Computers in Biology and Medicine 194 (2025) 110531

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Implementation of biomedical segmentation for brain tumor utilizing an adapted U-net model

Farah F. Alkhalid *, Nibras Z. Salih

University of Technology- Iraq, Control and Systems Engineering Department, Baghdad, Iraq

ARTICLE INFO

Keywords:
Segmentation
U net
Attention
MRI
Brain tumor

ABSTRACT

Using radio signals from a magnetic field, magnetic resonance imaging (MRI) represents a medical procedure that produces images to provide more information than typical scans. Diagnosing brain tumors from MRI is difficult because of the wide range of tumor shapes, areas, and visual features, thus universal and automated system to handle this task is required. Among the best deep learning methods, the U-Net architecture is the most widely used in diagnostic medical images. Therefore, U-Net-based attention is the most effective automated model in medical image segmentation dealing with various modalities. The self-attention structures that are used in the U-Net design allow for fast global preparation and better feature visualization. This research aims to study the progress of U-Net design and show how it improves the performance of brain tumor segmentation. We have investigated three U-Net designs (standard U-Net, Attention U-Net, and self-attention U-Net) for five epochs to find the last segmentation. An MRI image dataset that includes 3064 images from the Kaggle website is used to give a more comprehensive overview. Also, we offer a comparison with several studies that are based on U-Net structures to illustrate the evolution of this network from an accuracy standpoint. U-Net-based self-attention has demonstrated superior performance compared to other studies because self-attention can enhance segmentation quality, particularly for unclear structures, by concentrating on the most significant parts. Four main metrics are applied with a loss function of 5.03 %, a validation loss function of 4.82 %, a validation accuracy of 98.49 %, and an accuracy of 98.45 %.

1. Introduction

With the developments artificial intelligence field, deep learning has gained significant attention and has been used for medical image analysis. Although the use of deep learning for healthcare scanning has advanced, there are still issues requiring to be resolved. Generally, image segmentation involves dividing an image into meaningful sections associated with different categories to identify a set of pixels into a particular category. In most categorization techniques, only structural details, intensity levels, or a combination of object intensities and more structural details are used. Multi-level feature representation of image intensity is part of medical imaging segmentation; classifying the anatomical structure within the patterns is the first step in segmenting medical images. Then, nonlinear modifications are applied by registering the images to draw patterns in a specific region. The right label is then chosen from the label's space and assigned to the segments using a voting mechanism [1].

Segmentation may be used in many fields like image analysis, text processing, and customer marketing. Also in particle detection of nanomaterial [2]. In computer vision science, segmentation means partitioning the image into meaningful segments that have common features, the main goal of segmentation is to analyze the image by segments [3].

Many techniques are supposed to be used for image segmentation, and feature extracting, these techniques are based on one of the following approaches: Thresholding, Edge, Region, Clustering, and Deep learning [4,5]. Segmentation-based "Thresholding" is the simplest approach, hence it creates the segments depending on the pixel's intensity. While the segmentation-based "Edge" focuses on detecting the boundaries. Segmentation-based "Regions" means dividing the image into segments depending on the predefined criteria. On the other hand, clustering segmentation divides the image depending on Cluster algorithms such as K-mean. Finally, segmentation-based deep learning, which uses convolutional neural networks (CNNs), improves more accurate metrics [6]. CNN extracts features and divides information automatically which consists of a pooling layer, a convolution layer, and

* Corresponding author.
E-mail address: farah.f.alkhalid@uotechnology.edu.iq (F.F. Alkhalid).

https://doi.org/10.1016/j.compbiomed.2025.110531
Received 23 December 2024; Received in revised form 19 April 2025; Accepted 2 June 2025
Available online 11 June 2025

a Rectified Linear Unit activation function that is called the simple design of CNN [6,7].

Therefore, one of the problems with segmenting 3D MRI medical pictures is that it takes a lot of time and prone to tiny errors due to variability. It will be possible to find a more efficient way to extract target objects from MRI images through (CNN) deep learning [8]. The most important area in healthcare image analysis is brain tumor segmentation, which aims to identify the brain tumor accurately [9,10]. Tumor segmentation is done by applying different deep-learning designs and various techniques that depend on the skill of medical professionals. A brain tumor can arise as a result of abnormal development of brain tissue while the exact symptoms and causes are not known yet [11]. Therefore, early detection of any potential threat is crucial for saving people's lives [12].

This study aims to create a model that helps to identify cancer types automatically and reliably by introducing the latest methods for brain tumor segmentation using the U-Net architecture as a validated basis. Also, putting these techniques in one place and focusing on key patterns in research can help guide future work in this area. The outcomes are compared with the most recent literature to see how incremental research has improved. Moreover, the main trends in research are shown, which can serve as a roadmap for further study in this field by compiling recent approaches in a single location.

Three models are trained for tumor segmentation known as (standard-Unet, attention-Unet, and self-attention Unet) to provide the useful structure of the suggested model using several metrics such as loss function, accuracy, validation loss function, and validation accuracy.

The sections of this paper are organized as follows: In Section 2, the previous works of other studies are briefly explained. In Section 3, we explain a comprehensive analysis of the U-Net architecture. In Section 4, the proposed model and dataset are briefly illustrated. Finally, the result and conclusion are described in Sections 5 and 6.

2. Past works

The basic techniques for brain tumor segmentation are 2D segmentation, which is based on slicing, and 3D segmentation, which is based on MRI techniques. For 3D segmentation based on MRI, there is not enough labeled training data, and scaling the dataset is challenging due to memory issues and the huge parameters of the networks. Therefore, the massive amounts of medical data, high variability and the computational complexity in images is solved by deep learning. As a result, several approaches have been carried out to address this problem, especially in Rehman et al. [13] proposed multiple U-net neural networks, by adding residual extended skip (RES) and wide context (WC), into the standard U-network. They achieved good accuracy of tumor segmentation by expanding the valid response area which helps to identify more diverse patterns. The authors utilized clustering properties to extract specific details to improve segmentation efficiency. BraTS 2017 and 2018 datasets, were applied to assess the suggested BU-Net.

Authors in Ref. [14] developed the U-Net architecture to support information propagation across the network and enhance the use of recurrent feature maps using images from satellites. The authors used DenseNet to downsample the weights, and converge the features map, and then they used skip connection to downsample and upsample the weights.

Agalari et al. [15] designed an automated segmentation of brain tumors using MRI techniques by constructing three models, and both local and global aspects were used simultaneously in the model designs. The authors implemented structures with fewer parameters than the original U-Net, which had low computational costs and produced positive results when tested on the BRATS 2018 database with a sensitivity of 89.19 %.

In 2021, Zhang et al. [16] focused on leveraging prior information about brain segmentation by combining 3D and 2D learning and evaluation processes. The authors combined the three corresponding outputs from three separate Unet to determine the final segmentation and check its correctness using BraTS 2018 and 2020 datasets with an average dice score of 87.01 %.

