



Brain tumor image segmentation via asymmetric/symmetric UNet based on two-pathway-residual blocks

Motahareh Aghalari, Ali Aghagolzadeh^{*}, Mehdi Ezoji

Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

ARTICLE INFO

Keywords:

Brain tumors
Automatic segmentation
Glioma
UNet
Two-pathway-residual block

ABSTRACT

Early diagnosis and selection of an appropriate treatment method will increase the survival of cancer patients. Accurate and reliable brain tumor segmentation is an important component in tumor diagnosis and treatment planning. Glioma is one of the hardest brain tumors in diagnosis because of its irregular shape and blurred borders. Automatic segmentation of glioma brain tumors is a challenging problem due to significant variations in their structure. In this paper, improved UNet-based architectures are presented for automatic segmentation of brain tumors from MRI images. Specifically, we designed the strong Two-Pathway-Residual blocks for UNet structure and proposed three models. Our proposed models' architectures exploit both local features as well as more global features simultaneously. Furthermore, different from the original UNet, our proposed architectures have fewer parameters. The proposed models were evaluated on the BRATS'2018 database and given good results while the calculation cost was lower than the other methods. DCS, sensitivity, and PPV criteria values used for the segmentation results of the best-proposed model are 89.76%, 89.19%, and 90.65%, respectively.

1. Introduction

There are more than 120 types of brain tumors, that glioma brain tumors are the most common among adults [1]. Gliomas are caused by the uncontrolled and unusual spread of glial cells in the brain. These cells are located between neurons and protect them [2,3]. World Health Organization (WHO) classifies gliomas into two groups according to their grade: Low-Grade Gliomas (LGG) and High-Grade Gliomas (HGG) [4]. The median survival time (MST) of HGG patients is less than two years [5,6] and the MST of patients with glioblastoma, the most aggressive tumors, is only 4.9 months [7]. The course of disease progression and the MST for some patients can be controlled/improved by timely diagnosis and selection of the appropriate treatment [8]. Brain tumor segmentation is the process of separating the tumor from natural brain tissues; in clinical routine, it provides useful information about tumor characteristics including type, location, size, and shape for diagnosis and treatment planning [9–10].

Magnetic resonance imaging (MRI) technology is the most common imaging [11] test to diagnose a brain tumor as it has harmless, non-invasive, and high-resolution features [12]. Also, MRI can obtain multimodal images of the brain's internal tissues [13]. The most common MRI modalities are T1-weighted (T1), contrast-enhanced T1-

weighted (T1ce), T2-weighted (T2), and Flair scans. In each mode, a specific part of the brain tissue is better represented. For example, T1 images are good at isolating healthy tissues in the brain, while T1ce helps to separate tumor boundaries that appear brighter due to the contrast agent [14,15]. Despite the many benefits of MRI, segmenting a brain tumor manually is tedious, time-consuming, and error-prone [12,14,16]. Therefore, fully automatic and accurate methods are required. So far, many automated systems have been designed, but accurate glioma brain tumor segmentation is still a challenge. This is for some reasons: First, gliomas can appear anywhere on the brain with a variety of shapes, sizes, and appearances [17]. Second, diagnosing gliomas is difficult because of the fuzzy boundaries with healthy brain tissue [18,19], and the other important reason is the lack of a fixed definition of grayscale values in the same MRI mode. Four types of sequences for two patients are shown in Fig. 1. As seen in Fig. 1, the structures of tumors are different. Also in Flair mode, healthy cells are seen darker in the patient's image B. Depending on the specifications of the MRI machine (1.5, 3, or 7 T) and the acquisition protocol [20], the grayscale values of the brain tissue vary from one patient to another patient.

^{*} Corresponding author.

E-mail addresses: motahareh.aghalari@nit.ac.ir (M. Aghalari), aghagol@nit.ac.ir (A. Aghagolzadeh), m.ezoji@nit.ac.ir (M. Ezoji).

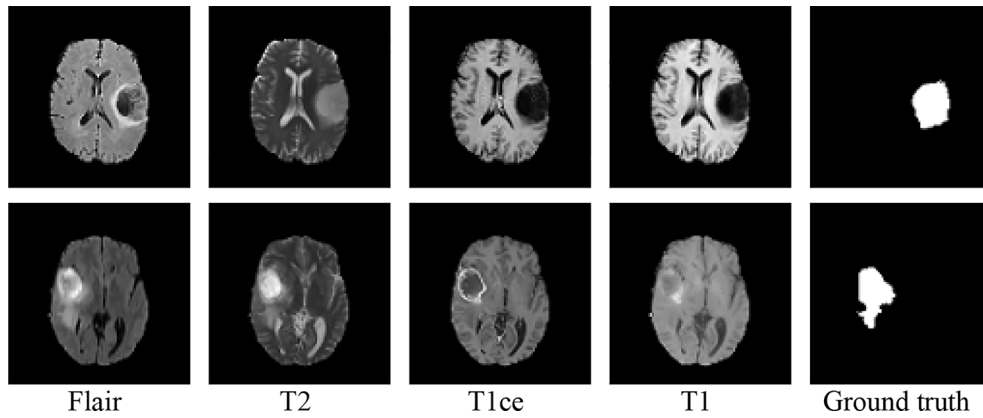


Fig. 1. Examples of a brain tumor in two patients A (top row) and B (bottom row) with 4 MRI modalities (Flair, T2, T1ce, T1) and physician' delineation of tumor region.

2. Related work

In the early works, segmentation algorithms are often based on thresholding, mathematical morphology, and region-growing. Ilhan et al. [21] used threshold-based segmentation to diagnose brain cancer. The accuracy of this method is reported to be 96%. Gupta et al. [16] proposed a different segmentation method for two sets of images. The first set, including T1 and T1c images, was segmented by adaptive thresholding while in the second set (Flair and T2 images), the tumor area was identified by Canny edge detection filter. Due to the complexity of the brain structure and the blurred tumor boundaries, threshold-based approaches failed to achieve high accuracy. The combination of region growing method and morphology was introduced by Rahima et al. [22] for semi-automatic segmentation of low-grade gliomas. Devkota et al. [23] attempted to segment 19 images of glioma with a mathematical morphology operator. Although the velocity and accuracy of morphological operation are high, increasing the number of samples may distort its good results. Li et al. [14] identified the tumor by combining spatial fuzzy c-mean (SFCM) and region growing. The region growing methods required to initialize the seed points. Mostly, their segmentations were not enough accurate when the initializations were unsuitable. These poor initial conditions limit their success for a larger number of data sets. Therefore, more sophisticated methods are needed.

Various convolutional neural networks (CNNs) have been designed for brain tumor segmentation. The idea behind these networks is to learn the complex hierarchical features directly from domain data [11]. Pereira et al. [24] used an automatic brain tumor segmentation system based on 2D-CNN and obtains very excellent performance in BRATS 2013 challenge. Havaei et al. [11] introduced two-pathway architecture for effective learning of tumor details and contextual information. Hussain et al. [2] presented Nexus architectures for the segmentation of gliomas. In the Nexus model, the output of a basic CNN is treated as an additional input for a subsequent CNN. All of these CNNs are based on a single-label prediction scheme. Single-label prediction architecture takes image patches (local areas of the image) as input and categorizes the patch's central voxel into a tumor or non-tumor class. The segmented image is achieved by predicting the central voxel class of the input image patches. Therefore, single-label prediction networks are very slow in the conclusion stage. Subsequent advances in CNNs were in the introduction of dense prediction networks. These are superior to the traditional CNNs in predicting the labels of voxels within the input patch simultaneously. Fully convolutional network (FCN), a typical example of dense prediction network, was first used to segment brain tumors by Shen et al. [17]. FCN replaces the full-connected layer in the traditional CNN with a convolutional layer and contains up-sampling layers to recover its original size. Thus, it can achieve the desired size and provide fast results. Zhao et al. [19] used a combination of FCN with conditional

random fields (CRF) to effectively segment gliomas. UNet and its variants have recently become very popular in medical image segmentation tasks. In order to obtain more accurate segmentation, UNet gradually introduces feature maps by concatenation from its down-sampling layers to its up-sampling layers of the same resolution [18]. Caver et al. [25] used UNet in the Brain Tumor Segmentation Challenge 2018 and achieved 87.8% for dice similarity coefficient (DSC). The idea of using two parallel UNet with asymmetric residual blocks was introduced by Abd-Ellah et al. [13]. By extracting local and global features in parallel paths, segmentation performance is improved. Wang et al. [20], proposed the WRN-PPNet network by combining the global priors under different pyramid scales with the local feature. Improving DSC to 91% indicates the power of the network. 3D patch-based UNet was presented by Cabezas et al. [26], which takes full advantage of 3D information on MRI scans. Hu et al. [27] proposed the 3D-dilated-UNet and reported 86% for DSC on the BRATS'2018 data set. For effective learning of the intra-slice and inter-slice representations in deep networks, a separable 3D UNet was presented by Chen et al. [28]. 3D networks can directly process 3D MRI data. However, they require high GPU memory and increase computational cost.

