour can spread to

other body portions and further disrupt the body's functions. According to a study in [1], roughly 22000 people were diagnosed with brain tumours in the USA. Similarly, another study presented in [2] reported prevalence and severity.

Generally, brain tumors are categorized into two main categories, i.e., malignant and benign tumors [3]. A benign tumor in the brain is a non-cancerous growth that develops slowly and does not progress. It is regarded as less aggressive because it stays confined to a specific location and does not metastasize to other body organs [4, 5].

The cell growth abnormality of a benign tumour can exert pressure on the surrounding tissue or brain area and can be surgically removed. In contrast, a malignant tumour is cancerous and spreads quickly with indistinct boundaries, infiltrating healthy cells and spreading to different body areas. Suppose a malignant tumour originates within the brain. In that case, it is called a primary malignant tumour. In contrast, if it develops in different body regions and spreads to the brain, it is considered a secondary malignant tumour [4, 5]. Similarly, The brain tumor types are meningioma,



glioma and pituitary tumors [6, 7]. According to the study presented in [8], approximately 81000 fresh cases of primary brain tumours were reported in 2018, where Meningioma includes 29,300 of these cases, while Gliomas accounted for 21,500 cases. Pituitary tumours comprised almost 13,30.

Meningiomas are the most common benign tumour that develops on the thin membranes surrounding the brain and spinal cord. Gliomas are a group of tumors that develop within the brain tissue, and high-grade gliomas are particularly aggressive, with an average survival time of only two years. In this case, pituitary type of brain tumours that can develop in the pituitary gland have an identical shape and inherent characteristics that can occur anywhere in the brain. Figure 1 depicts the sample images of these tumours.

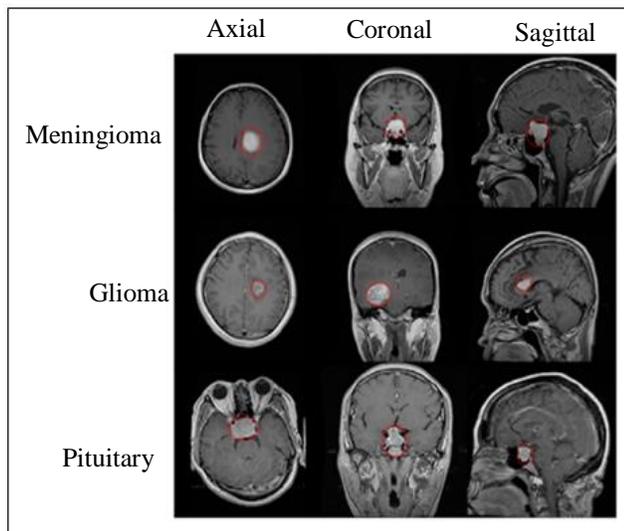


Fig. 1 Sample images of different types of brain tumor

Therefore, early diagnosis is considered the most critical aspect to diminish the impact of brain tumor. However, it can be done by visual inspection or by analysing the image data. The manual inspection fails to provide reliable and accurate results; therefore, computerized automated processes are adopted to increase the recognition performance of brain tumours [9]. These computerized automated processes extract significant clinical information concerning the presence of a tumour, its location, and its type.

Further, this information may guide and control future interventions, improving the diagnosis quality and tumour treatment. It can be achieved by automated computer-aided design-based systems, which have become popular. These systems operate based on the biomedical image analysis approach where different types of images are obtained from brain imaging technique which includes PET [10], SPECT [11], CT [12], MRI [13], and MRS [14]. MRI is the most used technique to extract detailed information about the anatomy of human tissues currently. This imaging modality utilizes powerful magnetic fields and radio frequency signals

to generate images of the tissues. Several kinds of research have been conducted based on brain MRI image analysis.

Diagnosing brain tumors involves several vital processes: tumor detection, segmentation, and classification. Detecting brain tumors involves identifying them in MRI images, which is straightforward. However, segmentation techniques are necessary to pinpoint and isolate the tumor tissues within these images[15]. Different classifications determine whether the abnormal images indicate malignant or benign tumours.

All three approaches - detection, segmentation, and classification - provide valuable information to radiologists, helping them better understand the MRI data and ultimately make an accurate diagnosis. In this process, the BTS plays a vital role; therefore, researchers have focused on developing novel segmentation methods. Moreover, manual image analysis is a complex and time-consuming process requiring professional neuroradiologists' intervention. Therefore, automated BTS methods are needed to overcome these issues.

In the case of medical imaging analysis, accurately segmenting a tumour and separating it from its surrounding Normal Adjacent Tissue (NAT) is very challenging due to various factors such as differences in size, location, shape, and undesirable artefacts caused by improper image acquisition. Researchers presented different models to overcome these challenges and find precise boundary curves for brain tumours in medical images. These models are as follows:

Machine learning methods: These segmentation methods use hand-crafted feature extraction mechanisms to extract the features for segmentation. Extracted features are then applied to train the discriminative model to recognize tumor or normal tissues. These methods are adopted in various research such as Amin et al. [16] adopted a statistical approach which uses wavelet band-based denoising method, potential field clustering, local Binary pattern, and Gabor wavelet features to train SVM, KNN, DT, Naïve Bayes, and Random Forest classifiers.

Jena et al. [17] used hybrid texture feature extraction methods to train supervised classifiers such as SVMs, KNN, BDTs, and RF. However, these methods require extracting the edge feature and other fine-grained information, increasing time consumption. Moreover, the boundary between healthy and tumour tissue is vague, which is not discriminated correctly, resulting in poor performance. Multi-Atlas registration (MAS) approaches are also based on image registration and the fusion of typical brain images to produce new modalities. However, registering typical brain images is difficult in these methods and requires many images. Therefore, these MAS methods fail to process the

images quickly and suffer from computational complexity, time, and accuracy.

Currently, automated methods based on deep learning have gained massive attention in various fields of biomedical applications, such as breast cancer [18], diabetic retinopathy [19], pancreatic tumours [20], and brain tumours etc. In the deep learning domain, Convolutional Neural Networks (CNN) are very successful in many visual tasks, such as image classification, segmentation, and detection of specific objects[21]. Fully Convolutional Networks (FCN) take this a step further by allowing us to label each part of an image with a meaningful category, all at once and with impressive accuracy [22].

U-Net [23] is a popular method used for segmenting medical images, which uses a unique structure that connects an encoder to a decoder in a symmetrical way and uses skip connections to preserve fine details better. This architecture has become a standard for medical image segmentation. Other variants of U-Net, like U-Net++ [24] and Res-UNet [25], have also been developed and have shown even better performance in image segmentation. Although CNN-based methods have proven to be very powerful in representing visual information, they have limitations when capturing long-range dependencies between different parts of an image. This is because the receptive field of the convolution kernels is limited, making it challenging to learn global semantic information crucial for tasks like image segmentation[26]. As a result, developing better methods to capture long-range dependencies is still a challenge for researchers.