In 2022 [17] Agrawal et al. used the 3D-Unet model to segment and detect the MRI images volumetrically, and then applied CNN to classify the tumor by the BraTS 2020 dataset. The effectiveness of this model was evaluated using several metrics with a loss function of 0.17 and a precision of 88 %.

In [18], the authors suggested a deep design known as attention Res-Unet with a guidance decoder to segment brain tumors using the BraTS 2019 dataset. The suggested design can regulate the learning process at every decoding level using multiple loss functions for each level to improve feature map generation by supervising the training process.

In 2023, Raza et al. [19] presented a comprehensive system for automated 3D brain tumor segmentation using a hybrid residual network approach with the U-Net model called (dResU-Net) model. To address the gradient vanishing problem, the residual network was employed as an encoder and the Unet model as the decoder using BraTS 2020 and BraTS 2021 datasets. The authors utilized skip connections across residual networks and convolutional blocks to hasten the training process.

Yousef et al. [20] presented an analysis of several Unet structures to illustrate the functionality and development of this structure from an efficiency standpoint and enhance the performance of brain tumor segmentation using the BraTS 2020 dataset with a dice score of 95 %.

The authors in Ref. [21] combined U-Net topologies layered on top of each other, known as modified U-Net. The initial U-Net combined three architectures: VGG-19, DenseNet, and a pre-trained Xception model. The authors used semantic segmentation by defining each pixel in the image to which the class belongs; the model achieved a good accuracy of 0.9494 %.

Arhami et al. [22] added a batch normalization step for each layer in U-Net segmentation to solve the gradient vanishing issue and help speed up weight convergence during training. Good performance results, with an accuracy of 91.4 %, were obtained using the U-Net architecture and batch normalization techniques applied to the cervical smear segmentation.

In [23], the authors proposed a modified U-Net by replacing each layer with a residual block using three axes. Then merging two image axes as the input using both the ISBI2015 and OpenMS datasets with the highest overall result of 93.09 %.

In 2024 [24], Parmar et al. studied three architectures of 3D-Unet to segment brain tumors using the BraTS 2020 dataset. Z-score standardization was used to adjust and normalize bias in MRI. Memory consumption and class disparity were decreased using two patch-generating techniques. With a dice score of 93 %, the architecture's efficiency was determined using variance and entropy measurements.

In [25], the authors designed multiple scales of residual CNN known as (mRes-Unet) using the BraTS 2021 dataset. Skip connections were utilized in the U-Net architecture to bridge the semantic gap across encoder and decoder systems. Residual structures were applied to ensure that deep networks get enough information and mitigate the problem of vanishing gradients through the learning process with an average dice score of 92.89 %.

In this work, a new adapted U-Net model for segmentation is proposed, this model is based on the U-Net architecture using self-attention blocks, these blocks allow for enhancing feature representation and improving the global context. Where self-attention blocks can consider the relationship between any two pixels regardless of the distance between them, this ability allows the model to understand the context of the imaging, while the standard U-Net depends on the convolutional layers features extractions which are limited in capturing the long-range dependencies of input images. Table 1below shows a table of related work.

Table 1
A brief table for related work.

Authors. Year	Description
Rehman et al. 2020	Proposed multiple U-net neural networks, by adding residual extended skip (RES) and wide context (WC), into the standard U-Network.
Soni et al. 2020	Developed the U-Net architecture to support information propagation across the network and enhances the use of recurrent feature maps using images from satellites.
Agalari et al. 2021	Designed an automated segmentation of brain tumors using MRI techniques by constructing three models.
Zhang et al. 2021	Focused on leveraging prior information about brain segmentation by combining 3D and 2D learning and evaluation processes.
Agrawal et al. 2022	Used the 3D-Unet model to segment and detect the MRI images volumetrically, and then applied CNN to classify the tumor by the BraTS 2020 dataset.
Maji et al. 2022	Suggested a deep design known as attention Res-Unet with a guidance decoder to segment brain tumors using the BraTS 2019 dataset.
Raza et al. 2023	Presented a comprehensive system for automated 3D brain tumor segmentation using a hybrid residual network approach with the U-Net model called (dResU-Net) model.
Yousef et al. 2023	Presented an analysis of several U-Net structures to illustrate the functionality and development using the BraTS 2020 dataset with a dice score of 95 %.
Deep Deb. 2023	Combined U-Net topologies layered on top of each other, known as modified U-Net.
Arhami et. al. 2024	Added a batch normalization step for each layer in U-Net segmentation to solve the gradient vanishing issue and help speed up weight convergence during training.
Amaludin et al. 2024	Proposed a modified U-Net by replacing each layer with a residual block using three axes.
Parmar et al. 2024	Studied three architectures of 3D-Unet to segment brain tumors using the BraTS 2020 dataset.
Wang et al. 2024	Designed multiple scales of residual CNN known as (mRes-Unet) using the BraTS 2021 dataset.
This Study	Proposed a new adapted U-Net model for segmentation, this model is based on the U-Net architecture using self-attention blocks, these blocks allow for enhancing feature representation and improving the global context.

3. U-net architecture

Generally, a U-Net network consists of CNN layers, it is proposed particularly for biomedical segmentation, the main components of this network are [26,27]:

- Encoder (Contracting path): In this part, the features are extracted, and one can find CNN layers followed by layers, this part is the same as the standard CNN network high-layer.
- Decoder (Expanding path): In this part, the features are upsampling to the main image with the same size as the original image, using a transposed CNN layer.
- Skip-Connection: it is between the Encoder and Decoder used to transfer features map from the Contracting path to the Expanding path directly.

The following Fig. 1 below shows the main components [26].

Fig. 1. U-Net components.

The U-Net model is especially used for applications that use boundary techniques.

3.1. Standard U-net model

It is an architecture of CNN layers, used for biomedical image segmentation, however, it is used for other applications; its name comes from the "U" shape of the network. Fig. 2 below shows the standard model [28,29].

It simply consists of an Encoder, Bottleneck, and Decoder, the first part on the left side of the model is the Encoder part, the input's features are extracted, and at the same time the spatial dimensions are reduced, also called "Constructing Path", this part consists of CNN, Activation functions and Pooling layers:

$$Conv(X) = W^{*}X + b$$

Where W is the Weights, X is the input, and b is the bias.

$$ReLU(X) = Max(0, X)$$

Where Max pooling is computed using stride with (2 × 2);

The second part at the bottom of the model is a Bottleneck, it is the deepest part of the model and contains a high level of features, it is considered as the bridge between the Encoder and Decoder. In this part, only CNN layers are found without pooling layers. The last part on the right side of the model is the Decoder, in this part, the features are reconstructed gradually reaching to input image size and then combined with Encoder layers, using up-sampling (Transposed convolutions layers) and skip connection layers. It is also called "Expanding Path":

$$UpSampling(X) = W^{T*}X + b$$

$$SkipConnection(S, Z) = [S, Z]$$

Where S is Up sampling, and Z is the Encoder features map.