The segmentation algorithm may mistakenly detect small areas of the tumor as healthy tissue, and vice versa. Hence a post-processing step has been used in some studies to refine the result of segmentation. For example, Pereira et al. [24] removed small tumor areas obtained from CNN segmentation by applying the T_{VOL} threshold. Havaei et al. [11] proposed a simple method based on the connected components to remove flat blobs. To get the final boundaries of glioma, Li et al. [14] designed two-step post-processing: initially, pseudo glioma areas were removed using the region merging algorithm based on MDL criteria. Then, the improved distance regularized level set evolution method was used to eliminate small holes and correct tumor boundaries. In addition to these methods, simple operators such as mathematical morphology [2] and median filter [21] have also been adopted for post-processing.

In this paper, we propose an automatic method for segmenting brain tumor based on skull-stripped MRI images. The main contributions of our paper are: 1) we use a modified version of the basic UNet to segment the brain tumor images, 2) then, we present a combination of a two-pathway architecture with the residual block as two-pathway-residual (TPR) architecture, 3) and finally, we propose three novel TPR-based UNet models to improve segmentation performance, which include TPR blocks. Our models achieved good results in glioma brain tumor segmentation, while computing costs were lower than the competitors. Also, unlike other methods, the proposed algorithm does not require any post-processing operator.

The rest of this paper is organized as follows: Section 3 introduces the proposed methodology. Then, the experimental setup is explained in section 4, and the experiment results are given in section 5. Finally, a

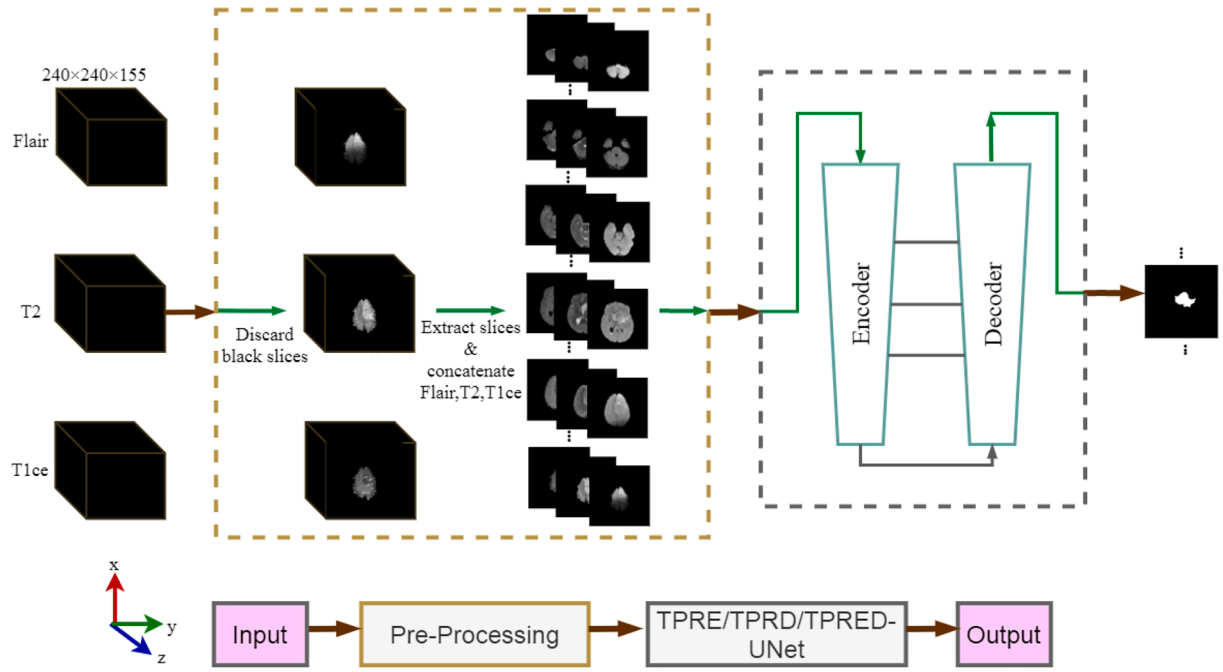


Fig. 2. Graphical abstract of the proposed methodology.

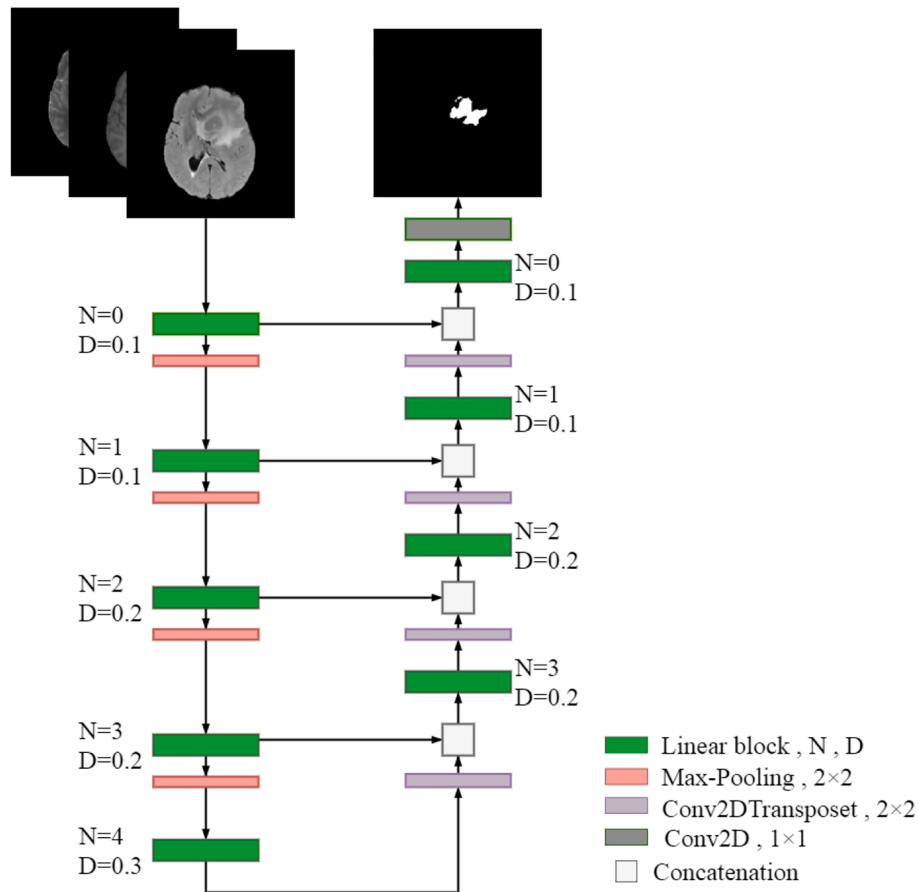


Fig. 3. The model architecture of the modified UNet [30]. 'D' refers to the value of the dropout rate for each layer. 'N' refers to the Nth power of two in (16×2^N) expression. The expression (16×2^N) equals the number of feature maps of each convolution layer.