Recent research has looked to reduce the limitations of CNN models in capturing long-range dependencies between parts of an image. Drawing inspiration from attention mechanisms in natural language processing, researchers have integrated attention mechanisms with CNN models.

Non-local neural networks [27] use a self-attention approach to obtain the long-term dependencies in the feature map. However, this approach is afflicted with excessive memory requirements and computational complexities. Schlemper et al. [28] proposed an attention-based model which improves model sensitivity and prediction accuracy.

In contrast to CNN-based methods, Transformer [29] models retain the long-range dependencies and capture relationships between arbitrary positions using self-attention without convolutions. A transformer is highly effective in modelling global context and achieves impressive results on downstream tasks, especially when pre-trained on large datasets. However, traditional methods suffer from issues that must be addressed to improve performance. Therefore, we introduce a novel deep learning approach for tumor segmentation. The main aspects and novelty of this work are described below:

- We adopted UNet architecture as our base model for BTS
- The proposed model uses a Swin transform module to incorporate the transformer-based image processing models to increase the accuracy and reduce the complexity.
- We also included a hybrid attention mechanism to improve the overall performance by including the long-term dependencies of local and global contextual information. Further, this article presents a literature review, describes the proposed model, compare its performance, and presents concluding remarks in section II, III, IV, and V, respectively.

2. Literature Review

Segmentation based on computer vision has been an essential part of automated medical image segmentation tasks. Thus, several studies have been undertaken in various biomedical applications. This section considers traditional image pre-processing methods, machine learning, and deep learning-based systems in biomedical image segmentation.

The accustomed image processing methods include several methods such as thresholding, region growing, active contour etc. Umit et al. [30] presented a thresholding-based method that considers image enhancement, morphological operations and pixel subtraction methods to segment tumours.

Elisee et al. [31] introduced a localized active contour-based method capable of processing the intensity inhomogeneity in biomedical images. Moreover, it uses background intensity compensation to compensate for the background intensity information.

However, achieving the desired accuracy remains challenging for these models. Therefore, researchers have suggested and adopted machine learning-based solutions using different feature extraction methods. The obtained features are then used to train machine learning classifiers, such as Venkatesh et al. [32] used KNN classification for BTS. Chenet al. [33] presented a five-fold methodology that includes the following steps: image standardization, noise removal by applying non-local means filtering, and contrast enhancement by applying an improved dynamic histogram method.

The feature extraction is carried out by employing GLCM, and finally, the SVM classifier is used to generate the outcome. These techniques leverage digital image processing[34] and mathematical principles to accomplish the segmentation task. They are relatively straightforward to compute and execute quickly, but achieving the desired accuracy remains challenging. Nowadays, deep learning-based methods have made significant advancements in image

segmentation. They surpass traditional segmentation methods in case of accuracy, with the fully convolutional network being the first to successfully employ deep learning methods in the case of image semantic segmentation.

As discussed before, UNet Segmentation has gained massive attention in this field. This architecture is further optimized in various research, such as Futrega et al. [35] introduced optimized UNet for BTS after analyzing the performance of basic UNet, Residual UNet and Attention UNet models. This study shows that employing deep supervision can improve the segmentation performance, which can be refined further by adding an input channel, increasing encoder depth with several convolutional channels, and adding post-processing steps. Jiang et al. [36] reported the issue in CNN that these networks cannot learn global and remote semantic information efficiently. Transformer schemes have been employed successfully due to their self-attention mechanism to handle global information. Therefore, the authors presented SwinBTS, a combination of transformer methods, encoder-decoder mechanism, and CNN for 3D brain tumor semantic segmentation.

Liu et al. [37] focused on segmenting the WT (Whole Tumour), TC (Tumour Core) and ET (Enhanced Tumour). The TC and ET are important in extracting significant information related to a brain tumor. Therefore, the authors considered the MetricUNet segmentation model based on voxel-metric learning to acquire satisfactory segmentation results. The effectiveness of these models relies on the brain tumour's scale and size, which affects the model's performance.

To solve this issue, the authors deployed Scale-adaptive Super-feature based MetricUNet (S2MetricUNet), which achieves better accuracy on TC and ET by using the novel scale-adaptive metric loss and minimizes computational complexity. Cinar et al. [38] also considered UNet as the base architecture for segmentation and deployed a pre-trained DenseNet121 architecture to improve the segmentation accuracy.

Qin et al. [39] presented an improved UNet3+ model for medical image segmentation based on the stage residual. Generally, they suffer from gradient vanishing problems; therefore, an encoder-based residual structure is employed to reduce this issue.

Further, the BN layer is replaced with the filter response normalization layer. The combination of residual models produces IResUnet3+ three-dimensional (3D) model for segmentation. Currently, transformer-based methods are widely deployed in biomedical image segmentation. Gai et al. [40] developed a residual mix transformer fusion net to

accomplish the BTS task. The encoder module includes a residual mix transformer and residual CNN. This transformer mechanism helps to reduce the loss of path boundary information. Further, a parallel fusion strategy is also used to obtain the local-global information.

Liang et al. [41] introduced TransConver deep learning architecture based on a combination of convolution and transformer-based models. This network uses a transformer-convolution inception module to obtain local and global information with the help of convolution and transformer blocks. Further, these features are aggregated by employing a cross-attention fusion mechanism.

3. Proposed Model

In this section, we present a proposed deep learning-based solution to overcome the issues of the existing BTS-based BTS approach. As discussed before, transformer-based methods have gained massive attention in various applications in the biomedical domain because of their impressive speed and accuracy. We adopted the Swin transform mechanism and incorporated it with the traditional unet mechanism for performing segmentation.

Several methods have adopted swin transform mechanisms such as SwinBTS [36], DenseTrans [43] and CSU-Net [44], etc. however, we have introduced a hybrid attention mechanism with a skip connection module to improvise the local and global information of input image to extract fine-grained segmentation results. To append the attention module, we consider Attention Res-UNet architecture, where the attention block connects the encoder and decoder blocks [45]. The proposed architecture is shown in Figure 2.

It consists of several modules: encoder, decoder, bottleneck, skip connections, and hybrid attention module. These encoder and decoder modules contain a Swin transformer block in which the information comes from the source [46]. The input images are processed and transformed into sequence embedding during the encoder phase. Due to this process, the medical image is partitioned into various non-overlapping patches. Each patch is 4x4 in size, generating the feature map of dimensions 4x4x3.

Furthermore, this model uses a linear embedding layer for converting the dimensions of feature maps into a customizable dimensions. The modified patch tokens are then processed through multiple Swin Transformer modules to produce feature maps. Similarly, these features are processed using the patch assimilation layer to reduce the size of the patches and increase their dimensionality. At the same time, the Swin Transformer block learns how to represent the features effectively.