3.2. Attention blocks

This mechanism was called "Bahdanau attention" because of the author Dzmitry, who founded this mechanism, and two other co-authors in 2015 [30]. This approach allowed the model to focus on different parts of the input features. It involves calculating the alignment score which is also called energy *eij* between hidden encoder *hij* and hidden decoder *dj-1* [31]:

$$eij = v^{T}\tanh\left(W_1 h_i + W_2 d_{j-1} + b\right)$$

Where, *hi* is the hidden positions in the encoder at time *i*, and *di* is the

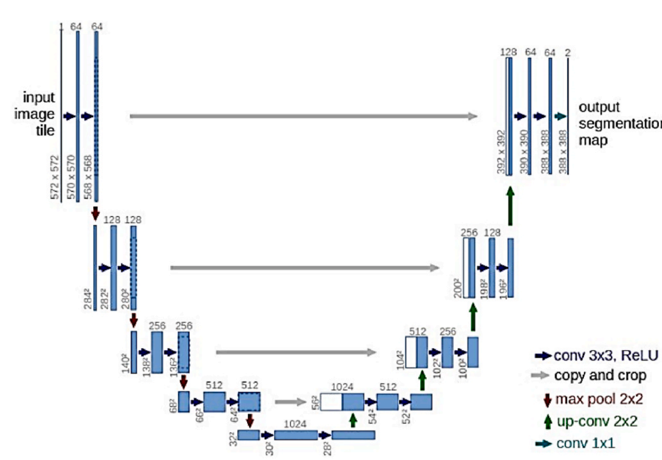

Fig. 2. U-Net model architecture.

hidden positions in the decoder at the previous time equal *j-1*.

W_1, W_2 are the weights, and v is the weights vector, and finally b is the bias.

Then, compute the attention weights and apply normalizing to the alignment score (energy) *eij* using the Softmax function:

$$\alpha ij = \frac{\exp\ (e_{ij})}{\sum_{k=1}^{T} \exp\ (e_{ik})}$$

Where k is the step time sequence.

The context vector is the summation of the attention weights α*ij* by the hidden layers:

$$C_i = \sum_{j=1}^{T} \alpha ij\ hj$$

3.3. Self-attention blocks

The attention blocks allow for updating the weights of the inputs matrix *(X)*, considering the task to do. The attention blocks depend on three major components, the first is called Query (*Q*), which refers to the component to calculate learning weights *(W)*, which can be applied by Refs. [32,33]:

$$Q = XW_Q$$

The second represents the Key (*K*), which refers to a component that is used to compare with *Q*, it can be calculated by:

$$K = XW_K$$

Finally, Value (*V*) is the updated component that reflects the *Q* but may be not equal. The formula of calculation *V* is:

$$V = XW_V$$

The score of attention blocks is calculated by applying the dot product *Q* matrix and the transpose of *the K* matrix:

$$Score = QK^T$$

As known, the dot product may lead to large values, which causes small gradients. To solve this problem, rescaling is applied through division on *(dk)*, the square root of the dimension of vector *K*:

$$Scaled\ Score = \frac{QK^T}{\sqrt{dk}}$$

The gained scores are passed through the SoftMax function to convert the scores to probability weights by summing 1:

$$Attention\ wieghts = SoftMax\left(\frac{QK^T}{\sqrt{dk}}\right)$$

All these elements are computed sequentially to approve the self-attention model with complex relations.

4. Proposed model

The proposed model has been derived from the standard U-Net model, this is done by modifying the U-Net model by adding attention blocks to the decoder layers of the standard model to improve well metrics. Fig. 3 below shows the proposed model diagram.

To simplify the modification applied to the Standard U-Net mode, the proposed model is divided into three parts (Encoder, Bottleneck, and Decoder). Traditionally, the Encoder part consists of many layers, these layers are common between the proposed model and the standard U-Net model. The figure below denotes this part of the model with layers defined in (Figs. 4–6).

The modified model is implemented using Python, based on Jupyter Notebook, the Algorithm below denotes in detail the steps of the proposed model with values. Three models are implemented and discussed, the first is the standard model of using standard U-net as shown in Algorithm 1, the second is using attention blocks with U-net model as shown in Algorithm 2, and the third is using self-attention blocks with U-net model as shown in Algorithm 3. First of all, import the required libraries like Numpy, and TensorFlow, then define the modified model, as shown above. Next, load the dataset (images and masks). After that, compile the model to show the metrics. The proposed model is useful in the healthcare sector for medical picture diagnosis and understanding the image context using attention blocks because it depends on an automated model for segmenting medical images.

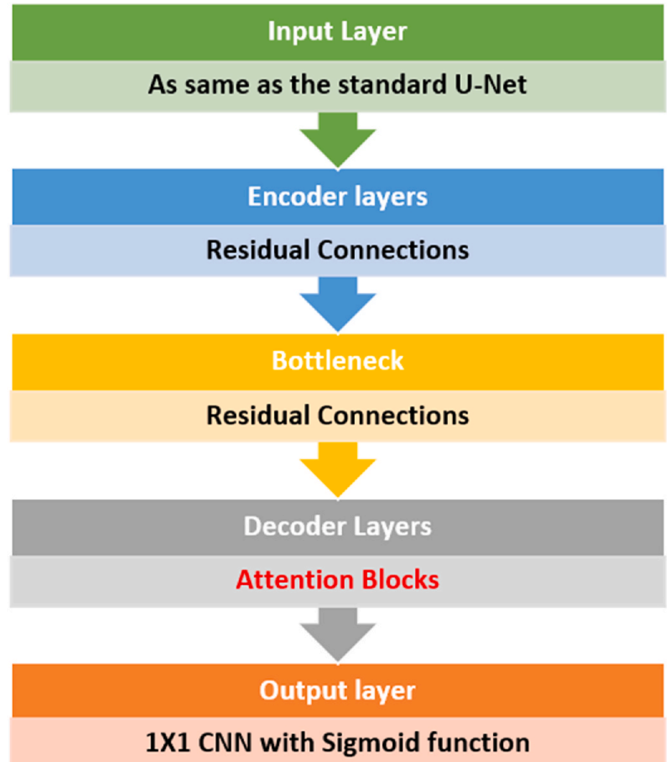

Fig. 3. The proposed model.