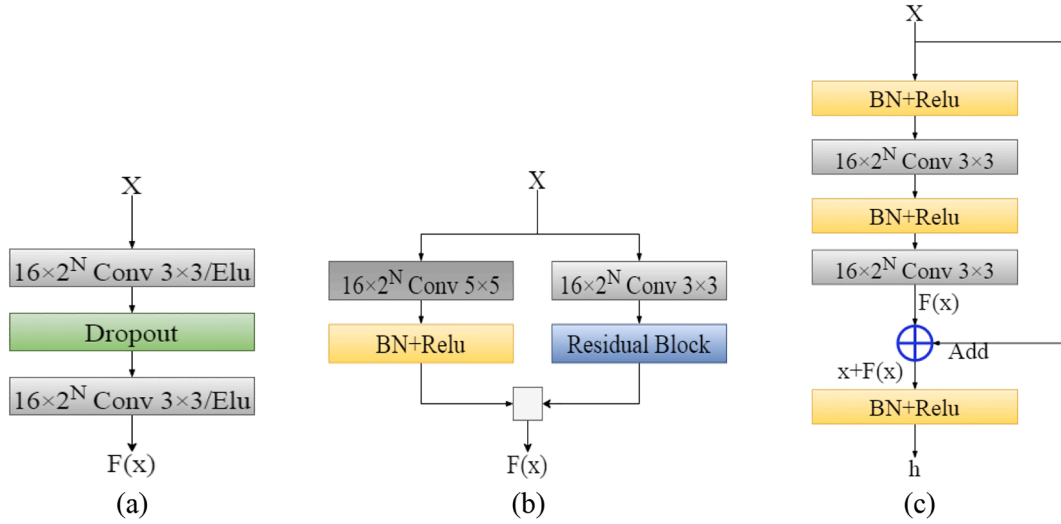


Fig. 4. The Model architecture of (a) the linear block [30], (b) the proposed two-pathway-residual (TPR) block, and (c) the residual block [31].

conclusion is drawn in section 6.

3. Proposed methodology

The proposed methodology explains the performance of deep neural networks in brain tumor segmentation task. According to Fig. 2, input images are improved during a two-step pre-processing, and then UNet-based architectures are used to segment the brain tumor images. Among the MRI sequences, the resolution of the tumor tissue in the T1 image is lower than the others. Therefore, only three types of sequences (Flair, T2, T1ce) have been used for determining the tumor area.

3.1. Pre-processing

3D MRI scans consist of 2D slices from the axial, coronal, and sagittal views. In our proposed method, only 2D slices from the axial view are used as input images as the 2D axial slices have discriminative information to differentiate tumor tissues [17]. Some slices at the beginning and the end of the MRI images are completely black (no brain image), which is not helpful. so black slices are discarded in the first step of the pre-processing. Then, since the intensity values across MRI slices are very large [2], image normalization is performed. If 'x' is the original image with a mean of ' μ ' and a standard deviation of ' σ ', its normalized image (' x_n ') is obtained according to Eq. (1). The normalized image has a mean of zero and a standard deviation very close to one.

$$x_n = \frac{x - \mu}{\sigma}. \quad (1)$$

3.2. Modified UNet architecture

UNet is an evolution of the traditional CNN, that first designed in 2015 to process biomedical images [29]. The modified version of the basic UNet (modified UNet) [30] is shown in Fig. 3. As seen in Fig. 3, three slices are used in modified UNet architecture with the same order in three MRI modalities respectively as inputs. These inputs are processed by a down-sampling/encoder path with down-sampling operations and an up-sampling/decoder path with up-sampling operations. To improve the accuracy of segmentation, the feature maps of the two paths are concatenated through a skip connection. There are the same linear blocks on both UNet paths; so UNet is a symmetric network. The model architecture of the linear block [30] and our proposed two-pathway-residual block and the residual block given in [31] are shown in Fig. 4. According to Fig. 4(a), the linear block receives input 'x' and generates feature maps as F(x). This block has two convolutional layers

with 16×2^N filters of size 3×3 , and stride of 1×1 . To maintain the output dimensions as in input, zero padding is applied to each convolutional layer. The activation function behind the convolutional layer is an exponential linear unit (Elu), and defined as:

$$y = \max(x, 0) + \alpha(e^x - 1)\min(0, x) \quad (2)$$

where 'x' is the output of the convolutional layer, 'y' is the output of the activation function and ' α ' is a pre-defined constant ($\alpha = 1.0$).

Besides, to prevent overfitting of the network, a dropout layer with a dropout rate of 'D' is placed between two convolutional layers. The down-sampling path has five linear blocks, which increase the value of 'N' from 0 to 4. For the down-sampling, max-pooling with stride 2×2 is applied to the end of every block except the last block; so the size of feature maps decreases from 240×240 to 15×15 . In the up-sampling path, every block starts with a transposed convolution layer with a filter size of 2×2 and stride of 2×2 , which doubles the size of feature maps. Zero padding is also applied to each transposed convolution layer. The feature maps produced by the transposed convolution layers are concatenated with those generated by linear blocks of the down-sampling path with the same resolution and are passed through a linear block with the same specifications. After the last linear block in the up-sampling path, a (1×1) convolutional layer with the sigmoid activation is used to obtain the tumor area.

In general, UNet performs well in segmenting brain tumors, but it also has two drawbacks: 1) usually, the size of brain tumors is smaller than the entire image [17]. As the number of linear blocks in the UNet increases, some small tumors in subsequent blocks may disappear because after each block (except the last one) a max-pooling layer is used to reduce the size of the feature maps. Therefore, some information is lost and subsequent blocks have less capability of detecting tumors. 2) the size of all filters in the linear block is small; so linear blocks learn the local features more than the global features. These problems limit UNet's ability to detect a variety of tumors. In this study, we proposed a strong block based on the combination of two-pathway and residual block called two-pathway-residual (TPR) for the UNet structure. The new UNet structures have a great ability to detect a variety of tumors of different sizes and shapes.

3.3. TPR-based UNet

To learn more about the global features, we designed a two-pathway-residual (TPR) block for UNet structure, inspired by [24]. This block consists of two streams: a local path with a (3×3) convolutional layer and a residual block, to extract the local features of the image, and a

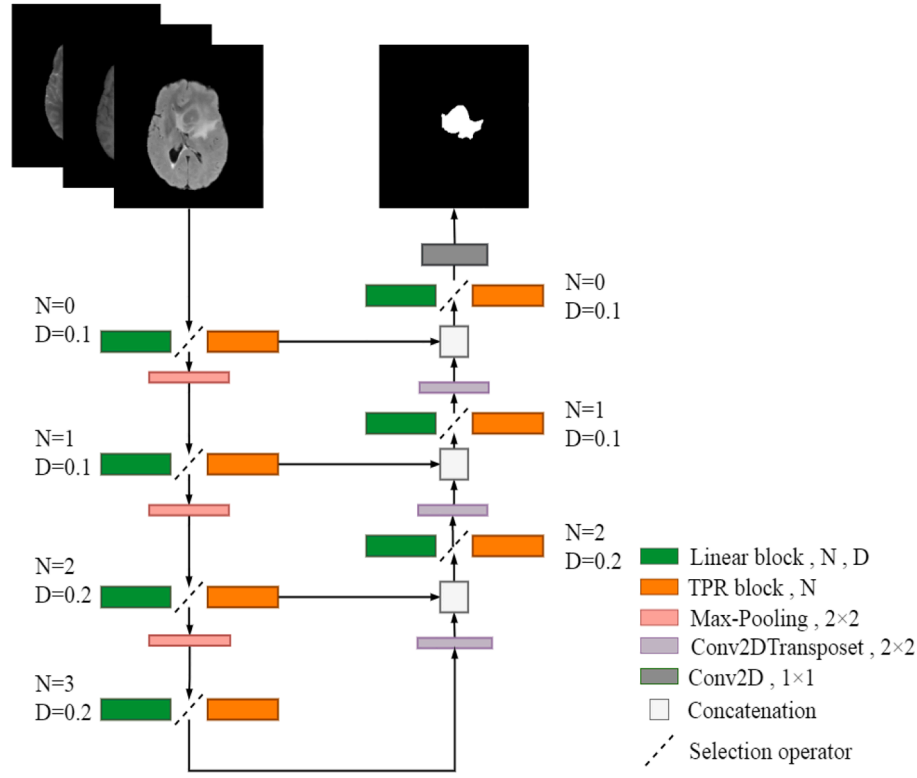


Fig. 5. The general scheme of the proposed TPR-based UNet architecture. 'D' refers to the value of the dropout rate for each layer. 'N' refers to the Nth power of two in (16×2^N) expression. The expression (16×2^N) equals the number of feature maps of each convolution layer.

global path with a (5×5) convolutional layer along with the batch normalization (BN) layer for extracting the global features. The feature maps of the two paths are transferred to the next layer after concatenation. The proposed TPR block along with its details is shown in Fig. 4 (b). After each BN layer, a rectified linear unit (Relu) operation is used as an activation function to ensure non-linear mapping [17]. Eq. (3) refers to the Relu operation:

$$y = \max(x, 0), \quad (3)$$

where 'x' is the output of the BN layer and 'y' is the output of the activation function.

The TPR block can process information from two parallel paths, which helps to more accurately detect the tumor area. Besides, placing a residual block in this two-pathway architecture can increase its performance. In most cases, it has been observed that as the depth of convolutional neural networks increases (the number of layers increases), the accuracy of their output estimation improves. However, the gradient in the deep layers may vanish, which will result in degradation problem [20]. To solve this problem, residual blocks were introduced by Mostafiz et al. [31]. The structure of the residual block [31] is illustrated in Fig. 4

(c). This block contains a stacked set of different layers. Its input is normalized by a BN layer, and at that point, the Relu activation is connected, followed by a (3×3) convolutional layer. These layers are then repeated. Then, the input is added by a direct connection to the output of the last convolutional layer. Using this direct connection, the problem of gradient disappearance will be solved, because the input 'x' is available and transferred to the next layers. Also, a BN layer is used along with the Relu activation function after the sum operation. The relationship between the input and the output of the residual block can be defined by the following formula:

$$h = B(x + F(x)), \quad (4)$$

where 'x' and 'h' are the input and the output of the residual block, respectively. $F(\cdot)$ is the output of the convolutional layer and $B(\cdot)$ is the output of the BN layer along with the Relu activation function.

In this study, we use the proposed TPR blocks to present different UNet schemes. The general scheme of the proposed UNet architectures is shown in Fig. 5. As it is shown, by changing the type of block from the linear to the TPR, you can get a variety of UNet schemes. The proposed architectures are as follows:

Table 1

Summary of the different UNet-based architectures. EP = Encoder path of UNet, DP = Decoder path of UNet.

| Model | Block (type, number) | Conv2D layers | Conv2D kernels | Dropout layers | BN layers | Max-pooling | Conv2D Transpose layers | Parameters |
|--------------------|----------------------|---------------|--------------------------------------|----------------|-----------|-------------|-------------------------|------------|
| Modified UNet [30] | EP (Linear, 5) | 10 | 3×3 | 5 | - | 4 | - | 1,914,105 |
| | DP (Linear, 4) | 9 | $3 \times 3, 1 \times 1$ | 4 | - | - | 4 | |
| TPRE-UNet | EP (TPR, 4) | 16 | $5 \times 5, 3 \times 3$ | - | 16 | 3 | - | 1,398,918 |
| | DP (Linear, 3) | 7 | $3 \times 3, 1 \times 1$ | 3 | - | - | 3 | |
| TPRD-UNet | EP (Linear, 4) | 8 | 3×3 | 4 | - | 3 | - | 811,489 |
| | DP (TPR, 3) | 13 | $5 \times 5, 3 \times 3, 1 \times 1$ | - | 12 | - | 3 | |
| TPRED-UNet | EP (TPR, 4) | 16 | $5 \times 5, 3 \times 3$ | - | 16 | 3 | - | 1,869,762 |
| | DP (TPR, 3) | 13 | $5 \times 5, 3 \times 3, 1 \times 1$ | - | 12 | - | 3 | |

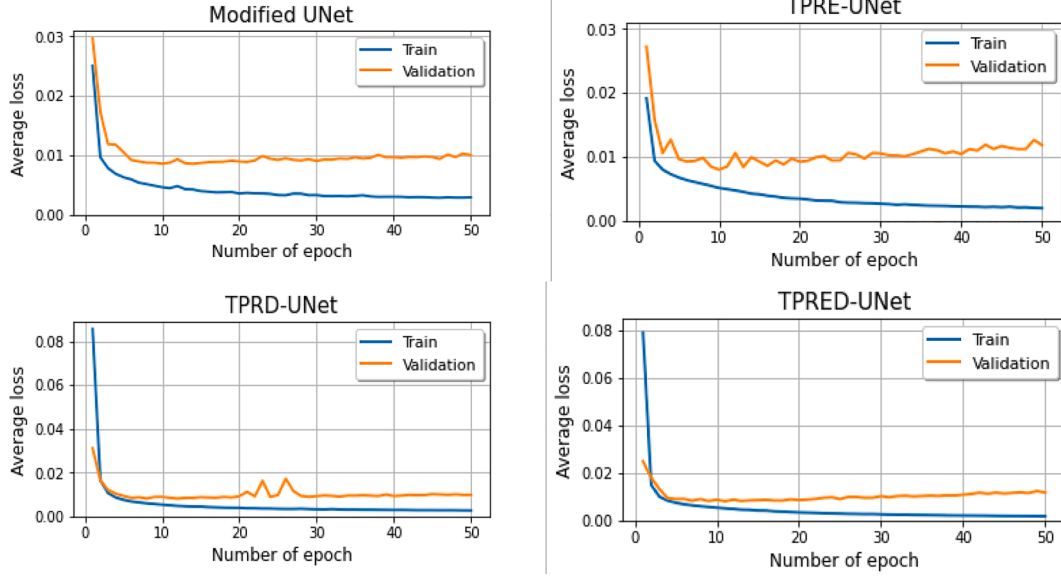


Fig. 6. The average loss function for training and validation data sets.

- First, we place the TPR blocks in the down-sampling path instead of the Linear blocks and call it TPR-Encoder-UNet (TPRE-UNet).
- Second, the first strategy is repeated for the up-sampling path, and the TPR-Decoder-UNet model (TPRD-UNet) is constructed.
- Finally, the proposed TPR-Encoder-Decoder-UNet (TPRED-UNet) model is constructed placing the TPR blocks on both UNet paths.

The TPRE-UNet and the TPRD-UNet architectures are asymmetric networks because they are designed using non-identical blocks in UNet paths. On the other hand, the TPRED-UNet architecture is a symmetric network due to having similar blocks in both encoder and decoder paths. Our experiments have shown that there is no need to increase the depth of the proposed networks for better accuracy. Therefore, only 4 and 3 blocks are placed in the down-sampling and up-sampling paths, respectively. A summary of the structure of the models is presented in Table 1, which shows the number of different layers and kernel sizes.

4. Experimental setup

4.1. Database

The proposed method is validated on BRATS database whose full name is the MICCAI Multimodal Brain Tumor Image Segmentation Challenge (BRATS) [32]. For every patient in BRATS, there are four MRI sequences: Flair, T2, T1, and T1ce. Besides, a manual segmentation for each patient is available as ground truth. The manual segmentation is performed by one to four physicians [25], in which the tumor area is determined with its main components, including edema, necrosis, enhancing, and non-enhancing region. Also, the number of samples in this database is increasing every year. For example, the 2013 database (BRATS'2013) contains 80 images of 35 LGG and 45 HGG, while the latest version, which was available in 2018 (BRATS'2018), includes 285 samples (75 LGG and 210 HGG). In our assessments, only BRATS'2018 database was used.