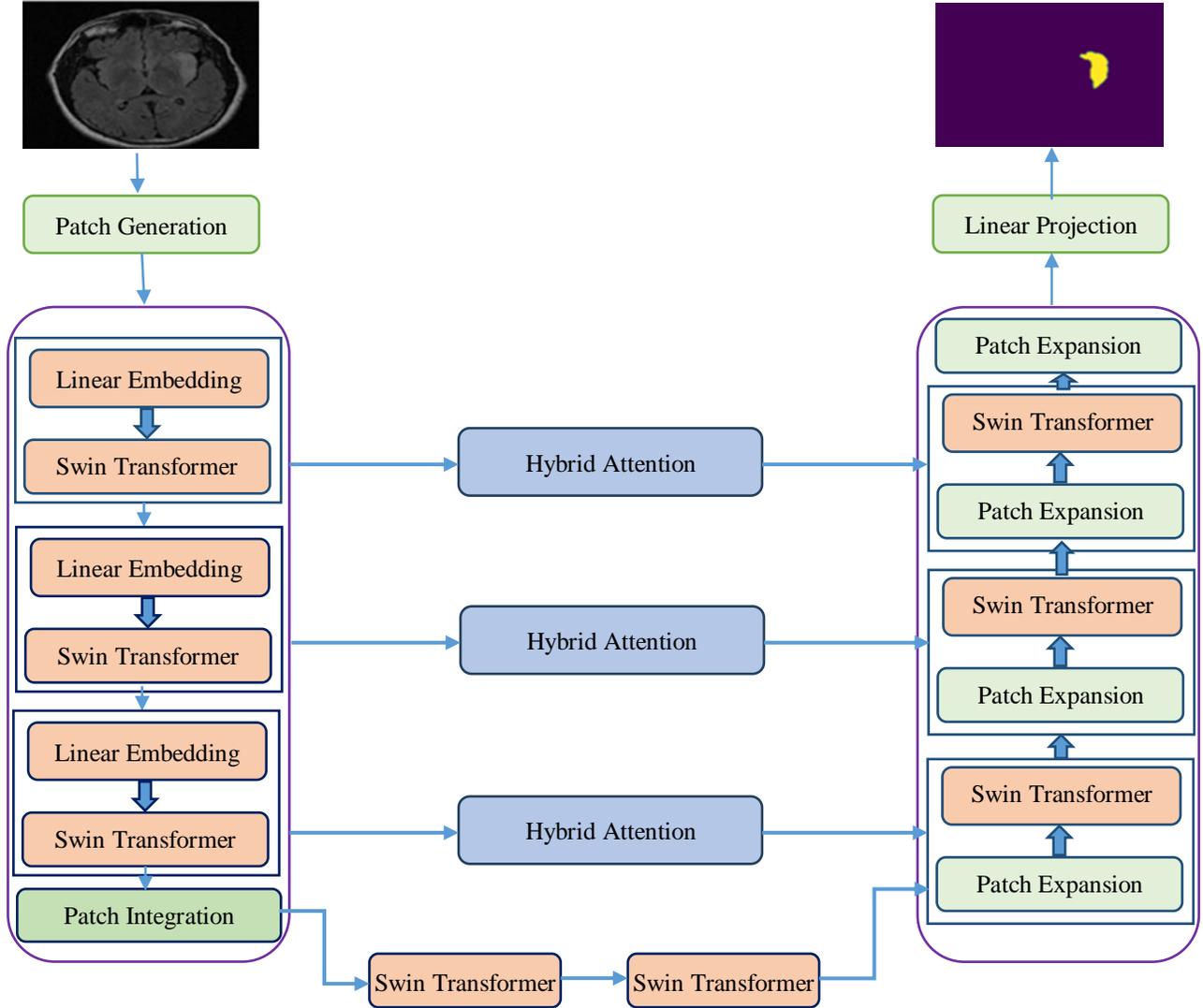


Fig. 2 Proposed deep learning architecture

Similarly, the processed data is passed through the decoder block, which consists swin transformer and patch expanding layer. In order to retain the information of encoder blocks, the obtained features are processed via skip connection and fused with the multiscale features. This process helps to mitigate the loss of spatial information which is caused due to down-sampling process.

On the other hand, it consists of a patch expansion layer formulated to perform upsampling. Finally, a patch expansion layer is deployed for upsampling the data. This upsampling process is helpful in re-establishing the resolution of feature maps to match the original resolution. These blocks are described below:

3.1. Swin Transform Block

Below given figure shows the architecture of the standard transform (Figure 3 (a)) and swin transformer block

(Figure 3 (b)). The conventional Transformer module comprises a stack of L-equal blocks. Each block is made up of attention and Multi-Layer Perceptron components. Moreover, before each attention module and MLP, a Layer Normalization (LN) layer is interpolated. Further, these blocks are connected with the help of residual connections. The result of p -layer in the encoder can be acquired using Equation 1:

$$\hat{z}_p = MSA(L_N(z_{p-1})) + z_{p-1}$$

$$z_p = MSA(L_N(\hat{z}_p)) + \hat{z}_p \quad (1)$$

In the traditional Transformer design, each token is compared to every other token, resulting in a quadratic computational complexity in the number of tokens. However, this is not feasible for tasks that involve dense prediction and high-resolution images.

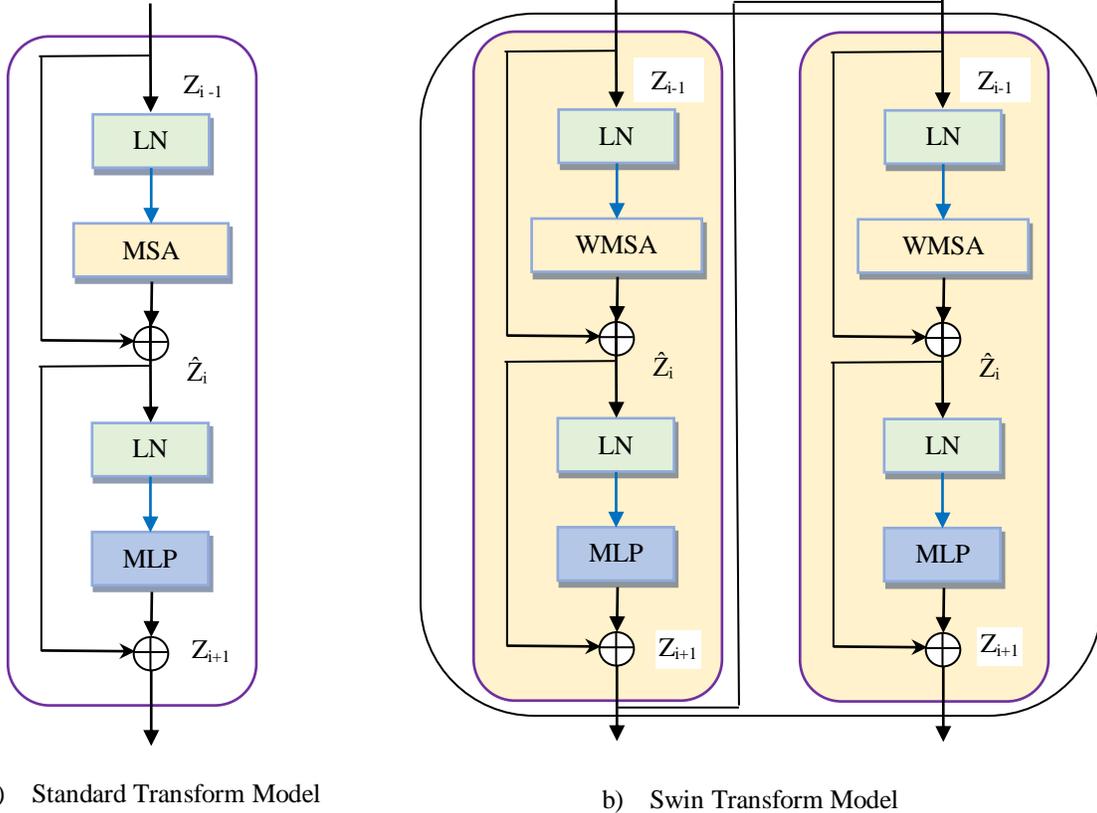


Fig. 3 Standard and swin transform architectures (a) Standard transform model and (b) Swin transform model