Algorithm 1

Start
Step1: Import the important libraries (numpy, tensorflow, cv2, os, matplotlib)
Step2: Define the modified U-net model (Encoder, Bottleneck, Decoder)
Step3: Model Compilation (optimizer, Loss, metrics)
Step4: Data Loading (images and masks)
Step5: Split to Train and Validation (80 % Train, 20 % Validation)
Step6: Model Training
Step7: Model Evaluation
Step8: Visualizing Prediction
End

Algorithm 2

Start
Step1: Import the important libraries (numpy, tensorflow, cv2, os, matplotlib)
Step2: Define the modified U-net model (Encoder, Bottleneck, Decoder)
Step3: Define attention blocks (merge them with the U-Net model)
Step4: Model Compilation (optimizer, Loss, metrics)
Step5: Data Loading (images and masks)
Step6: Split to Train and Validation (80 % Train, 20 % Validation)
Step7: Model Training
Step8: Model Evaluation
Step9: Visualizing Prediction
End

Algorithm 3

Start
Step1: Import the important libraries (numpy, tensorflow, cv2, os, matplotlib)

(continued on next page)

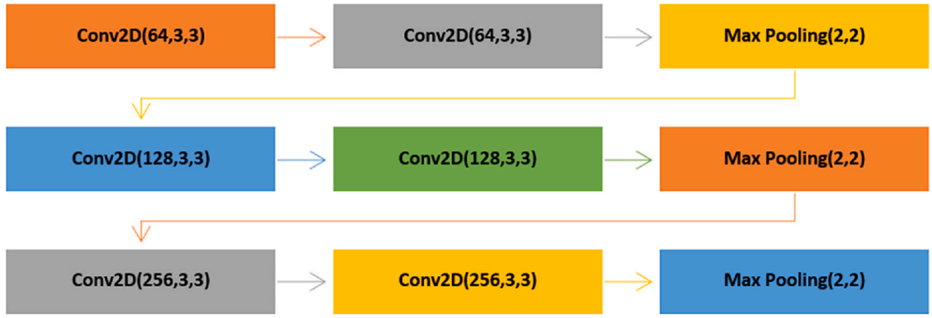

Fig. 4. The Encoder of the proposed model.

Fig. 5. The Bottleneck of the proposed model.

(*continued*)

Step2: Define the modified U-net model (Encoder, Bottleneck, Decoder)
Step3: Define Self-attention blocks (merge them with UNet model)
Step4: Model Compilation (optimizer, Loss, metrics)
Step5: Data Loading (images and masks)
Step6: Split to Train and Validation (80 % Train, 20 % Validation)
Step7: Model Training
Step8: Model Evaluation
Step9: Visualizing Prediction
End

4.1. Dataset

The dataset is available on the Kaggle official website [34] it consists of 3064 MRI Brain Tumer images, in addition to 3064 masks, the format of all images and masks is “png” with the size 512 × 512, Fig. 7 below shows samples of the dataset.

The first row in Fig. 7 above refers to the MRI section of the brain, and the second row refers to the tumor segmentation mask.

5. Results and discussions

Among the serious diseases that claim the lives of many people every

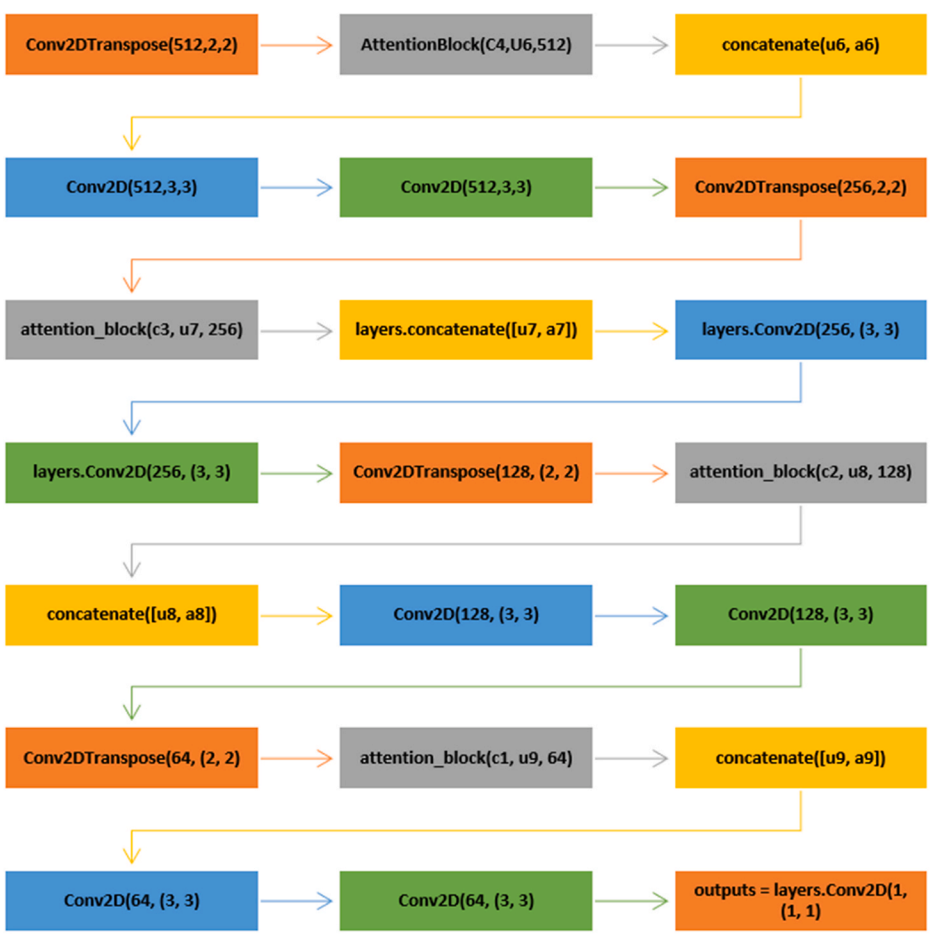

Fig. 6. The Decoder of proposed model.

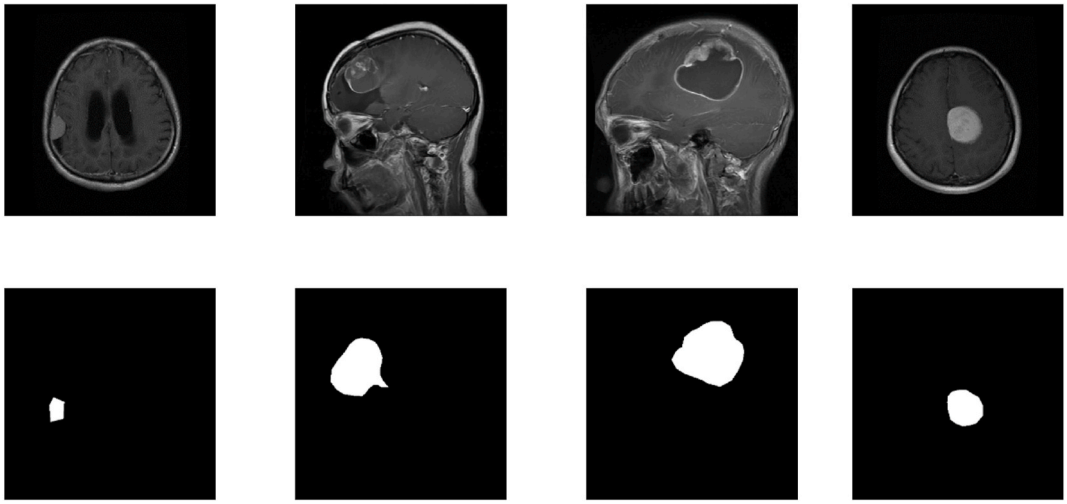

Fig. 7. Samples of Brain Tumor MRI dataset.

year is tumors, because accurately identifying tumors from pictures is difficult. To manage this disease, early detection of tumors is essential. This study proposes three models, including standard U-Net, Attention U-Net, and self-attention U-Net, to test a group of these models and improve the segmentation efficiency. The impacts of ambiguities on expected segmentations were investigated to enhance the model's understanding and regularity. Also, to improve the segmentation accuracy of the model, the estimations were employed to determine a suitable value and exclude assessments with a smaller degree of dependability. Thus, the difficulties faced by each model can be managed using similar cost criteria, optimization techniques, and other additional factors related to the configuration design.