The MRI images of BRATS'2018, are provided by 19 institutions and have different scanners and clinical protocols to ensure data diversity [25]. All the data have been preprocessed, involving skull stripping, co-registration with T1c, and interpolation of the sequences so that all sequences have the same size ($240 \times 240 \times 155$ pixels) [20]. Compared to

the other methods, 70% of data (199 patients) were randomly selected as training data set, 15% (43 patients) as validation data set, and the rest (43 patients) as testing data set. For each MRI image, 155 slices can be extracted from the axial view. But, all these slices are not useful because some are black. So, black slices are discarded as mentioned in Section 3.1. Due to memory limitation, among the remaining slices of each patient's MRI image, only slices with even indexes were selected for training and evaluation of networks. But in the testing process, all slices of each patient's MRI images (except the black slices) are considered. It should be noted that the number of slices selected from the brain MRI image for each patient varies based on the number of black slices of the patient's brain MRI image. Therefore, about 13532, 2900, and 5840 images have been used for training, validation, and testing, respectively.

4.2. Training

The proposed models are implemented in Google Colab¹ using Keras library. Colab has made it possible for users to use the GPU for free. NVIDIA Tesla K80, NVIDIA Tesla T4, and NVIDIA Tesla P100-PCIE GPUs are currently available in Colab environment, but a user cannot choose the GPU type. In the training process, the loss function binary cross-entropy and adam optimizer are used. In order to speed up the training process, the number of batch sizes is set to 34.

4.3. Evaluation criteria

Five criteria were used to evaluate the performance of the proposed method in determining the tumor area, including accuracy, dice similarity coefficient (DSC), sensitivity, specificity, and positive predictive value (PPV). These criteria are defined in formulas (5) to (9), respectively:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (5)$$

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \quad (6)$$

¹ <https://colab.research.google.com>

Table 2

Performance (%) / Standard Deviation of different UNet-based architectures.

| Method | DSC | Sensitivity | PPV | Specificity | Accuracy |
|--------------------|-------------------|-------------------|-------------------|--------------------|--------------------|
| Modified UNet [30] | 89.33/0.52 | 87.25/2.26 | 91.71/2.23 | 99.83/0.19 | 99.72/0.024 |
| TPRE-UNet | 89.4/0.5 | 87.77/2.6 | 90.88/2.53 | 99.87/0.053 | 99.71/0.044 |
| TPRD-UNet | 89.54/0.56 | 88.31/1.36 | 90.49/1.63 | 99.87/0.021 | 99.72/0.027 |
| TPRED-UNet | 89.76/0.36 | 89.19/2.18 | 90.65/1.65 | 99.87/0.036 | 99.73/0.031 |

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

$$\text{PPV} = \frac{TP}{TP + FP} \quad (9)$$

where ‘TP’ is the number of pixels in the true tumor area, ‘FP’ is the number of pixels in the false tumor area, ‘TN’ is the number of pixels in the true non-tumor area, and ‘FN’ is the number of pixels in the false non-tumor area.

5. Experimental results and analysis

To confirm the performance of the proposed networks on the segmentation of brain tumor images, the test results were examined from different perspectives. As mentioned in Section 4.1, in the training process, 70% of samples were randomly selected for training and the rest for testing and validation. This was repeated 10 times, and the average evaluation criteria values were reported. Also, for a fair evaluation, the same 10 data sets for training and evaluating all architectures were considered.

5.1. Network convergence

At the end of each epoch of the training process, the value of the loss function on the training and the validation data set is calculated and the average values obtained in 10 times the execution of the algorithm are shown in Fig. 6. It has been observed that increasing the number of epochs in the training of the proposed networks leads to an increase in computational error on the validation data. Therefore, the best number of epochs to train the modified UNet is set to 30. Also, the number of epochs for training the proposed models is 15. This indicates that TPR blocks guarantee faster convergence of the UNet structure.

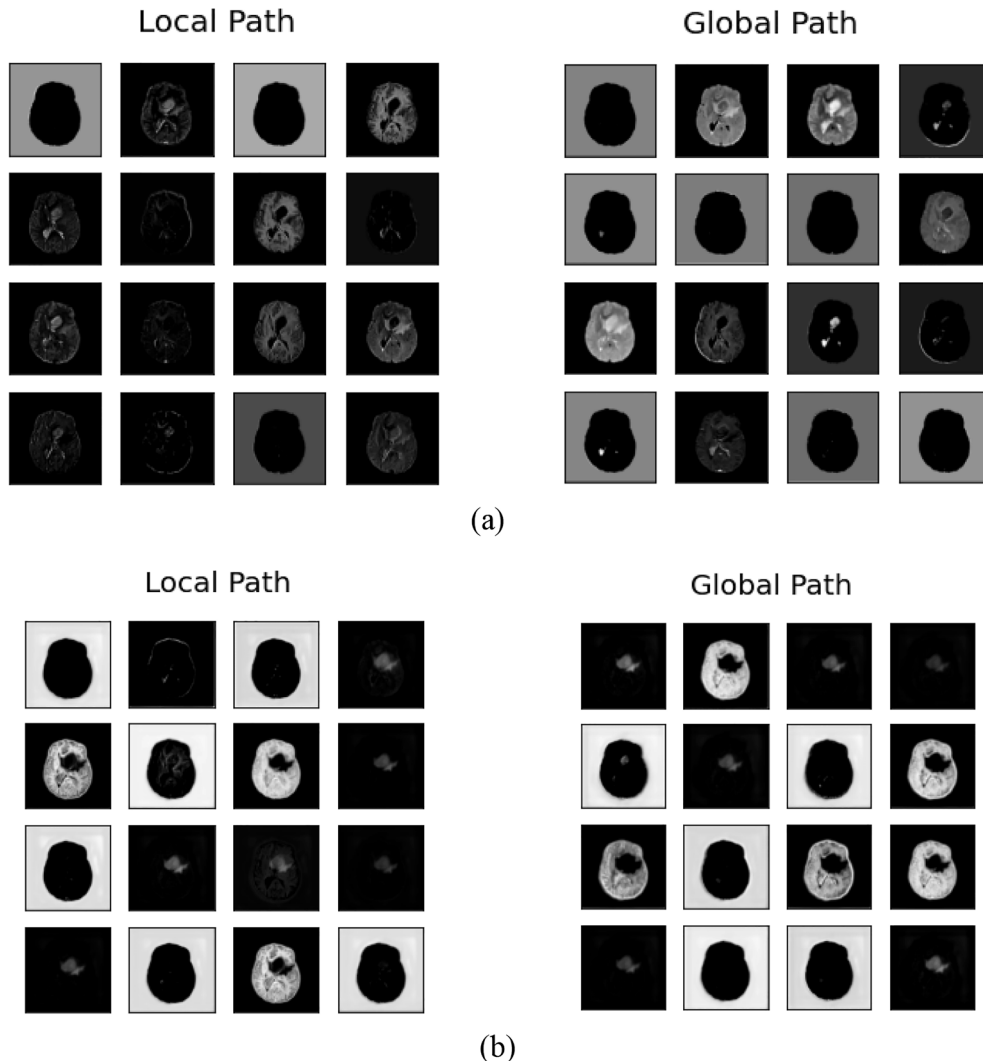


Fig. 7. Output visualization of (a) the first TPR block in down-sampling path, and (b) the last TPR block in up-sampling path of TPRED-UNet architecture.

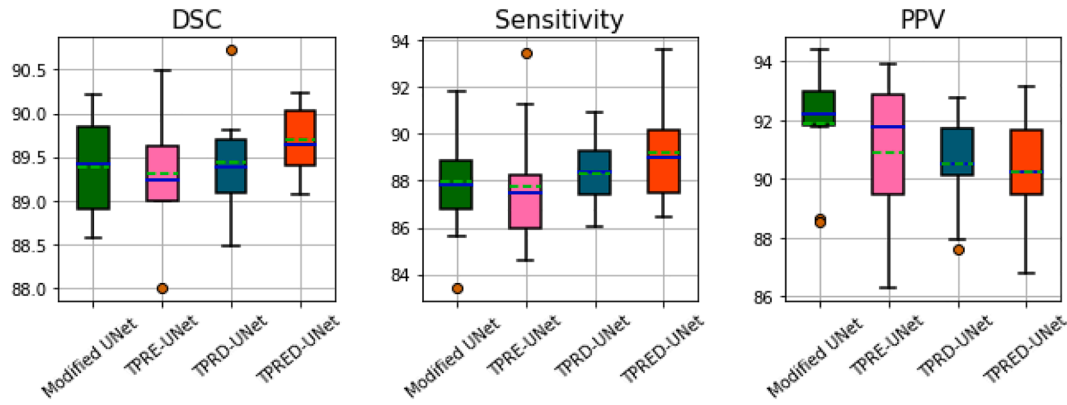


Fig. 8. Boxplot of the results on 10 test sets selected from the BRATS'2018 database using the proposed architectures: TPPE-UNet, TPRD-UNet, TPRED-UNet, and the modified UNet architecture, in terms of DSC, sensitivity, and PPV. The arithmetic mean is presented in green, the median in blue, and the outliers in brown.

Table 3

Comparison of average time required for the different segmentation methods in a single segmentation test.

| Method | Time (s) | Parameters |
|--------------------|----------|------------|
| Modified UNet [30] | 15.39 | 1,914,105 |
| TPPE-UNet | 5.88 | 1,398,918 |
| TPRD-UNet | 5.1 | 811,489 |
| TPRED-UNet | 7.53 | 1,869,762 |

5.2. Quantitative evaluation

Table 2 shows the average/standard deviation of evaluation criteria values obtained on the test data sets from the BRATS'2018 database. The high value of PPV for the segmentation results of the modified UNet architecture indicates an effective reduction in false positives; so this model works well in accurately predicting true negatives. However, due to the low sensitivity, it is a poor model in accurately determining the tumor area. Alongside, the proposed TPPE-UNet architecture is superior to the modified UNet in determining the non-tumor region because it has a higher mean (99.87%) and a lower standard deviation (0.053) in terms of specificity.

The good performance of TPR blocks in TPRD-UNet architecture can be clearly seen with significant false negatives reduction. TPR blocks are more successful in retrieving feature maps for the original resolution than the linear blocks because the former have filters with different sizes. The high sensitivity criterion confirms this result. However, the prediction of the model in non-tumor pixels is slightly weaker than the previous two models. On the other hand, it has been observed that using TPR blocks in both UNet paths, false positives and false negatives are reduced simultaneously. This means that TPRED-UNet architecture attempts to correctly identify the non-tumor area, in addition to accurately determining the tumor area. Therefore, the similarity of the segmentation results of this model and the ground truth is greater than the other architectures. The mean and standard deviation of the DSC for the results of this model are 89.76% and 0.36, respectively, which confirms the above claim.

For better understanding, the output of two TPR blocks in TPRED-UNet architecture is shown in Fig. 7. Fig. 7(a) refers to the output of the first TPR block in down-sampling path, while Fig. 7(b) indicates the output of the last TPR block in up-sampling path of the network. It is easy to see that the global path filters learn more about the global (contextual) features, whereas the local path filters focus more on edge detection.

The distribution of the criteria values obtained from 10 test sets is shown in Fig. 8. According to the box length, the scatter of DSC data for the results of the proposed architectures is reduced, indicating that these data are in a more coherent range. Accordingly, the proposed TPRD-

Table 4

Comparison of the different brain tumor segmentation methods in the BRATS database.

| Method | Database | DSC | Sensitivity | PPV | Specificity |
|-----------------------|------------|-------|-------------|-------|-------------|
| Hussain et al. [2] | BRATS'2013 | 87 | 90 | – | 94 |
| Havaei et al. [11] | BRATS'2013 | 88 | 87 | – | 89 |
| Shen et al. [17] | BRATS'2013 | 87 | 89 | 85 | – |
| Pereira et al. [24] | BRATS'2013 | 88 | 89 | 88 | – |
| Zhao et al. [19] | BRATS'2013 | 88 | 86 | 90 | – |
| | BRATS'2015 | 84 | 82 | 89 | – |
| Li et al. [14] | BRATS'2015 | 86 | 84 | 90 | – |
| Abd-Ellah et al. [13] | BRATS'2017 | 89 | 89 | – | 99 |
| Wang et al. [20] | BRATS'2018 | 91 | 94 | 89 | – |
| Caver et al. [25] | BRATS'2018 | 87.8 | 88.8 | – | 99.3 |
| Cabezas et al. [26] | BRATS'2018 | 88.92 | 88.84 | – | 99.48 |
| Hu et al. [27] | BRATS'2018 | 86 | 88 | – | 99 |
| Chen et al. [28] | BRATS'2018 | 88.78 | 90.12 | – | 99.42 |
| Modified UNet [30] | BRATS'2018 | 89.33 | 87.25 | 91.71 | 99.83 |
| Proposed TPPE-UNet | BRATS'2018 | 89.4 | 87.77 | 90.88 | 99.87 |
| Proposed TPRD-UNet | BRATS'2018 | 89.54 | 88.31 | 90.49 | 99.87 |
| Proposed TPRED-UNet | BRATS'2018 | 89.76 | 89.19 | 90.65 | 99.87 |

UNet architecture is superior in terms of sensitivity. Although the PPV box length is much shorter for the results of the modified UNet architecture, more outliers are observed that are much smaller than the most of data. Therefore, one of the important advantages of the proposed architectures is to reduce their dependency on the selected data set and increase their generalizability to the other data sets.

5.3. Computation time

Table 3 provides the average time required for different segmentation methods to complete a single segmentation. Computation time is the sum of time required for loading a model, pre-process a patient's MRI image, and its segmentation. It should be noted that all architectures are tested on the same GPU (NVIDIA Tesla T4) based on the experimental setup as mentioned in section 4.2. As can be seen, the proposed architectures reduce the computational complexity and computational time of the proposed method because they have fewer parameters than the modified UNet architecture. Although asymmetric architectures have fewer parameters, they perform well in segmenting brain tumors. In the proposed TPRD-UNet architecture, which has the best performance in terms of reducing the standard deviation of sensitivity, the number of parameters is 811,489.

5.4. Quantitative comparison with other brain tumor segmentation methods

The simulation results of this study for the same data set are

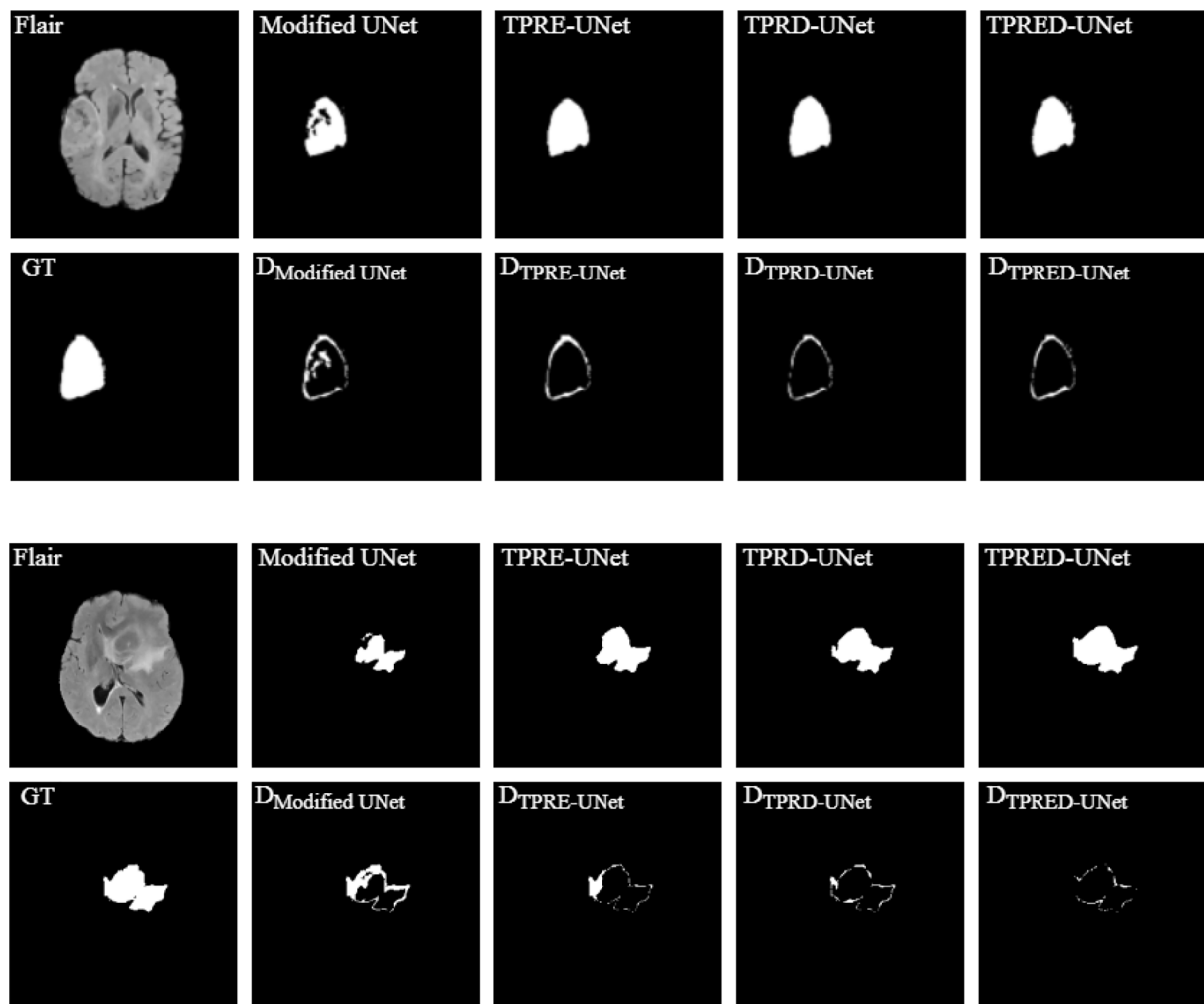


Fig. 9. Visual results from UNet-based architectures. For each subfigure, the top row from left to right shows Flair modality, the modified UNet, the proposed TPRED-UNet, the proposed TPRD-UNet, and the proposed TPRED-UNet model. The first image in the second row from the left shows ground truth, and the rest belong to the difference between the ground truth and the segmentation results. Note that 'D_x' refers to the difference between the ground truth and the segmentation result of 'x' model.

compared with the state-of-the-art methods that have been reported on the BRATS database and presented in Table 4. All evaluation criteria for the other methods are taken from the reference articles. It can be seen from Table 4 that the proposed architectures have better performance than the compared methods under DSC, PPV, and specificity criteria. This good performance is achieved by evaluating the proposed algorithm on the BRATS'2018 database, which has more data diversity than the other versions of BRATS [2,11,13,14,17,19,24].

The models presented in this study are designed by placing TPR blocks in the UNet structure. Hence, they are a type of dense prediction network and provide faster segmentation compared to the single-label prediction schemes in [2,11,19,24,26]. Also, the algorithms designed in [14,17,20] use all four types of MRI sequences to separate the brain tumor. While, the proposed algorithm uses only Flair, T2, and T1ce sequences to determine the tumor area. Therefore, computing costs have been dropped. In addition, learning the proposed models is end-to-end, and does not require post-processing compared to the other methods in [2,11,14,24]. Expert approval is required to refine the segmentation result by post-processing algorithms. Also, some post-processing methods increase the computational cost.

The number of parameters for the proposed architecture is much lower than the original UNet [29] parameters (approximately 31 million parameters) used in the study of Caver et al. [25]. The architecture provided by Wang et al. [20], ranks the highest in the sensitivity

criterion, indicating that it has been very successful in determining the tumor area. In comparison, our best architecture i.e., TPRED-UNet, tries to correctly identify the tumor and non-tumor tissues at the same time. In the segmentation of brain tumor image, the reduction of FPs is as important as the reduction of FNs because an inappropriate segmentation can divert the course of treatment. Thus, TPRED-UNet performance is better.

The overall performance of the proposed architectures i.e., TPRED-UNet, TPRD-UNet, and TPRED-UNet over the entire database is summarized in Fig. 9 and Fig. 10. As can be seen in Fig. 9, with the increasing number of TPR blocks in the UNet structure, the difference between the segmentation result and the ground truth decreased. According to Fig. 10, the proposed TPRD-UNet and TPRED-UNet architectures perform better than the others. As a result, asymmetric architectures are proper for the segmentation of brain tumor images.

6. Conclusions

In this paper, an automated algorithm for segmenting brain tumors from MRI sequences is presented. First, a two-step pre-processing was used to discard the black slices of the MRI image. Then, the proposed models were used to segment the brain tumor image. These models were obtained by adding two-pathway-residual (TPR) blocks to the UNet structure. Two-pathway-residual blocks exploit both local features as

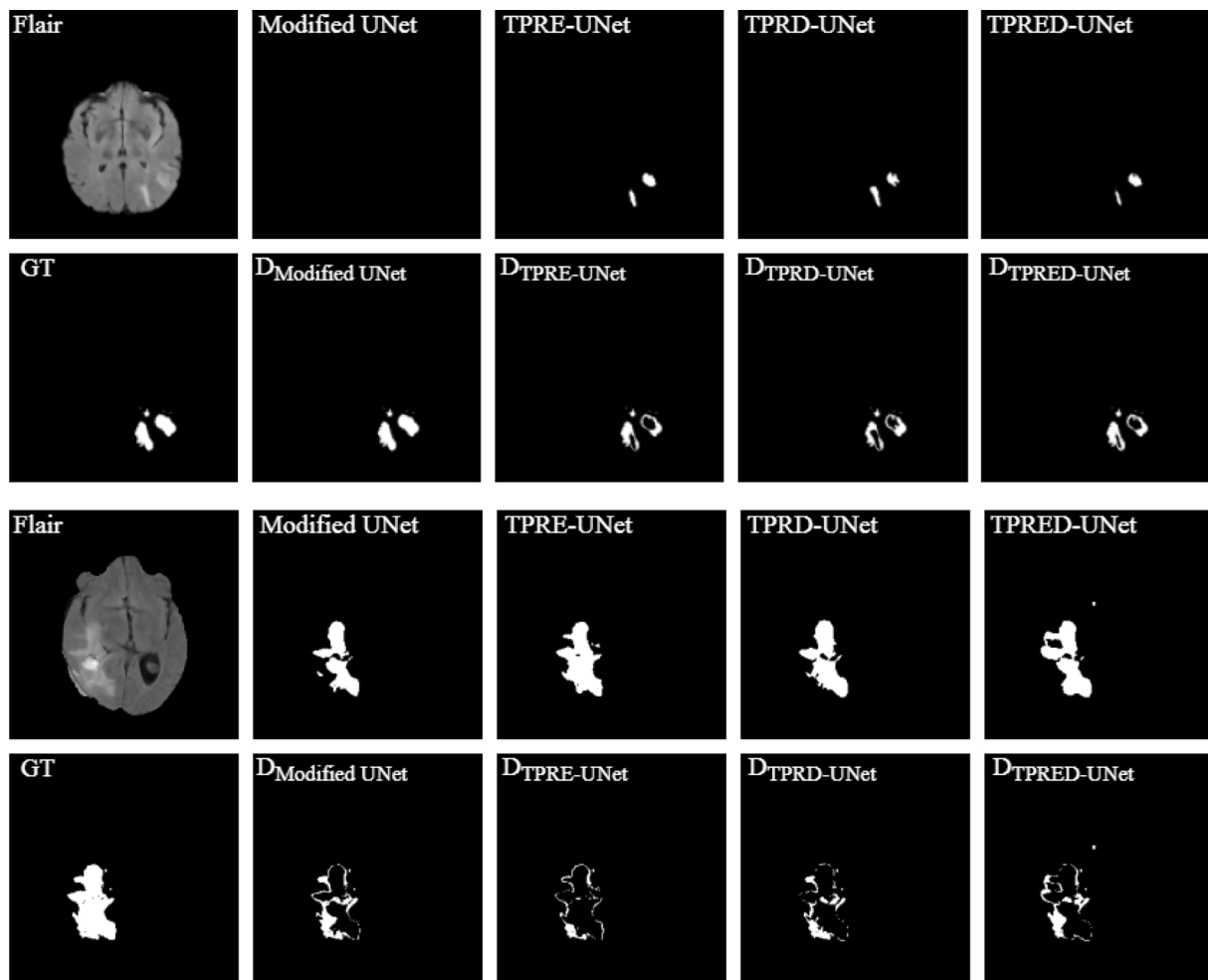


Fig. 10. Visual results from UNet-based architectures. For each subfigure, the top row from left to right shows Flair modality, the modified UNet, the proposed TPRE-UNet, the proposed TPRD-UNet, and the proposed TPRED-UNet model. The first image in the second row from the left shows ground truth, and the rest belong to the difference between the ground truth and the segmentation results. Note that D_x refers to the difference between the ground truth and the segmentation result of 'x' model.

well as more global features, simultaneously. By the presence of TPR blocks in the UNet structure, not only the evaluation criteria such as DSC and sensitivity have been improved, but also the number of parameters of the proposed models has been reduced. In summary, the advantages of the proposed models include reducing computational cost, increasing segmentation speed, and not using post-processing.