To resolve this issue, the Swin Transformer [46] introduces a more efficient approach called Window-based SA and Shifted Window-based MSA, allowing for effective modelling while reducing computational complexity. According to the WMSA approach, the input attribute set is partitioned into several non-overlapping windows. These windows consist of $M \times M$ patches.

Moreover, the WMSA performs self-attention on local windows. Thus the output of WMSA and MLP in p^{th} the layer can be expressed as shown in Equation 2:

$$\begin{aligned} \hat{z}_p &= WMSA \left(L_N(z_{p-1}) \right) + z_{p-1} \\ z_p &= MLP \left(L_N(\hat{z}_p) \right) + \hat{z}_p \end{aligned} \quad (2)$$

However, the WSA approach faces the problem of limited correspondence between different windows. To address this limitation and avoid additional computation, the Shifted Window-based Multi-head Self Attention (SW-MSA) is introduced after the W-MSA, enabling effective window interaction.

This window-based shifting mechanism performs cyclic shifting. By using this shift, the SW-MSA mentioned above,

and the MLP component can be expressed as shown in Equation 3:

$$\begin{aligned} \hat{z}_{p+1} &= SWMSA \left(L_N(z_{p-1}) \right) + z_{p-1} \\ z_{p+1} &= MLP \left(L_N(\hat{z}_{p+1}) \right) + \hat{z}_{p+1} \end{aligned} \quad (3)$$

3.2. Encoder Module

The encoder model is based on the UNet architecture, where we incorporate swin transformer blocks for feature extraction. The process of the swin transform block is described in the previous section. As per the encoder module, the input image is divided into $\frac{H}{s} \times \frac{H}{s}$ as non-overlapping patches where s represents the patch size, these patches are considered as “tokens”, and the linear embedding layer is used for processing for projection to dimension C .

Moreover, these patches are achieved using convolution operations; therefore, additional position information is not required in this phase. The Swin Transformer takes in patch tokens as input, which are then processed through four stages, each containing a specific number of Swin Transformer blocks, including W-MSA and SW-MSA. The input features undergo a patch integration layer during the initial three stages. However, this process reduces the feature

resolution but increases the dimensions of the feature map. The patch integration layer brings together the attributes of adjacent 2x2 patches. Further, a linear concatenation layer is applied to concatenate the features to improve the channel dimensionality. With the help of these operations, the output resolution of the four stages can be extracted as $\frac{H}{s} \times \frac{H}{s}$, $\frac{H}{2s} \times \frac{H}{2s}$, $\frac{H}{4s} \times \frac{H}{4s}$, and $\frac{H}{8s} \times \frac{H}{8s}$ resulting in updated dimensions as C, 2C, 4C and 8C.

3.3. Hybrid Attention Module

In the local context, the self-attention mechanism computes the self-affinities inside each window. Later, these tokens are aggregated into global tokens. This combination is represented as primary information of the window. Further, the entire feature map is down-sampled and then global self-attention can be performed as shown in Equation 4:

$$\begin{aligned} z_l &= SA(X) \\ z_g &= GSA(Concat(z_l)) \\ z &= concat(z_l, Upsample(z_g)) \end{aligned} \quad (4)$$

Where X denotes the input feature map as $X \in R^{H \times W \times C}$, z is the output, and GSA denotes the global self-attention. Further, we incorporate a Gaussian attention mechanism

which helps to improve the perception of the query and reduces the computational complexity. The outcome can be written as Equation 5 :

$$z_{i,j} = e^{-\frac{D_{i,j}^2}{2\sigma^2}} softmax(S(q_{i,j}, K_{i,j})) V_{i,j} \quad (5)$$

Where $q \in Q$ signifies the input query, $e^{-\frac{D_{i,j}^2}{2\sigma^2}}$ denotes the Gaussian weights. Figure 4 shows the architecture of this attention mechanism.

3.4. Decoder Module

In the proposed architecture, the decoder module contains three main stages: upsampling, concatenating, and swin transformer. The input data is up-sampled by two and processed through the skip connection.

Finally, the obtained output is fed to the Swin transformer. The decoder modules aid in utilizing features fully and helps to form long-term dependencies to improve the overall performance. These stages generate an output with a resolution $\frac{H}{4} \times \frac{H}{4}$ however, this process impacts the shallow features; therefore, we perform downsampling and produce the low-level attributes of resolution $H \times W$ and $\frac{H}{2} \times \frac{W}{2}$. These blocks contain 3x3 convolution, normalization layer and ReLu function.