By comparing the three models in Table 2, one can notice that the accuracy of the model training based on the attention block is higher compared with the other two models (0.9845).

Table 2
Evaluation of the model-based Standard U-Net.

Model	Epoch Number	Loss	Accuracy	Validation loss	Validation Accuracy
Standard U-Net	**Epoch #1**	0.1213	0.9831	0.0642	0.9829
	Epoch #2	0.0648	0.9831	0.0626	0.9829
	Epoch #3	0.0630	0.9832	0.0622	0.9829
	Epoch #4	0.0599	0.9835	0.0563	0.9840
	Epoch #5	0.0580	0.9840	0.0617	0.9843
Attention-based U-Net model	**Epoch #1**	0.0944	0.9821	0.0732	0.9849
	Epoch #2	0.0789	0.9845	0.0741	0.9849
	Epoch #3	0.0734	0.9845	0.0735	0.9849
	Epoch #4	0.0728	0.9845	0.0745	0.9849
	Epoch #5	0.0711	**0.9845**	0.0706	**0.9849**
Self-attention-based U-Net model	**Epoch #1**	0.1388	0.9746	0.0659	0.9823
	Epoch #2	0.0634	0.9810	0.0610	0.9823
	Epoch #3	0.0571	0.9815	**0.0482**	**0.9849**
	Epoch #4	**0.0503**	0.9829	0.0745	0.9860
	Epoch #5	0.0619	0.9827	0.0470	0.9852

By referring to Table 2 above, the using of attention based Unet model provides the best accuracy 0.9845 and validation accuracy 0.9849 rathar than the standard and using self attention, however the difference is too small, but it considered as achievement.

Figs. 8 and 9 below shows the graphical presentation of the results.

On the other hand, by looking at Figs. 10–12 one can observe the stable training along the epochs from the first epoch in the attention-based U-Net model. So, the best segmentation results were obtained when using U-Net-based attention blocks.

Below, Figs. 13–16 represent the accuracy of training vs Validation for each model by unifying the Y-Axis (Which refer to accuracy) and X-Axis (Which refer to epoch).

As denoted in Table 3, the attention-based U-net model is steadier because the attention blocks use dropout unnecessary layers. Also, the processing time (9894s) in the self-attention-based U-Net model is less than the standard U-Net model (35525s) and attention-based U-Net (34557s) for removing some layers, so, some extra calculations are omitted.

On the other hand, the other evaluation metrics of the model using attention blocks, with different threshold value of binary classification are listed in the following Table 4.

By reffering to Table 4, Dice score is decreasing (96.54 % → 86.36 %) when increasing the threshold (0.3 → 0.6), which meanse lots of pixels classified as positive leading to better overlap. Jaccard Index is as same as Dice, but classify the false negative more, (0.3 → 82 %) meanse more positives classified, (0.6 → 70.61 %) fewer true positives classified.

In addition, the loss function goes towards the stable in the proposed model than the standard model, by looking at Figs. 8–10, the validation loss decreases along the epochs in the proposed model, while the Validation loss function increases along the last epochs, this is because of the huge features gain from the standard which causes overfitting. While the outcomes are decent with the current techniques, there is still an opportunity for improvement. Therefore, incorrect assessments can be reduced by changing the training procedure to distribute the label correctly, which can continue to modify the results with improved accuracy.

The proposed model has depended on an attention-based U-Net which improves the best accuracy by comparing with the past work that also depended on an adapted U-Net. Table 5 below shows the accuracy of each past work's architecture.

6. Policy suggestions

In medical applications, the accuracy plays a vital role for making

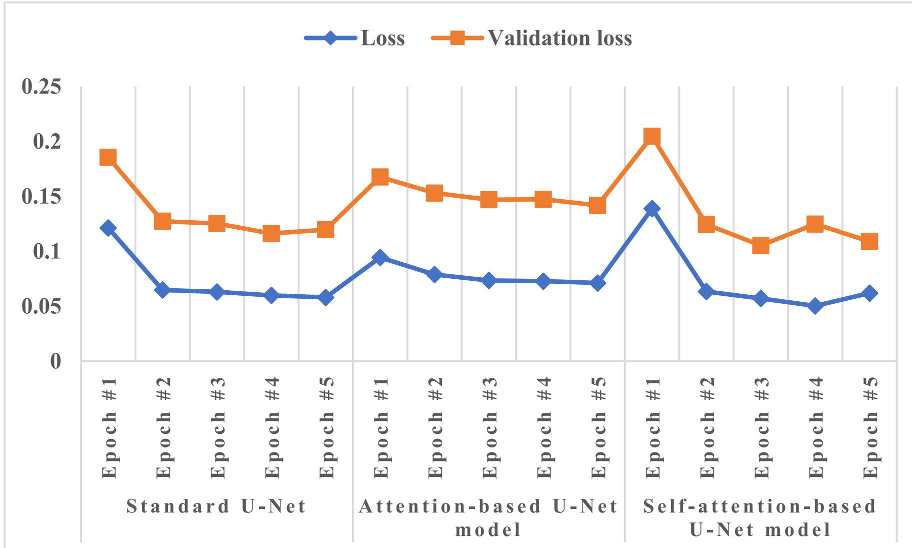

Fig. 8. Graphical presentation of Loss and Val Loss along epochs.

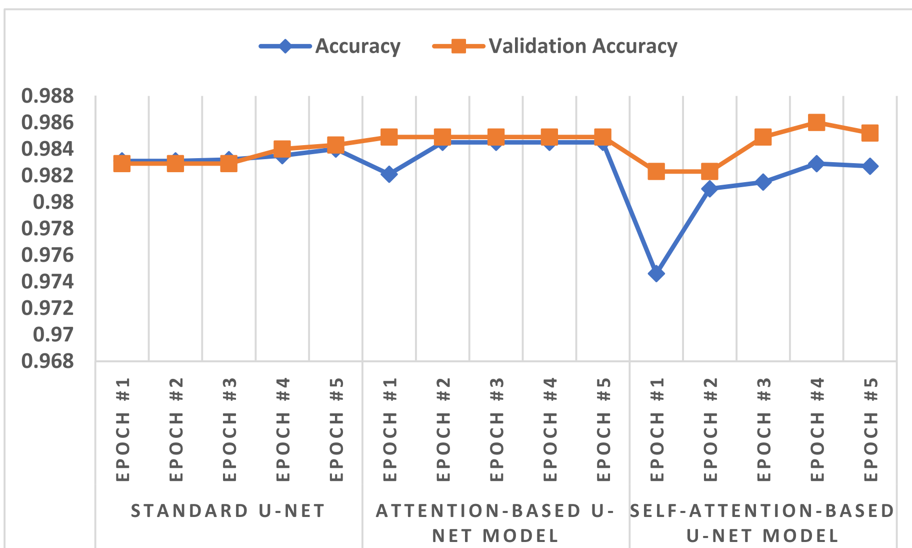

Fig. 9. Graphical presentation of Accuracy and val accuracy along epochs.