The assumption of using skull-stripped MRI images can be considered as a limitation of the proposed algorithm. Skull stripping preprocess can be employed to segment brain tissues from non-brain regions in MRI images.

The analysis of deep neural networks, especially UNet, shows the high capacity of these architectures in segmentation tasks. Hence, the future study is the appropriate structural design using deep neural networks for object-aware segmentation tasks.

CRediT authorship contribution statement

All three persons were involved in all stages of preparing this research including: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing; Motahareh Aghaglari as a M.Sc. student, Ali Aghagolzadeh (the corresponding author) as supervisor and Mehdi Ezoji as advisor.

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.

Acknowledgement

The authors would like to acknowledge the funding support of Babol Noshirvani University of Technology through grant program No. BNUT/389059/99.

References

- [1] V. Collins, Brain tumours: Classification and genes. *Journal of Neurology, Neurosurgery & Psychiatry*. 2004;75(suppl 2):ii2-ii11.
- [2] S. Hussain, S.M. Anwar, M. Majid, Segmentation of glioma tumors in brain using deep convolutional neural network, *Neurocomputing* 282 (2018) 248–261.
- [3] Robert, Christian., Machine learning, a probabilistic perspective. 2014:62-63.
- [4] T. Batchelor, Patient information: high-grade glioma in adults (Beyond the Basics), *UpToDate* (2013) 1–6.
- [5] H. Ohgaki, P. Kleihues, Population-based studies on incidence, survival rates, and genetic alterations in astrocytic and oligodendroglial gliomas, *Journal of Neuropathology & Experimental Neurology* 64 (6) (2005) 479–489.
- [6] D.N.H. Louis, O. Ohgaki, D. Wiestler, W.K. Cavenee, WHO classification of tumours of the central nervous system, World Health Organization classification of tumours, World Health Organization. 2007.

- [7] N. Hattori, Y. Hirose, H. Sasaki, S. Nakae, S. Hayashi, S. Ohba, et al., World health organization grade ii–iii astrocytomas consist of genetically distinct tumor lineages, *Cancer Sci.* 107 (8) (2016) 1159–1164.
- [8] A. Hekmatnia, M. Sabouri, A.H. Ghazavi, M. Salehi, F. Hekmatnia, G. Jamalipour Sofi, M. Alinezhad, P. Shahriyari Far, Diagnostic value of MRI enhancement in determining the degree of malignancy of Brain Tumors in patients, *Med. Sci.* 23 (100) (2019) 946–952.
- [9] M.W. Nadeem, M.A.A. Ghamdi, M. Hussain, et al. Brain tumor analysis empowered with deep learning: a review, taxonomy, and future challenges, *Brain Sci.* 2020;10 (2):118.
- [10] A. Bousselham, O. Bouattane, M. Youssfi, A. Raihani, Towards reinforced brain tumor segmentation on mri images based on temperature changes on pathologic area, *Int. J. Biomed. Imaging* (2019) 1–18.
- [11] M. Havaei, et al., Brain tumor segmentation with deep neural networks, *Med. Image Anal.* 35 (2017) 18–31.
- [12] J. Tong, et al., MRI brain tumor segmentation based on texture features and kernel sparse coding, *Biomed. Signal Process. Control* 47 (2019) 387–392.
- [13] M.K. Abd-Ellah, A.A.M. Khalaf, A.I. Awad, H.F.A. Hamed, TPUAR-Net: Two Parallel U-Net with Asymmetric Residual-Based Deep Convolutional Neural Network for Brain Tumor Segmentation. In: Karray F., Campilho A., Yu A. (eds) *Image Analysis and Recognition. ICIAR, Lecture Notes in Computer Science.* 2019; 11663.
- [14] Q. Li, Z. Gao, Q. Wang, Glioma segmentation with a unified algorithm in multimodal MRI images, *IEEE Access* 6 (2018) 9543–9553.
- [15] S. Alqazzaz, X. Sun, X. Yang, et al., Automated brain tumor segmentation on multimodal MR image using SegNet, *Comp. Visual Media.* 5 (2019) 209–219.
- [16] N. Gupta, P. Bhatele, P. Khanna, Glioma detection on brain MRIs using texture and morphological features with ensemble learning, *Biomed. Signal Process. Control* 47 (2019) 115–125.
- [17] H. Shen, J. Zhang, W. Zheng, Efficient symmetry-driven fully convolutional network for multimodal brain tumor segmentation. in 2017 IEEE International Conference on Image Processing (ICIP). IEEE. 2017.
- [18] S. Chena, C. Dinga, M. Liu, Dual-force convolutional neural networks for accurate brain tumor segmentation, Preprint submitted to Pattern Recognition. 2018;1-33.
- [19] X. Zhao, et al., A deep learning model integrating FCNNs and CRFs for brain tumor segmentation, *Med. Image Anal.* 43 (2018) 98–111.
- [20] Y. Wang, C. Lia, T. Zhua, J. Zhang, Multimodal brain tumor image segmentation using WRN-PPNet, *Comput. Med. Imaging Graph.* 75 (2019) 56–65.
- [21] U. İlhan, A. İlhan, Brain tumor segmentation based on a new threshold approach, *Procedia Comput. Sci.* 120 (2017) 580–587.
- [22] Z. Rahima, et al. Segmentation of low-grade gliomas based on the growing region and level sets techniques. in 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP). IEEE. 2018.
- [23] B. Devkota, et al., Image segmentation for early stage brain tumor detection using mathematical morphological reconstruction, *Procedia Comput. Sci.* 125 (2018) 115–123.
- [24] S. Pereira, et al., Brain tumor segmentation using convolutional neural networks in MRI images, *IEEE Trans. Med. Imaging* 35 (5) (2016) 1240–1251.
- [25] E. Caver, C. Liu, W. Zong, Z. Dai, N. Wen, Automatic Brain Tumor Segmentation Using a U-net Neural Network. Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge. 2018; 63.
- [26] M. Cabezas, S. Valverde, S. Gonzalez-Villa, A. Clerigues, M. Salem, K. Kushibar, et al, Survival prediction using ensemble tumor segmentation and transfer learning. Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge. 2018; 54.
- [27] X. Hu, M. Piraud, Multi-level Activation for Segmentation of Hierarchically-nested Classes on 3D-UNet. Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge. 2018; 188.
- [28] W. Chen, B. Liu, S. Peng, J. Sun, X. Qiao, S3D-UNet: Separable 3D U-Net for Brain Tumor Segmentation. Pre-Conference Proceedings of the 7th MICCAI BraTS Challenge. 2018; 91.
- [29] Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015; 234-241.
- [30] D. Barh, Artificial Intelligence in Precision Health, Chapter3: Deep learning in Precision Medicine, Elsevier, 1st Edition, 2020. Available: <https://www.elsevier.com/books/artificial-intelligence-in-precision-health/barh/978-0-12-817133-2/>.
- [31] T. Mostafiz, I. Jarin, S.A. Fattah, C. Shahnaz, Retinal Blood Vessel Segmentation Using Residual Block Incorporated U-Net Architecture and Fuzzy Inference System. IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE). 2018; 106-109.
- [32] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (BRATS), *IEEE Trans. Med. Imaging* 34 (2015) 1993–2024.