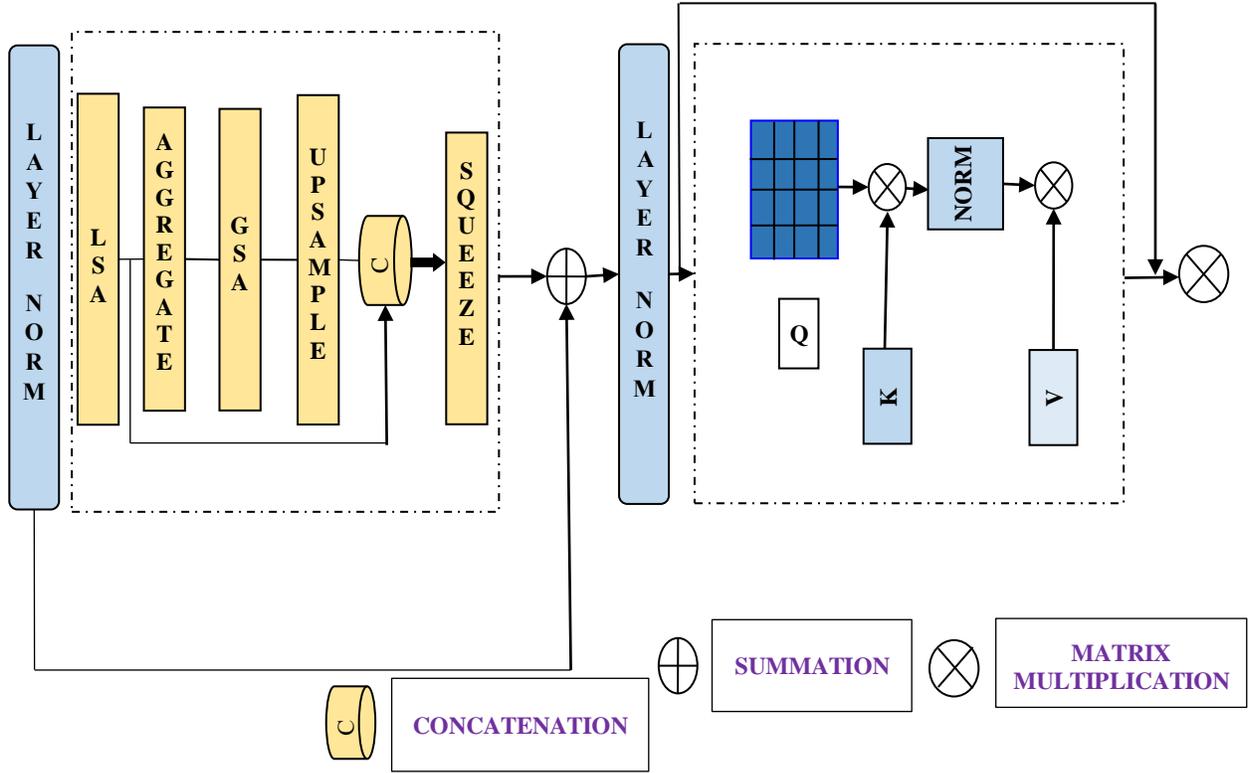


Fig. 4 Hybrid attention model

4. Results and Discussion

In this section, we conduct experiments to analyze the performance of the proposed model and compare how well it performs. We start by utilizing the brain tumour segmentation benchmark dataset and explain the different methods used to assess the model's performance. Then, we describe how our approach was implemented. We are, finally, comparing the results from our model with those achieved by the best existing algorithms.

4.1. Benchmark Dataset Details

The proposed models' performance is analysed using three benchmark datasets: BraTS 2017, BraTS 2018, and BraTS 2019. These datasets were obtained from the publically available library Multimodal Brain Tumor Segmentation Challenge (BraTS) [42].

4.1.1. BraTS 2017 Dataset

This dataset comprises medical records of 285 patients with Glioma. Among these patients, 210 cases are classified as High-Grade Gliomas (HGG) and 75 as Low-Grade Gliomas (LGG). The validation set contains records of 46 patients whose grade is unknown. The data has been manually labelled to establish ground truth. Each patient's case includes four different modalities: Flair, T1, T1ce, and T2.

4.1.2. BraTS 2018 and 2019 Dataset

The BraTS 2018 dataset shares the same training set as the BraTS 2017 dataset but has a different validation set that includes 66 unlabeled patient data. In contrast, the BraTS 2019 dataset has a larger sample size than BraTS 2017 and BraTS 2018. The training set in BraTS 2019 consists of a total of 335 glioma patients, out of which 259 cases belong to HGG and 76 to LGG. Additionally, the number of validation patient data has been increased to 125.

4.2. Performance Measurement Parameters

This section describes the parameters for the performance measurement used for measuring the result of the proposed approach. Mainly, we used Dice Similarity Coefficient (DSC) and Hausdorff distance. The Dice Similarity Coefficient (DSC) measures how much two sets of data overlap. It is frequently used to segment medical images to assess the effectiveness of image segmentation algorithms.

The DSC is computed by taking twice the size of the intersection of the two sets and dividing it by the sum of their sizes. A score of 1 implies a perfect overlap, while 0 indicates no overlap. In medical image segmentation, a high DSC score suggests that the algorithm correctly identifies the relevant area inside the image. This can be expressed by using Equation 6:

$$DSC = \frac{2T_P}{F_N + F_P + 2T_P} \quad (6)$$

Where T_P , F_P , and F_N denotes the true positive, false positive, and false negative values.

Similarly, the Hausdorff distance is a metric used to decide the level of dissimilarity between two data sets. It is commonly applied in computer vision, pattern recognition, and medical image processing. In the segmentation of medical images, the Hausdorff distance is utilized to assess the accuracy of the segmentation algorithm by measuring the most significant distance between the points of a segmented object and the corresponding points of the ground truth object. This metric considers errors' size and spatial distribution in the segmentation result. A smaller Hausdorff distance signifies a better segmentation outcome. It can be computed as expressed in Equation 7:

$$HD(T, P) = \max\{\sup_{t \in T} \inf_{p \in P} d(t, p), \sup_{p \in P} \inf_{t \in T} d(t, p)\} \quad (7)$$

Where sup represents the supremum and inf is used to denote the infimum, t denotes the points on surface T of ground truth, and p denotes the points on the surface P of the predicted region, $d(.,.)$ is the distance between t and p points.

4.3. Comparative Analysis

This section reviews the proposed approach's comparison for publicly available datasets where the proposed method is implemented on training and validation datasets. Below given Table 1 depicts the comparative performance for the Brats 2015 challenge dataset.

Further, we measured the performance metric of the proposed deep learning model for HGG (High-Grade Glioma) cases. The samples are taken from Brats 2017 datasets with 210 HGG cases. The samples are further classified into two parts where 168 samples are employed in training, and 42 are employed for testing. The resultant performance is presented in below given Table 2. As per this experiment, the proposed model achieves average DSC performance of 0.921, 0.895, and 0.887 for whole, core and enhancing tumour cases, respectively.

Further, we measured the performance of the proposed technique for the BraTS 2019 and 2020 validation datasets. We measured the performance concerning the Dice score, sensitivity, specificity, and Hausdorff95 for this process. Table 3 describes the comparison analysis for BraTS 2019 dataset. Similarly, we extended experimental analysis and gauged the performance on BraTS 2020 dataset. The resulting performance is demonstrated in Table 4.

As per the comparison analysis shown in Table 4, the proposed method performs better regarding Dice score, sensitivity, specificity, and Hausdorff95.