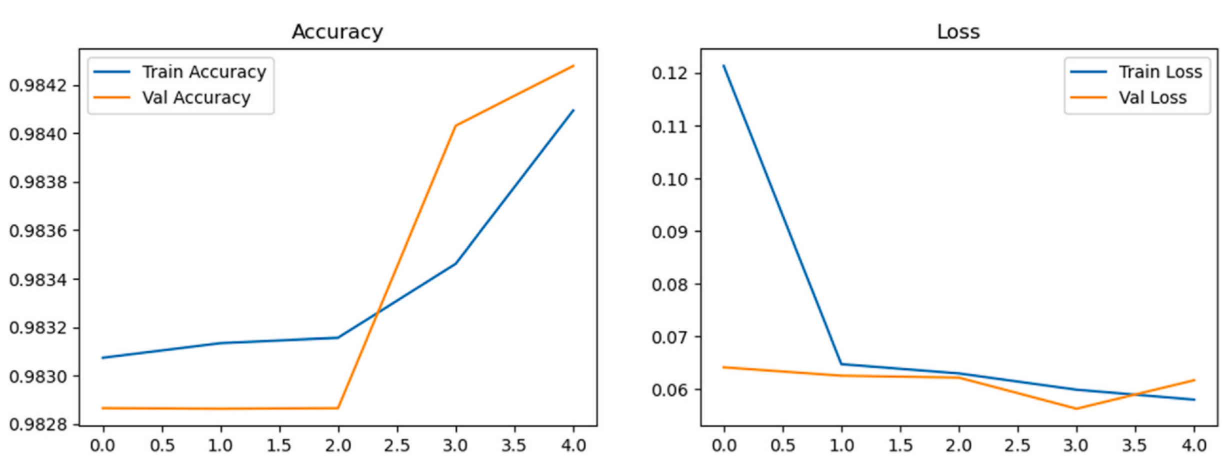

Fig. 10. Model-based Standard U-Net metrics.

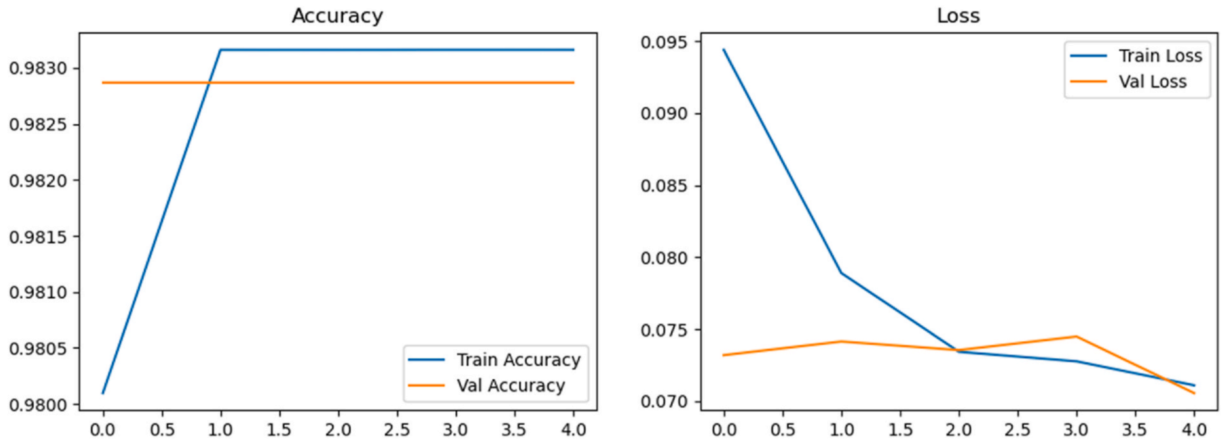

Fig. 11. Attention-based U-Net metrics.

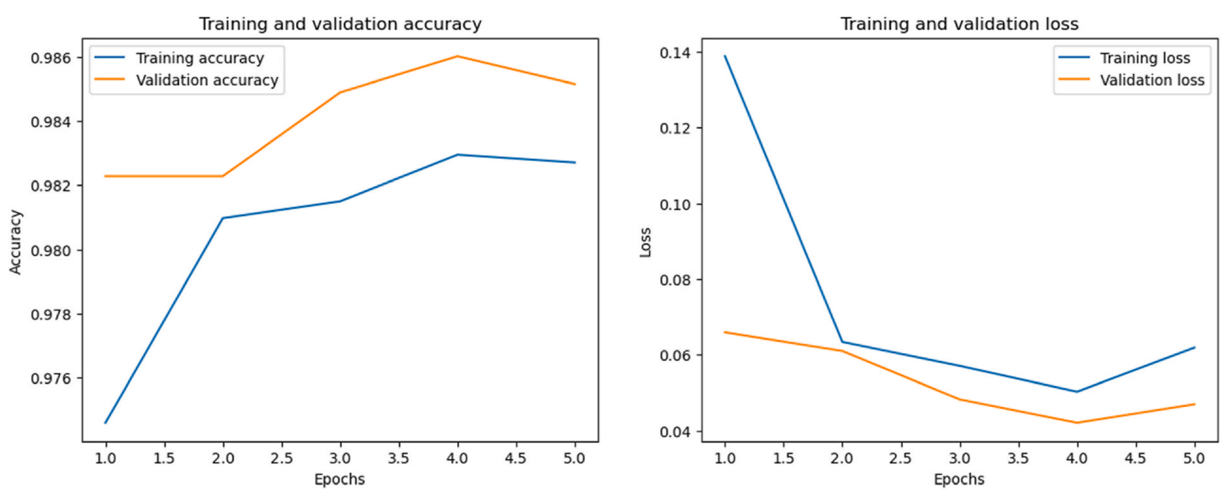

Fig. 12. Self attention based U-Net metrics.

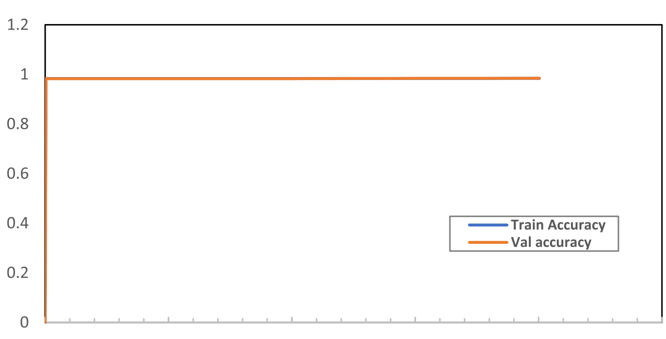

Fig. 13. The Accuracy of training vs validation for Standard model.