Table 1. Dice score performance for BraTs 2015

Method	Whole	Core	Enhancing
DCNN [47]	0.85	0.68	0.87
CLDF [48]	0.79	0.67	0.70
VCTB [49]	0.74	0.54	0.54
AG [50]	0.85	0.70	0.73
3DNeT3 [51]	0.92	0.84	0.77
iLinear [52]	0.86	0.87	0.90
Deep Capsule Network [53]	0.91	0.86	0.85
Proposed Model	0.94	0.91	0.92

Table 2. Dice score comparisons for HGG cases in BraTS 2017 dataset

Method	Whole	Core	Enhancing
UNet	0.881	0.847	0.814
ResUNet	0.886	0.857	0.823
Chen et al. [54]	0.720	0.847	0.810
Kamnitsas et al [55]	0.900	0.857	0.730
Dong et al [56]	0.831	0.750	0.750
Pereira et al [57]	0.840	0.801	0.620
Kermi et al [58]	0.880	0.720	0.820
Zhao et al. [59]	0.865	0.850	0.816
Zhao et al.[60]	0.90	0.83	0.78
Ghaffari et al. [61]	0.90	0.83	0.78
Guan et al. [62]	0.69	0.85	0.68
Proposed Model	0.921	0.895	0.887

Table 3. Comparative analysis of the BraTS 2019 dataset

Method	Dice Score			Sensitivity			Specificity			Hausdorff95		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
S2Metric Unet [63]	0.84	0.83	0.76	0.76	0.86	0.73	-	-	-	-	-	-
3D FCN [64]	0.76	0.89	0.78	-	-	-	-	-	-	-	-	-
SoResUnet [65]	0.72	0.87	0.78	-	-	-	-	-	-	5.9	9.3	11.4
Trans BTS [66]	0.78	0.88	0.81	-	-	-	-	-	-	5.9	7.5	7.5
Cascaded 3D Unet [67]	0.8	0.8	0.8	0.8	0.9	0.8	-	-	-	6.14	4.92	6.75
AMMGS [68]	0.76	0.89	0.81	0.82	0.94	0.85	0.99	0.98	0.99	5.1	8.2	7.2
Proposed Model	0.83	0.91	0.85	0.88	0.951	0.89	0.998	0.991	0.998	6.5	9.20	12.2

Table 4. Comparative analysis of the BraTS 2020 dataset

Method	Dice Score			Sensitivity			Specificity			Hausdorff95		
	ET	WT	TC	ET	WT	TC	ET	WT	TC	ET	WT	TC
MITAU [70]	0.57	0.73	0.61	0.52	0.77	0.62	0.99	0.99	0.99	38.87	20.81	24.22
Prob Unet [71]	0.688	0.819	0.716	0.690	0.846	0.699	0.99	0.998	0.999	36.88	41.52	26.27
3D DMFNet [72]	0.748	0.871	0.748	0.751	0.872	0.718	0.998	0.993	0.998	3.92	9.42	10.09
AHM3D [63]	0.710	0.880	0.740	0.740	0.920	0.740	0.99	0.99	0.99	38.31	6.88	32
DS3D Unet [69]	0.78	0.882	0.815	0.797	0.910	0.787	0.999	0.998	0.999	23.86	7.30	8.16
AMMGS [68]	0.780	0.883	0.817	0.798	0.925	0.80	0.997	0.998	0.999	23.61	7.16	7.98
Proposed Model	0.821	0.902	0.845	0.831	0.944	0.845	0.999	0.998	0.999	24.50	7.30	8.1

Similarly, we measured the performance of proposed approach for the BRATS 2021 dataset in terms of Dice, Sensitivity, and Specificity. Below given Table 5 shows the comparative performance, and Table 6 shows the

performance analysis in terms of dice score. Below, Figure 5 shows the qualitative segmentation outcome using the proposed approach, depicting the original image, its corresponding ground truth and predicted images.

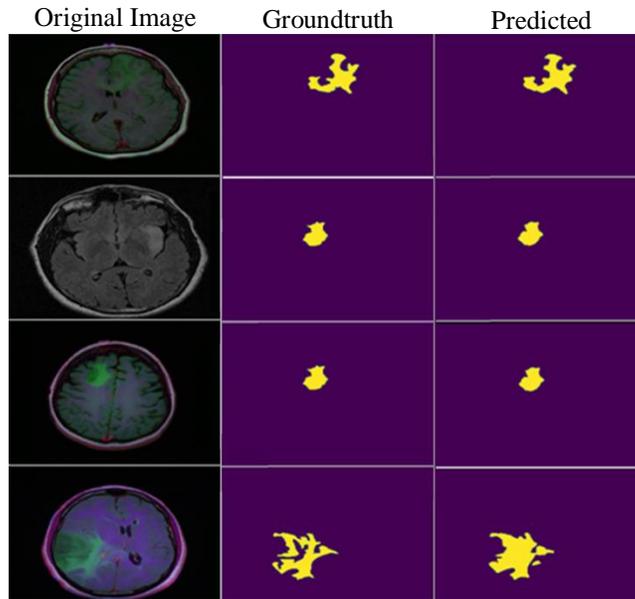


Fig. 5 Sample outcome of the proposed model

Table 5. Comparative analysis of the BraTS 2021 dataset

Method	Dice			Sensitivity			Specificity		
	ET	WT	TC	ET	WT	TC	ET	WT	TC
CapsNet [53]	0.80	0.83	0.81	0.80	0.80	0.77	0.82	0.87	0.85
CapsNet + LDCRF [53]	0.87	0.87	0.85	0.83	0.85	0.83	0.84	0.88	0.86
CapsNet + LDCRF + post processing [53]	0.92	0.92	0.88	0.85	0.90	0.86	0.87	0.93	0.91
Proposed Model	0.94	0.95	0.92	0.91	0.93	0.89	0.91	0.94	0.94

Table 6. Comparative analysis of the BraTS 2021 dataset in terms of dice Score

Method	Dice		
	ET	WT	TC
Bitr UNet [73]	0.82	0.91	0.84
SegResNet [74]	0.86	0.93	0.89
GNN-CNN [75]	0.73	0.89	0.81
Proposed Model	0.94	0.95	0.92

5. Conclusion

This study introduces the Transform Enabled Attention U-Net (TEA-UNet), a U-shaped encoder-decoder framework for medical image segmentation. Our TEA-UNet utilizes the Swin Transformer module to improve the segmentation efficiency of UNet. These Swin Transformer blocks are added in the encoder and innovatively in the decoder. Additionally, we introduce a hybrid attention mechanism

along with the skip to extract multi-scale feature representations. Our experiments on Brats' medical image segmentation tasks demonstrate that our TEA-UNet significantly outperforms other state-of-the-art methods. Moving forward, we will focus on developing more lightweight Transformer-based models and improving the learning of pixel-level intrinsic structural features generated by various vision transformers.

References

- [1] Rebecca L. Siegel, Kimberly D. Miller, and Ahmedin Jemal, "Cancer Statistics, 2015," *A Cancer Journal for Clinicians*, vol. 65, no. 1, pp. 5-29, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Rebecca L. Siegel, Kimberly D. Miller, and Ahmedin Jemal, "Cancer Statistics, 2017," *A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7-30, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Gopal S. Tandel et al., "A Review on a Deep Learning Perspective in Brain Cancer Classification," *Cancers*, vol. 11, no. 1, pp. 1-32, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Amin Zadeh Shirazi et al., "The Application of Deep Convolutional Neural Networks to Brain Cancer Images: A Survey," *Journal of Personalized Medicine*, vol. 10, no. 4, pp. 1-27, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Mahwash Mukhtar et al., "Nanomaterials for Diagnosis and Treatment of Brain Cancer: Recent Updates," *Chemosensors*, vol. 8, no. 4, pp. 1-31, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Asaf Raza et al., "A Hybrid Deep Learning-Based Approach for Brain Tumor Classification," *Electronics*, vol. 11, no. 7, pp. 1-17, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Mohammed Rasool et al., "A Hybrid Deep Learning Model for Brain Tumour Classification," *Entropy*, vol. 24, no. 6, pp. 1-16, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Neelum Noreen et al., "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," *IEEE Access*, vol. 8, pp. 55135-55144, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Rajan Hossain, Roliana Binti Ibrahim, and Haslina Binti Hashim, "Automated Brain Tumor Detection using Machine Learning: A Bibliometric Review," *World Neurosurgery*, vol. 175, pp. 57-68, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Philipp Lohmann et al., "Combined Amino Acid Positron Emission Tomography and Advanced Magnetic Resonance Imaging in Glioma Patients," *Cancers*, vol. 11, no. 2, pp. 1-13, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] H. J. Biersack, F. Grünwald, and J. Kropp, "Single Photon Emission Computed Tomography Imaging of Brain Tumors," *Seminars in Nuclear Medicine*, vol. 21, no. 1, pp. 2-10, 1991. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Jonathan Huang et al., "Artificial Intelligence Applications in Pediatric Brain Tumor Imaging: A Systematic Review," *World Neurosurgery*, vol. 157, pp. 99-105, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Muhammad Farhan Safdar, Shayma Saad Alkobaisi, and Fatima Tuz Zahra, "A Comparative Analysis of Data Augmentation Approaches for Magnetic Resonance Imaging (MRI) Scan Images of Brain Tumor," *Acta informatica medica*, vol. 28, no. 1, pp. 29-36, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Adrien Heintz, and Jean-Marc Constans, "Role of Magnetic Resonance Spectroscopy in Clinical Management of Brain Tumors," *Atlas of Clinical Cases on Brain Tumor Imaging*, pp. 49-67, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] R. Sakthi Prabha, and M. Vadivel, "Brain Tumor Stages Prediction using FMS-DLNN Classifier and Automatic RPO-RG Segmentation," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 2, pp. 110-121, 2023. [[CrossRef](#)] [[Publisher Link](#)]

- [16] Javaria Amin et al., “Brain Tumor Detection using Statistical and Machine Learning Method,” *Computer Methods and Programs in Biomedicine*, vol. 177, pp. 69-79, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Biswajit Jena, Gopal Krishna Nayak, and Sanjay Saxena, “An Empirical Study of Different Machine Learning Techniques for Brain Tumor Classification and Subsequent Segmentation using Hybrid Texture Feature,” *Machine Vision and Applications*, vol. 33, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Dayong Wang et al., “Deep Learning for Identifying Metastatic Breast Cancer,” *arXiv:1606.05718*, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Wejdan L. Alyoubi, Wafaa M. Shalash, and Maysoun F. Abulkhair, “Diabetic Retinopathy Detection through Deep Learning Techniques: A Review,” *Informatics in Medicine Unlocked*, vol. 20, p. 100377, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Zihui Guo et al., “Deep LOGISMOS: Deep Learning Graph-Based 3D Segmentation of Pancreatic Tumors on CT Scans,” *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, pp. 1230-1233, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] R. Tamilaruvi et al., “Brain Tumor Detection in MRI Images using Convolutional Neural Network Technique,” *SSRG International Journal of Electrical and Electronics Engineering*, vol. 9, no. 12, pp. 198-208, 2022. [[CrossRef](#)] [[Publisher Link](#)]
- [22] Jonathan Long, Evan Shelhamer, and Trevor Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-3440, 2015. [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, Munich, Germany, pp. 234-241, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Zongwei Zhou et al., “UNet++: A Nested U-Net Architecture for Medical Image Segmentation,” *ML-CDS 2018: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Granada, Spain, pp. 3-11, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Ahmed Shaker, Wai Yeung Yan, and Paul E. LaRocque, “Automatic Land-Water Classification using Multispectral Airborne LiDAR Data for Near-Shore and River Environments,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 94-108, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] D. Napoleon, and M. Praneesh, “Detection of Brain Tumor using Kernel Induced Possibilistic C-Means Clustering,” *International Journal of Computer Organization Trends and Technology (IJCOT)*, vol. 3, no. 5, pp. 40-42, 2013. [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Xiaolong Wang et al., “Non-Local Neural Networks,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7794-7803, 2018. [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Jo Schlemper et al., “Attention Gated Networks: Learning to Leverage Salient Regions in Medical Images,” *Medical Image Analysis*, vol. 53, pp. 197-207, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Ashish Vaswani et al., “Attention is All you Need,” *31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, pp. 1-11, 2017. [[Google Scholar](#)] [[Publisher Link](#)]
- [30] Umit Ilhan, and Ahmet Ilhan, “Brain Tumor Segmentation Based on a New Threshold Approach,” *Procedia Computer Science*, vol. 120, pp. 580-587, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [31] Jakhongir Nodirov, Akmalbek Bobomirzaevich Abdusalomov, and Taegkeun Whangbo, “Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images,” *Sensors*, vol. 22, no. 17, pp. 1-17, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [32] Venkatesh, and M. Judith Leo, “MRI Brain Image Segmentation and Detection using K-NN Classification,” *Journal of Physics: Conference Series*, vol. 1362, no. 1, p. 012073, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [33] Baoshi Chen et al., “A Novel Extended Kalman Filter with Support Vector Machine Based Method for the Automatic Diagnosis and Segmentation of Brain Tumors,” *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105797, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [34] Rohan K. Gajre, Savita A. Lothe, and Santosh G. Vishwakarma, “Identification of Brain Tumor using Image Processing Technique: Overviews of Methods,” *SSRG International Journal of Computer Science and Engineering*, vol. 3, no. 10, pp. 48-52, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [35] Michał Futrega et al., “Optimized U-Net for Brain Tumor Segmentation,” *BrainLes 2021: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 15-29, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [36] Yun Jiang et al., “SwinBTS: A Method for 3D Multimodal Brain Tumor Segmentation using Swin Transformer,” *Brain Sciences*, vol. 12, no. 6, pp. 1-15, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [37] Yujian Liu et al., “Scale-Adaptive Super-Feature Based MetricUNet for Brain Tumor Segmentation,” *Biomedical Signal Processing and Control*, vol. 73, p. 103442, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [38] Necip Cinar, Alper Ozcan, and Mehmet Kaya, “A Hybrid DenseNet121-UNet Model for Brain Tumor Segmentation from MR Images,” *Biomedical Signal Processing and Control*, vol. 76, p. 103647, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [39] Junying Zeng et al., “Improved U-Net3+ with Stage Residual for Brain Tumor Segmentation,” *BMC Medical Imaging*, vol. 22, pp. 1-15, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [40] D. Gai et al., “RMTF-Net: Residual Mix Transformer Fusion Net for 2D Brain Tumor Segmentation,” *Brain Sciences*, vol. 12, no. 9, pp. 1-18, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [41] Junjie Liang et al., “TransConver: Transformer and Convolution Parallel Network for Developing Automatic Brain Tumor Segmentation in MRI Images,” *Quantitative Imaging in Medicine and Surgery*, vol. 12, no. 4, pp. 2397-2415, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [42] Spyridon Bakas, MICCAI BRATS - The Multimodal Brain Tumor Segmentation Challenge, 2021. [Online]. Available: <http://braintumorsegmentation.org/>
- [43] Li ZongRen et al., “DenseTrans: Multimodal Brain Tumor Segmentation using Swin Transformer,” *IEEE Access*, vol. 11, pp. 42895-42908, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [44] Yu Chen et al., “CSU-Net: A CNN-Transformer Parallel Network for Multimodal Brain Tumour Segmentation,” *Electronics*, vol. 11, no. 14, pp. 1-12, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [45] Dhiraj Maji, Prarthana Sigedgar, and Munendra Singh, “Attention Res-UNet with Guided Decoder for Semantic Segmentation of Brain Tumors,” *Biomedical Signal Processing and Control*, vol. 71, p. 103077, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [46] Ze Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10012-10022, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [47] Kuan-Lun Tseng et al., “Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation,” *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 6393-6400, 2017. [[Google Scholar](#)] [[Publisher Link](#)]
- [48] Loic Le Folgoc et al., “Lifted Auto-Context Forests for Brain Tumour Segmentation,” *BrainLes 2016: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 171-183, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [49] Edgar A. Rios Piedra et al., “Brain Tumor Segmentation by Variability Characterization of Tumor Boundaries,” *BrainLes 2016: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 206-216, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [50] Bi Song et al., “Anatomy-Guided Brain Tumor Segmentation and Classification,” *BrainLes 2016: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Berlin/Heidelberg, Germany, pp. 162-170, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [51] Jiten Chaudhary, Rajneesh Rani, and Aman Kamboj, “Deep Learning-Based Approach for Segmentation of Glioma Sub-Regions in MRI,” *International Journal of Intelligent Computing and Cybernetics*, vol. 13, no. 4, pp. 389-406, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [52] Saddam Hussain, Syed Muhammad Anwar, and Muhammad Majid, “Segmentation of Glioma Tumors in Brain using Deep Convolutional Neural Network,” *Neurocomputing*, vol. 282, pp. 248-261, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [53] Mahmoud Elmezain et al., “Brain Tumor Segmentation using Deep Capsule Network and Latent-Dynamic Conditional Random Fields,” *Journal of Imaging*, vol. 8, no. 7, pp. 1-16, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [54] Lele Chen et al., “MRI Tumor Segmentation with Densely Connected 3D CNN,” *Medical Imaging 2018: Image Processing*, Houston, Texas, US, vol. 10574, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [55] Konstantinos Kamnitsas et al., “Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation,” *Medical Image Analysis*, vol. 36, pp. 61-78, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [56] Hao Dong et al., “Automatic Brain Tumor Detection and Segmentation using U-Net Based Fully Convolutional Networks,” *MIUA 2017: Medical Image Understanding and Analysis*, pp. 506-517, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [57] Sérgio Pereira et al., “Brain Tumor Segmentation using Convolutional Neural Networks in MRI Images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240-1251, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [58] Adel Kermi, Issam Mahmoudi, and Mohamed Tarek Khadir, “Deep Convolutional Neural Networks using U-Net for Automatic Brain Tumor Segmentation in Multimodal MRI,” *BrainLes 2018: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 37-48, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [59] Xiaomei Zhao et al., “A Deep Learning Model Integrating FCNNs and CRFs for Brain Tumor Segmentation,” *Medical Image Analysis*, vol. 43, pp. 98-111, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [60] Xiangyu Zhao et al., “Prior Attention Network for Multi-Lesion Segmentation in Medical Images,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 12, pp. 3812-3823, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [61] Mina Ghaffari et al., “Automated Post-Operative Brain Tumour Segmentation: A Deep Learning Model Based on Transfer Learning from Pre-Operative Images,” *Magnetic Resonance Imaging*, vol. 86, pp. 28-36, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [62] Xi Guan et al., “3D AGSE-VNet: An Automatic Brain Tumor MRI Data Segmentation Framework,” *BMC Medical Imaging*, vol. 22, no. 1, pp. 1-18, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [63] Yujian Liu et al., “Scale-Adaptive Super-Feature Based MetricUNet for Brain Tumor Segmentation,” *Biomedical Signal Processing and Control*, vol. 73, p. 103442, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [64] Jindong Sun et al., “Segmentation of the Multimodal Brain Tumor Image Used the Multi-Pathway Architecture Method Based on 3D FCN,” *Neurocomputing*, vol. 423, pp. 34-45, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [65] Ning Sheng et al., “Second-Order ResU-Net for Automatic MRI Brain Tumor Segmentation,” *Mathematical Biosciences and Engineering*, vol. 18, no. 5, pp. 4943-4960, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [66] Wenxuan Wang et al., “TransBTS: Multimodal Brain Tumor Segmentation using Transformer,” *Medical Image Computing and Computer Assisted Intervention-MICCAI 2021*, Strasbourg, France, pp. 109-119, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [67] Pengyu Li et al., “Automatic Brain Tumor Segmentation from Multiparametric MRI Based on Cascaded 3D U-Net and 3D U-Net++,” *Biomedical Signal Processing and Control*, vol. 78, p. 103979, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [68] Xiangbin Liu et al., “Attention-Based Multimodal Glioma Segmentation with Multi-Attention Layers for Small-Intensity Dissimilarity,” *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 4, pp. 183-195, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [69] Théophraste Henry et al., “Brain Tumor Segmentation with Self-Ensembled, Deeply-Supervised 3D U-Net Neural Networks: A BraTS 2020 Challenge Solution,” *BrainLes 2020: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lima, Peru, pp. 327-339, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [70] Navchetan Awasthi, Rohit Pardasani, and Swati Gupta, “Multi-Threshold Attention U-Net (MTAU) Based Model for Multimodal Brain Tumor Segmentation in MRI Scans,” *BrainLes 2020: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lima, Peru, pp. 168-178, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [71] Chinmay Savadikar, Rahul Kulhalli, and Bhushan Garware, “Brain Tumour Segmentation using Probabilistic U-Net,” *BrainLes 2020: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lima, Peru, pp. 255-264, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [72] Chen Chen et al., “3D Dilated Multi-Fiber Network for Real-Time Brain Tumor Segmentation in MRI,” *MICCAI 2019: Medical Image Computing and Computer Assisted Intervention*, pp. 184-192, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [73] Qiran Jia, and Hai Shu, “BiTr-Unet: A CNN-Transformer Combined Network for MRI Brain Tumor Segmentation,” *BrainLes 2021: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 3-14, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [74] Md Mahfuzur Rahman Siddiquee, and Andriy Myronenko, “Redundancy Reduction in Semantic Segmentation of 3D Brain Tumor MRIs,” *arXiv:2111.00742*, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [75] Camillo Saueressig et al., “A Joint Graph and Image Convolution Network for Automatic Brain Tumor Segmentation,” *BrainLes 2021: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pp. 356-365, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]