Fig. 14. The Accuracy of training vs validation for model used attention blocks.

accurate diagnosis, for this reason, However, the classical segmentation models are used with accepted accuracy, so this work focuses on proposing an enhanced segmentation model with more accuracy and less loss than the classical models.

7. Conclusions

In real-time applications such as diagnostic procedures and treatment, segmentation of medical images is a challenging problem. Segmentation algorithms can correctly identify different tissues surrounding the tumor site and its boundaries; additional creative work is needed to increase computing performance. Therefore, medical data needs to be segmented precisely which creates an ideal deep-learning model. U-Net design has been the foundation of most current studies rather than using alternative deep models. This work covers the deep structure of U-Net and uses both the encoder and decoder architecture. It aims to improve the results by better representing the labels on the network and improving the visualization of globally and locally relevant data in the dataset. Different models were executed using MRI image datasets from Kaggle to evaluate the model segmentation. A variety of U-Net designs were trained, including standard U-Net, attention U-Net, and self-attention U-Net. In addition, to show if the U-Net design

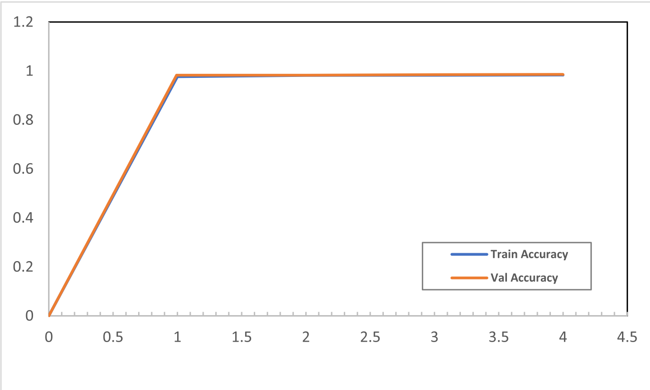

Fig. 15. The Accuracy of training vs validation for model used Self attention.

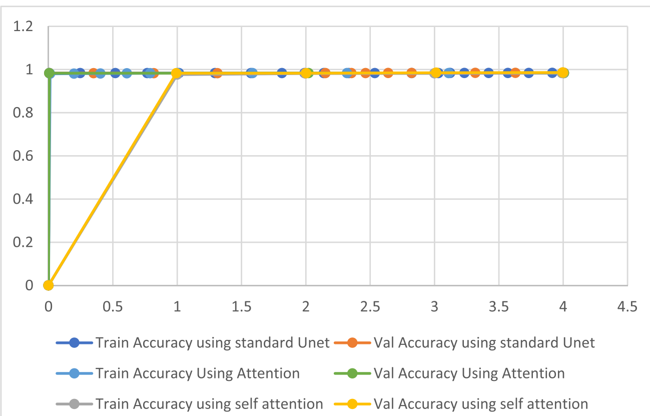

Fig. 16. The comparisons for Accuracy of training vs validation for all models.

Table 3
Processing time in (sec).

Epoch Number	Standard U-Net	Attention	Self-Attention
Epoch #1	6746s	7020s	1916s
Epoch #2	6633s	6863s	1827s
Epoch #3	6609s	6873s	1848s
Epoch #4	8848s	6946s	1904s
Epoch #5	6689s	6855s	2399s
Total	35525s	34557s	**9894s**

operates more effectively than others, we included a comparison of several models from other studies. We have shown that the effectiveness of the suggested models outperforms several of the most advanced methods. Four key metrics are used to evaluate the models that achieved the best results, especially attention U-Net and self-attention U-Net models: validation accuracy of 98.49 %, loss function of 5.03 %, accuracy of 98.45 %, and validation loss function of 4.82 %.

These models focus on the important regions, which can improve segmentation quality, especially for ambiguous structures. This is accomplished using a feature map and examining the function of every layer individually. Therefore, a thorough investigation demonstrates how well the framework can be adjusted. The main limitations of this study are high computing resource requirements because of their intricate architecture and high processing demands. Furthermore limited access to data. Moreover, there is limited data access and large training sample requirements to train the networks using 3D images. In the future, this network architecture can be modified to provide significant discovery of additional resources. These models will help radiologists diagnose problems rather than replace them. Therefore, medical

Table 4
Model metrics.

Threshold	Dice Coefficient	Jaccard Index	Sensetivity	Specifity
0.3	96.54 %	82 %	93.20 %	99.68 %
0.5	90 %	74 %	80.55 %	99.97 %
0.6	86.36 %	70.61 %	79.61 %	99.95 %

Where:
Threshold: It means how the model can classify a segment as positive or negative; in many models, the output ranges between 0 and 1, and the default threshold value is 0.5. But specific models need to use a lower threshold of 0.3, the other need to use higher threshold of 0.6, where fewer prediction is considered ad positive.
Dice Coefficient: It is used to find the similarity between two samples; in segmentation, it is used to find the overlap between predicted and truth foreground.
Jaccard Index: It is like Dice coefficient, but it use different behaviour in calculation, it finds how the predicted foreground close to thr real.
Sensetivity: It identifies the actual positive in the model. It is also called the true positive rate.
Specificity: It identifies the actual Negative in the model. It is also called the true negative rate.

Table 5
Comparison with past work.

Work	Architecture	Accuracy
[1]	3D Unet	0.9500
[13]	BU-Net	0.8920
[14]	M-Unet	0.9602
[15]	M-Unet	0.8919
[16]	3D+2D Unet	0.8701
[17]	3D Unet	0.8800
[20]	M-Unet	0.9500
[21]	Double Unet	0.9494
[22]	M-Unet	0.9148
[23]	M-Unet	0.9309
[24]	3D Unet	0.9300
[25]	mRes-Unet	0.9289
Proposed	Adapted Unet + Self Attention Blocks	**0.9845**

operations will be made more efficient when radiologists and advanced models work together.

CRediT authorship contribution statement

Farah F. Alkhalid: Supervision, Project administration, Methodology, Data curation, Conceptualization. **Nibras Z. Salih:** Writing – review & editing, Writing – original draft, Validation.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

Ethics approval

Not applicable.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T. Henry, et al., Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: a BraTS 2020 challenge solution [Online]. Available, http://arxiv.org/abs/2011.01045, Oct. 2020.

[2] R. Branko, Maričić and Aleksandar Đorđević, "Strategic_market_segmentation,", Marketing 46 (6) (2015) 243–251, https://doi.org/10.5937/markt1504243M.

[3] Muna Khalaf, Ban N. Dhannoon, Deep learning based on attention in semantic segmentation: an introductory survey, Iraqi .J. Comput. Communication, Control Syst. Eng. (Mar. 2023) 104–114, https://doi.org/10.33103/uot.ijccce.23.1.9.

[4] Z.N. Alabdali, F.F. Alkalid, Analysis of particle size in composite materials using image processing, Int. J. Eng., Transactions A: Basics 37 (4) (Apr. 2024) 579–587, https://doi.org/10.5829/ije.2024.37.04a.01.

[5] F.F. Alkhalid, A.M. Hasan, A.A. Alhamady, Improving radiographic image contrast using multi layers of histogram equalization technique, IAES Int. J. Artif. Intell. 10 (1) (2021) 151–156, https://doi.org/10.11591/ijai.v10.i1.pp151-156.

[6] K.K.D. Ramesh, G. Kiran Kumar, K. Swapna, D. Datta, S. Suman Rajest, A review of medical image segmentation algorithms, EAI Endorsed Trans. Pervasive Health Technol. 7 (27) (2021), https://doi.org/10.4108/eai.12-4-2021.169184.

[7] W. Rawat, Z. Wang, Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review, MIT Press Journals, Sep. 01, 2017, https://doi.org/10.1162/NECO_a_00990.

[8] S. Hossain, et al., Automated breast tumor ultrasound image segmentation with hybrid UNet and classification using fine-tuned CNN model, Heliyon 9 (11) (Nov. 2023) e21369, https://doi.org/10.1016/j.heliyon.2023.e21369.

[9] B. Liang, C. Tang, W. Zhang, M. Xu, T. Wu, N-Net: an UNet architecture with dual encoder for medical image segmentation, Signal Image Video Process 17 (6) (Sep. 2023) 3073–3081, https://doi.org/10.1007/s11760-023-02528-9.

[10] Bushra T. Hashim, Hadi T. Ziboon, Sinan M. Abdulsatar, Enhancement of a two thresholds sensing system by segmentation of confused regions, Iraqi .J. Comput. Communication, Control Syst. Engin (Jun. 2023) 83–94, https://doi.org/10.33103/uot.ijccce.23.2.7.

[11] A.M.G. Allah, A.M. Sarhan, N.M. Elshennawy, Edge U-Net: brain tumor segmentation using MRI based on deep U-Net model with boundary information, Expert Syst. Appl. 213 (2023) 118833, https://doi.org/10.1016/j.eswa.2022.118833.

[12] N.Z. Salih, F.F. Alkhalid, The novel of using transfer learning approach for seatbelts automated surveillance, International J. Transp. Development Integr. 8 (1) (Mar. 2024) 141–147, https://doi.org/10.18280/ijtdi.080113.

[13] M.U. Rehman, S. Cho, J.H. Kim, K.T. Chong, Bu-net: brain tumor segmentation using modified u-net architecture, Electronics (Switzerland) 9 (12) (Dec. 2020) 1–12, https://doi.org/10.3390/electronics9122203.

[14] A. Soni, R. Koner, V.G.K. Villuri, M-unet: modified u-net segmentation framework with satellite imagery, in: Advances in Intelligent Systems and Computing, Springer, 2020, pp. 47–59, https://doi.org/10.1007/978-981-15-2188-1_4.

[15] M. Aghalari, A. Aghagolzadeh, M. Ezoji, Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks, Biomed. Signal Process Control 69 (2021) 102841.

[16] Y. Zhang, et al., Brain tumor segmentation from multi-modal MR images via ensembling UNets, Frontiers in Radiology 1 (2021), https://doi.org/10.3389/fradi.2021.704888.

[17] P. Agrawal, N. Katal, N. Hooda, Segmentation and classification of brain tumor using 3D-UNet deep neural networks, International. J. Cognitive Computing Eng. 3 (Jun. 2022) 199–210, https://doi.org/10.1016/j.ijcce.2022.11.001.

[18] D. Maji, P. Sigedar, M. Singh, Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors, Biomed. Signal Process Control 71 (Jan) (2022), https://doi.org/10.1016/j.bspc.2021.103077.

[19] R. Raza, U. Ijaz Bajwa, Y. Mehmood, M. Waqas Anwar, M. Hassan Jamal, dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI, Biomed. Signal Process Control 79 (Jan. 2023), https://doi.org/10.1016/j.bspc.2022.103861.

[20] R. Yousef, et al., U-Net-Based models towards optimal MR brain image segmentation, Diagnostics 13 (9) (May 2023), https://doi.org/10.3390/diagnostics13091624.

[21] S. Deep Deb, R. Kumar Jha, Modified double U-net architecture for medical image segmentation [Online]. Available, https://www.researchgate.net/publication/369927256, Apr. 2023.

[22] M. Arhami, F.F. Yanuar Rudi, A semantic segmentation of nucleus and cytoplasm in pap-smear images using modified U-net architecture 16 (2) (2024) 273–288, https://doi.org/10.20895/INFOTEL.V16I2.1006.

[23] B. Amaludin, S. Kadry, F.F. Ting, D. Taniar, Toward more accurate diagnosis of multiple sclerosis: automated lesion segmentation in brain magnetic resonance image using modified U-Net model, Int. J. Imag. Syst. Technol. 34 (1) (Jan. 2024), https://doi.org/10.1002/ima.22941.

[24] B. Parmar, M. Parikh, Pg 1 Estimation of uncertainty in Brain Tumor segmentation using modified multistage 3D-UNet on multimodal MRI images [Online]. Available, http://pubs.thesciencein.org/jist.

[25] P. Li, Z. Li, Z. Wang, C. Li, M. Wang, mResU-Net: Multi-Scale Residual U-Net-Based Brain Tumor Segmentation from Multimodal MRI, Springer Science and Business Media Deutschland GmbH, Mar. 01, 2024, https://doi.org/10.1007/s11517-023-02965-1.

[26] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation [Online]. Available, http://arxiv.org/abs/1505.04597, May 2015.

[27] J. Ruan, S. Xiang, VM-UNet: vision mamba UNet for medical image segmentation [Online]. Available, http://arxiv.org/abs/2402.02491, Feb. 2024.

[28] K. Cao, X. Zhang, An improved Res-UNet model for tree species classification using airborne high-resolution images, Remote Sens (Basel) 12 (7) (Apr. 2020), https://doi.org/10.3390/rs12071128.

[29] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, P.A. Heng, H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes [Online]. Available, http://arxiv.org/abs/1709.07330, Sep. 2017.

[30] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate [Online]. Available, http://arxiv.org/abs/1409.0473, Sep. 2014.

[31] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, CBAM: Convolutional Block Attention Module, 2018.

[32] T. Shen, T. Zhou, G. Long, J. Jiang, C. Zhang, Bi-directional block self-attention for fast and memory-efficient sequence modeling [Online]. Available, http://arxiv.org/abs/1804.00857, Apr. 2018.

[33] J.-B. Cordonnier, A. Loukas, M. Jaggi, On the relationship between self-attention and convolutional layers [Online]. Available, http://arxiv.org/abs/1911.03584, Nov. 2019.

[34] Brain Tumor Dataset, Kaggle, https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset.