IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 20 June 2023, accepted 7 July 2023, date of publication 10 July 2023, date of current version 14 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294179

RESEARCH ARTICLE

ResUNet+: A New Convolutional and Attention Block-Based Approach for Brain Tumor Segmentation

SEDAT METLEK[1] **AND HALİT ÇETİNER**[2]
[1]Vocational School of Technical Sciences, Burdur Mehmet Akif Ersoy University, 15100 Burdur, Turkey
[2]Vocational School of Technical Sciences, Isparta University of Applied Sciences, 32260 Isparta, Turkey
Corresponding author: Sedat Metlek (sedatmetlek@mehmetakif.edu.tr)

ABSTRACT The number of brain tumor cases has increased in recent years. Therefore, accurate diagnosis and treatment of brain tumors are extremely important. Accurate detection of tumor regions is difficult, even for experts, because brain tumor images are low-contrast, noisy and contain normal tissue-like structures. Therefore, in this study, a new convolution-based hybrid model was proposed to perform segmentation with high accuracy. In the proposed model, instead of applying convolution to the whole image, convolution was applied to the ROI regions detected in different modalities. With this approach, it was determined that the processing cost is reduced, and the performance is increased. The proposed model was tested on BraTS 2020, BraTS 2019, and BraTS 2018 datasets. The proposed method in the study was also compared with SOTA methods using the same dataset. As a result of the comparison, dice scores of 92.80%, 93.10%, and 91.90% were respectively obtained for whole tumors, enhance tumors and tumor nuclei in the images in the BraTS 2020 dataset. With these results, the proposed model can compete with many models in the literature using the same datasets. The proposed model is a new method that can be preferred in different segmentation applications due to its performance success and especially the advantage of the pre-processing structure.

INDEX TERMS Image segmentation, deep learning, tumors, convolution, deconvolution.

I. INTRODUCTION

Medical imaging systems are the leading decision support systems that play an important role in medical treatment and diagnosis. These systems are widely used in the detection of lesioned areas, such as tumors. Image processing methods are used in the general structure of medical imaging systems. Image processing methods used in medical imaging systems are widely used in the pre-processing of lesioned areas, determination of their distinctive features, selection, and classification [1]. One of the preferred methods among image processing methods is the segmentation method. The Region of Interest (ROI) can be detected on the image using the segmentation method. Image segmentation, which is used in many different imaging applications, has a vital role in analyzing and interpreting medical images [2]. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) imaging technologies are used in many different medical applications, from tumor detection to disease classification.

MRI is an advanced method based on nuclear magnetic resonance. MRI provides the flexibility to adjust many parameters to obtain tissue resolution, sharpness, and different anatomical features [3], [4]. It also provides great convenience for patients. Because MRI is a non-invasive method. In MRI, powerful magnetic waves and radio frequencies between 1.5T and 3T are used to visualize the internal structure of the body and tissues [3]. These magnetic waves and radio frequencies do not damage the organ being imaged with high ionization and radiation effects. Therefore, it is an actively used method in brain tumor segmentation.

Image segmentation methods are heavily needed in medical decision support systems developed to assist decision-makers using these technologies. Experts interpret the difference between normal and diseased tissue using

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar.

images from MRI and CT systems. But the need for decision support systems is growing every day to minimize the potential problems caused by high workloads and human error. With the increase in cancer cases today, the segmentation of diseased tissue has become an extremely important issue. Considering that patients with malignant brain tumors have a very short life expectancy, it is extremely significant to detect and segment ROI regions with tumors as early as possible and with high accuracy [5]. These situations have been the main motivation for the study.

Tissue deterioration can be seen as a result of neoplastic and aggressive growth of cells in the brain. These tissue disruptions are also called lesions or neoplasia. In general, brain tumors are classified as primary and metastatic tumors. Gliomas, which are primary brain tumors, are composed of glial cells [6]. Metastatic brain tumors are tumors that migrate to the brain from different parts of the body through the bloodstream. The use of segmentation methods in MRI and CT images, which are also used in the detection of these brain tumors, is an important support that can help decision-makers make decisions with high accuracy. As a result, an accurate glioma segmentation can increase the survival rate of patients, helping specialists from surgery planning to postoperative observations. However, there are some difficulties in the segmentation processes [7], [8].

One of the main difficulties encountered in brain tumor segmentation arises from the choice of contrast used in MRI scans. Because contrast selection is affected by clinical difficulties. What is meant by clinical difficulties here is the requirement of local varying relaxation applied to the area to be MRI scanned. However, this process cannot be applied in brain tumor segmentation applications in general. In addition, the selection of the optimal imaging plane is another clinical difficulty in contrast selection.

Factors affecting the quality of MRI scans are low contrast, noise in the image, contrast-to-noise ratios, image resolution, and scan time [9]. The scale of voxel values in MRI is not standardized. Therefore, the same tumor cells can have very different gray level values depending on the type of MR scanner (7, 3, or 1.5 Tesla) and protocol (field of view value, voxel resolution, gradient power, b0 value, etc.) [10], [11]. In addition to all these, in some cases, MRI results of the tumor region and MRI results of the healthy region may show similar characteristics. Especially in brain tumor segmentation, the use of spatial information as the primary tool (e.g., template matching or image warping), as suggested for normal tissue segmentation, is not appropriate. This is because there are no a priori spatial criteria for changes that can be easily applied [12].

Among patients with glioma brain tumors, the size, morphology, and location of gliomas are among the reasons that make the segmentation process difficult. In addition to these reasons, many small artifacts found in the MRI scan make the tumor classification process difficult because they disrupt the balance between the lesion region and its background [13].

This is also the main problem faced by experts working on the subject. Because the detection of inhomogeneous junctions of tumor regions complicates the classification of tumors such as glioma brain tumors. Besides, tumors are segmented as Enhancing Tumor (ET), Edema, and Non-Enhancing Tumor (NET) / Necrotic Tumor Core (NTC), as shown in Figure 1 [14]. In addition, the densities of subregions of some tumor types may overlap with healthy tissues. Images in different modalities are used to solve these problems. The modalities frequently used in the literature are T1, T1ce, T2, and FLAIR modalities presented in Figure 1, respectively [15].

Images in the T1 modality use short echo and repetition time, while images in the T2 modality use long echo and repetition time. Images in the FLAIR modality are a T2 modality-based pulse sequence that overrides Cerebrospinal Fluid (CSF) and can therefore distinguish between abnormality and CSF. In FLAIR modality images, gray matter is brighter than white matter. The tumor regions observed in the brain have a higher intensity in FLAIR and T2 modalities than other parts of the brain. The FLAIR modality is also used to obtain information about fiber density in different applications. Especially in the FLAIR modality, normalization is recommended to remove the anisotropic effect and make the data more uniform. The T1ce modality is created by infusing a contrast substance called Gadolinium into the T1 modality. In the T1ce modality, there is contrast accumulation in the lesion areas on the tissues. Therefore, it makes that the area appear brighter. Because of this feature, it is frequently preferred, especially in tissue classification applications.

However, due to MRI devices and protocols in different hospitals, significant changes may be observed in the gray values of abnormal tissues. Therefore, in the study, a deep learning-based segmentation model was developed based on T1, T2, T1ce, and FLAIR modalities. This is also seen in the BraTS 2020, 2019, and 2018 datasets, which are benchmark datasets in the literature. In these datasets, the gray values of abnormal tissues from different hospitals' machines with various MRI protocols vary [11], [16], [17].

Therefore, ancillary decision support systems that can segment brain images are urgently needed. As a result of the widespread use of deep learning and artificial intelligence applications in the literature, it is aimed to apply segmentation processes, which are frequently used in image processing applications, on brain MRI. Although it has been determined that similar methods are used in the literature research on the subject, it is essential to increase the success of the segmentation method used. For this reason, a new model has been proposed based on the UNet architecture, which is used as a basis in segmentation applications. But in general, there are semantic gaps between the encoder and decoder in UNet architecture. In addition, it shows poor performance in images with fined details.

The proposed model aims to increase the segmentation success by eliminating this disadvantage of the UNet architecture. In accordance with this purpose, firstly, a new data

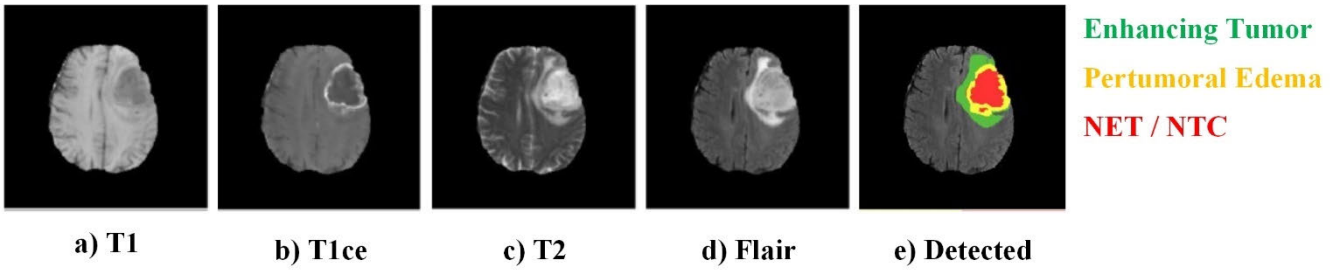

FIGURE 1. Four different MRI modalities and ground truth with corresponding manual segmentation annotation.

pre-processing strategy is proposed to remove redundant space outside the ROI region. With the proposed strategy, firstly, the location of the tumor was determined on the image, and the segmentation model recommended only in the relevant area was applied. Thus, the effect of many small artifacts that can be found throughout the image is reduced, and the negative effect on the classification success is eliminated.

As a result of reducing the size of the ROI, the overfitting problems frequently mentioned in the literature have been avoided [18]. Furthermore, the ResUNet+ model is proposed to perform segmentation by extracting effective features after the pre-processing step.

Convolution operations were applied in blocks so that the proposed model could be effective against differences in tumor location, borders and size. The convolution operations preferred in this structure are also found in strong deep-learning models used in the literature. In the proposed model, different convolution blocks used in powerful deep learning models in the literature are combined with a hybrid structure.

The main contributions of the model proposed in this study to the literature are briefly summarized below.

1) In the literature, a new pre-processing method has been proposed, different from the pre-processing stages used in brain tumor segmentation studies. In the proposed method, using different modalities images, the tumor, and ET region are more prominent.
2) The semantic gap in UNet models and the low performance in fine detail images have been eliminated with a proposed hybrid method based on residual block.
3) In the study, it was found that applying the proposed model only on ROI instead of applying it to the whole image increased the success performance.
4) The proposed method is comparatively tested on BraTS 2020, BraTS 2019, and BraTS 2018 datasets, which are considered reliable for brain tumor segmentation in the literature. The comparison results show that the proposed model achieves higher classification accuracy than the state-of-the-art methods in the literature.

The remainder of this paper is organized as follows. Studies involving state-of-the-art techniques and technologies in brain tumor segmentation are described in Section II. The pre-processing method and datasets applied in the study are presented in detail in Section III. The proposed architecture and other architectures used in the study with implementation details, and evaluation metrics are explained in Section IV. Results and discussion are presented in Section V. The paper is concluded in Section VI.

II. RELATED WORKS

Recently, deep learning-based methods based on automatic feature extraction for brain tumor segmentation have attracted great interest and became a new research topic in the field of medical imaging [7]. For this reason, U-shaped architectures have lately been developed and used in medical image segmentation applications [19]. Walsh et al. have developed a lightweight UNet application for brain tumor segmentation. They tested their application on the BITE dataset and claim the results are promising. In their study, they achieved an intersection-over-union (IoU) 89% success rate [20]. Ronneberger et al. [19] developed a U-shaped segmentation model by sequentially applying convolutional operations. In the model they developed, the features obtained at the encoder stage for the segmentation process are combined and reassembled in the decoder section. However, as a result of this process, semantic gaps may occur. In order to eliminate these semantic gaps, some improvements have been suggested in the literature by using different skip connections [21], [22].

Zhou et al. [23] developed a convolution-based segmentation model called UNet++, which includes dense skip connections to meet the need for medical image segmentation. In the model they developed, the skip connections they used showed better performance than the UNet model. However, it contains too many parameters compared to the UNet model used. Therefore, in ResUNet+ model, simpler skip connections are preferred compared to UNet++ model.

Cao et al. [24] developed the DenseUNet model to increase the success of the UNet segmentation model and to overcome the vanishing gradient problem. In this UNet-based model, they used dense blocks instead of convolution layers. They tested their model on ISBI 2012 EM dataset without any post-processing module or pre-training. In the ResUNet+ model, residual blocks are used together with convolution layers to improve the discrimination in feature maps. In addition, node structures were added between the encoder-decoder.

Liu et al. [25] preferred pre-trained block structures in the structure of the segmentation model to obtain better features than the U-shaped segmentation model found in the literature. In the model they created, 3D sliced images were used. Qamar et al. [26] proposed a new UNet model that uses inception module constructs to segment brain tissues that are highly desirable to divide in the clinical setting. Alom et al. [25] similarly added residual and recurrent modules to the UNet segmentation model. More specifically, Zhang et al. [27] developed the residual attention UNet (AResU-Net), which simultaneously deploys the attention mechanism and residual units to the UNet for an end-to-end 2D brain tumor segmentation network. Huang et al. focused on feature enhancement. Therefore, they collected multi-scale features in their study [28]. Zhang et al. [29] developed a new UNet model based on a swin transformer to fully utilize the spatial and channel information of the image. Wang et al. [30] proposed to combine the encoder-decoder structures used in the U-Net structure into repetitive loop unit structures. Zhuang et al. [31] developed a model capable of capturing sensitive features. In the model they developed, they designed multiple encoder-decoder branch pairs.

Although these studies presented in the literature are very important studies in terms of eliminating the deficiencies of classical UNet models, images can be under or over-segmented due to the structure of convolution operations and medical images. In addition, developing the network model only on a certain dataset to obtain stronger features, and the representation may reduce the generalization performance of the model. This can cause irreparable errors, especially in medical images. As a result, deep learning-based segmentation models in the literature need to be both improved and tested for their accuracy on different datasets [14].

For this reason, in the developed model, possible semantic gaps that may arise from the features spread in the encoder section are minimized. In addition, the structure placed between the encoder and decoder is aimed to prevent the semantic gap that may occur. Thus, the main advantage of the proposed model is to reduce semantic gaps and improve segmentation performance by using new convolution blocks. For this reason, in this study, a new model was developed to represent more powerful features and a test was performed on different datasets.

III. MATERIAL

In this section, the BraTS dataset, which is a public dataset, is focused on in order to make an accurate comparison of the study with different methods in the literature. In the next subsection of the study, detailed information about the firstly used dataset is given. Afterward, the pre-processing performed on this dataset is presented.

A. DATASET

Brain tumor segmentation is a difficult and complex problem due to the low quality of images obtained from imaging systems and the variability of intensity regions. For this reason, pre-processing was applied to improve the quality of the images to be used in the study. In the study, BraTS 2020, BraTS 2019, and BraTS 2018 datasets were used as training and testing material on the proposed model and UNet model.

The MRI scan images in the BraTS 2020, BraTS 2019, and BraTS 2018 datasets were obtained from a total of 19 different medical institutions with different MRI scanners. BraTS 2020, BraTS 2019, and BraTS 2018 datasets contain two different datasets: High-grade glioma (HGG) and low-grade glioma (LGG). The BraTS 2020 dataset includes 369 patient images with four modalities in which tumor areas were manually segmented by experts. There are 125 patient images with four modalities as validation data. There are 335 training data with four modalities in the BraTS 2019 dataset. Similarly, there are 125 patient images with four modalities for validation data. In the BraTS 2018 dataset, there are 285 training data and 66 validation images of four modalities [32], [33].

Details about the number of images in the datasets used are presented in Table 1. In practice, these data were divided into two groups training and testing according to the cross-validation 5 value. The image in each dataset has dimensions of 240 × 240×155. The third dimension of the images used refers to the slice layer.

TABLE 1. Parameters of the used datasets [14].

Datasets	BraTS 2020	BraTS 2019	BraTS 2018
Total Number of Image	660	626	542
Total Number of Training Image	369	335	285
Total Number of Validation Image	125	125	66
Total Number of Testing Image	166	166	181
Tasks	Segmentation Survival Prediction	Segmentation Survival Prediction	Segmentation Survival Prediction
Timepoint	Pre-operative	Pre-operative	Pre-operative

In the study, four different modalities of images were preferred, namely T1, T1ce, T2, and FLAIR. In the literature, the T1ce method is generally used in the detection of tumor nuclei because it is brighter than other modalities. Since FLAIR and T2 modalities have higher density, the transition between tumor and normal tissues is more pronounced [34].

B. PRE-PROCESSING

One of the biggest challenges in MRI analysis is the artifacts caused by thermal noise and magnetic fields. In addition, small artifacts caused by patient movements during the scanning process can also occur. Therefore, noise in MRI scans can distort fine details [35]. Thus, subtle edges in tumor images may become blurred, and resolution may decrease. As a result, the number of discriminative features is reduced.

The performance of CNN-based approaches is also significantly reduced [36].

For this reason, pre-processing processes such as noise removal have started to be the subject of interest in studies to be conducted with MRI in the literature [37], [39]. It is obvious that the application of this subject in brain segmentation will increase performance.

Therefore, unlike many studies in the literature that use the whole image, we focus on the detection of the ROI region by pre-processing. Thus, it was aimed to exclude the non-ROI regions from the study and to prevent the factors that negatively affect the success. With this strategy, we don't need to use a very deep convolutional model, and we don't need to make convolutions on the whole image unnecessarily. As a result, it is aimed to reduce both the computational cost and memory usage compared to the classical method in which the whole image is processed.

For this reason, the pre-processing is primarily carried out in the study and this process is carried out in nine steps in total. When the images used in the study were examined, it was determined that the values in the image matrix were between 0 and 1224. Therefore, in the first step, all values in the image matrix are reduced to the range 0-1 using (1). An example image obtained after applying this process is presented in Figure 2.

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \tag{1}$$

In equation (1), I represents any image in four different modalities, I_{min} represents the smallest value and I_{max} represents the largest value in this model [40].

The second step of pre-processing is to detect the ROI region of interest in the image. For this, in addition to morphological processes, a different pre-processing method from the literature has been adopted. In this method;

1) The image in the FLAIR modality is taken as input in Figure 3 (a).
2) Since all brain images and tumor regions are elliptical-like structures, a disk-shaped structural element was chosen in the study, and this element has hovered over the figure. In the study, a 3×3 window size disk and strel structure used for morphological processes were applied.
3) Afterward, highHat and lowHat transformation functions were applied, respectively, in order to obtain a clear image by reducing the effect of uneven illumination. At this stage, imtophat and imbothat functions, which are Matlab built-in functions, which are also included in the literature, are used in the study [41]. This process is performed with a 3-radius structural element using high line (imtophat) and low line (imbothat) transformation functions. The imtophat function is used for light-colored objects on a darker background, and imbothat for non-dark light-colored regions on a lighter background [42]. Imtophat is obtained by subtracting the opened image from the original image using the morphological opening process. Imbothat, on the other hand, is obtained as a result of subtracting the image calculated with the closing operation from the original image. Both functions, it is aimed to prevent the negative effects caused by irregular lighting. The operation performed in the third step is presented in (2) [43].

$$I_{out} = [I_{in} \oplus (I_{in} \Delta \text{se})] \odot (I_{in} \nabla \text{se}) \tag{2}$$

I_{in} in (2) denotes the input image, se denotes the structural element specified in the second step, Δ imtophat operation, ∇ imbothat operation, $\oplus$ addition operation $\odot$ denotes subtraction. The result of I_{in} image and imtophat operation is presented in Figure 3 (b), and the result of imbothat operation is presented in Figure 3 (c). I_{out} also represents the filtered image obtained as a result of the (2) operation.
4) In the next step, the 25×25 floating window median filtering method is used to reduce the noise in the FLAIR modality.
5) In order to increase the effect of the background in the obtained image, the image obtained in the fourth step and the FLAIR modality image are multiplied again on an element basis in Figure 3 (d).
6) In this step, a two-dimensional 150×150 average filter design is made. The designed filter is applied to the FLAIR image used in the first step. The result obtained is subtracted from the reverse of the FLAIR image in Figure 3 (e).
7) The image obtained in the fifth step is subtracted from the image obtained in the sixth step. Thus, the background is removed from the image and the ROI region is more pronounced in Figure 3 (f).
8) The image obtained in the seventh step is divided into the FLAIR image of the first step on a pixel basis. Thus, the ROI region in the brain image is made more prominently in Figure 3 (g).
9) In the last step, a disk-shaped 3-radius structural element is created again, similar to the one in the second step, and then applied with the imclose process. Thus, small spots that may be noise on the image are cleared from the image in Figure 3 (h).

After detecting the ROI region with the tumor on the image, this region is converted to binary format and its coordinates were calculated. Afterward, the region with the same coordinate in four different modalities was determined as the image region to be used in the study, as shown in Figure 4.

In the images in the Ground Truth (GT) region, there is a total of 4 different classes, including the healthy class. The study, it is aimed to classify Tumor Core (TC), Enhancing Tumor (ET), and Whole Tumor (WT) regions. For this reason, GT images with ROI regions of each relevant area in the GT images were created to use specifically in the proposed model.

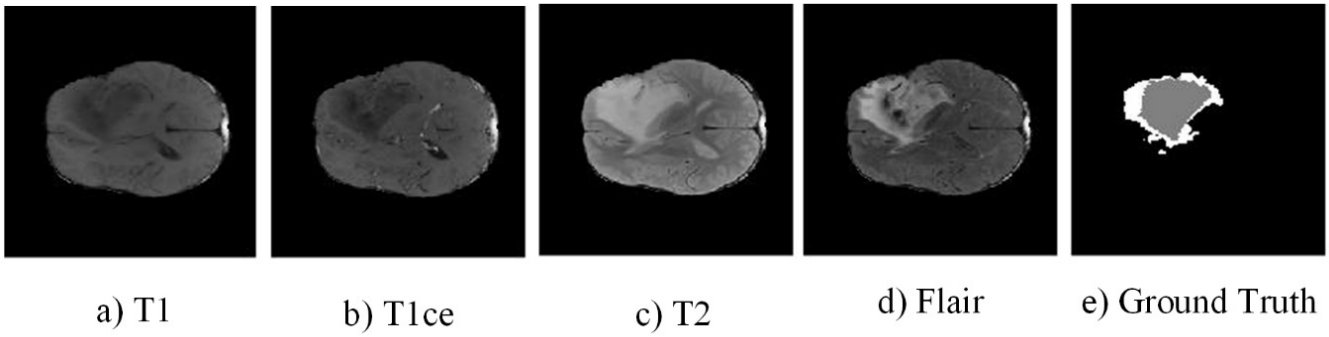

FIGURE 2. Four different modalities and ground truth mask images in the datasets used in the article are presented. In general, healthy tissues can be differentiated more clearly in the T1 modality. In the T1ce modality, tumor borders are slightly more pronounced than in other modalities. In the T2 modality, edema tumor areas are more prominent. In the FLAIR modality, the boundaries between the cerebrospinal fluid region and edema tumor regions are clearer. Ground truth, on the other hand, refers to the image obtained by manually marking the WT, ET and TC regions by experts.

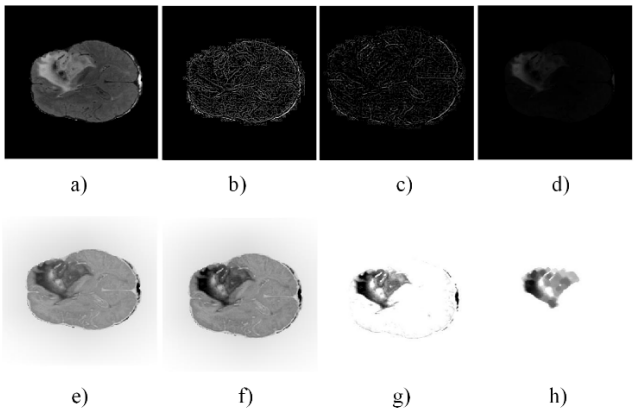

FIGURE 3. Figure 3 (a) is presented a slice view of a 3D MRI image from the BraTS 2018 dataset in the FLAIR model. This image is also used as sample input data in the pre-processing step. The results of the imtophat and imbothat operations in step 3 are shown in Figure 3 (b) and Figure 3 (c), respectively. The result of the pre-processing method at step 5 is shown in Figure 3 (d), the result at step 6 is shown in Figure 3 (e), the result at step 7 is shown in Figure 3 (f), the result at step 8 is shown in Figure 3 (g), and the final result is shown in Figure 3 (h).

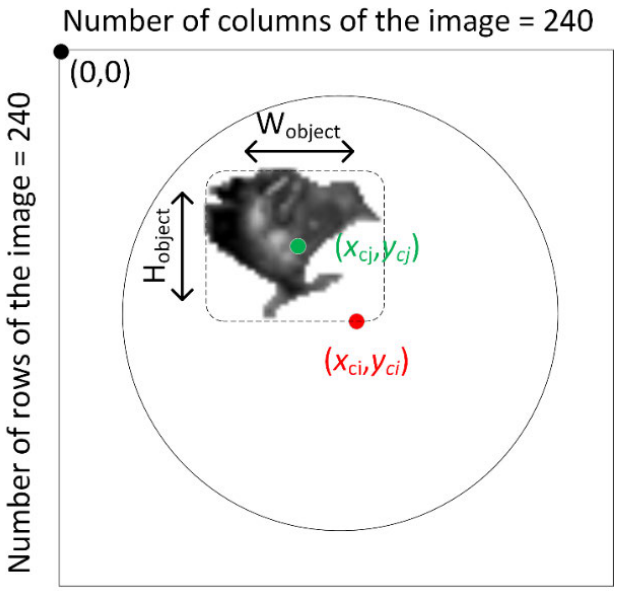

FIGURE 4. The coordinates of the center point of the ROI region obtained from the images in modalities are expressed by X_{cj}, Y_{cj}. The width of the ROI is represented by W_{object} and the height is represented by H_{object}.

IV. METHODS

In the literature, the UNet model is one of the most popular architectures preferred for biomedical image segmentation. One of the main reasons why the UNet model is preferred is its ability to extract local and global features at varying scales. For this reason, the UNet model is first mentioned in this study. Then, ResUNet+ model is presented in detail by making improvements to this model. In the next step, the implementation of the application is shared in detail. In the last step, the performance criteria used to measure the success of the systems are presented. Thus, the difference between the proposed model and the inspired model is presented in a much clearer way.

A. UNET

UNet is one of the most widely used CNN-based segmentation models in the literature for biomedical semantic image segmentation. The main reason behind the widespread use of UNet is its ability to extract local and global features at varying scales. In addition, it can transfer the feature map of each encoder level to the decoder level with skip connections. As a result, the pixel-based classifier can increase its performance level by processing the low-level layer, which contains information about the edges, and the upper-level layer, which contains information about the low-level layer, together. In image segmentation based on the CNN model, automatic feature extraction is performed in the first layers. In the next layers, segmentation masks are created by classifying the analyzed pixels according to objects or backgrounds. In the CNN-based segmentation approach, the image is divided into small patches.

To estimate the class labels of these patches, CNN is applied to each patch. Therefore, patch-based image segmentation is performed in the CNN-based approach. One of the leading CNN-based segmentation models used in the literature is the UNet segmentation model.

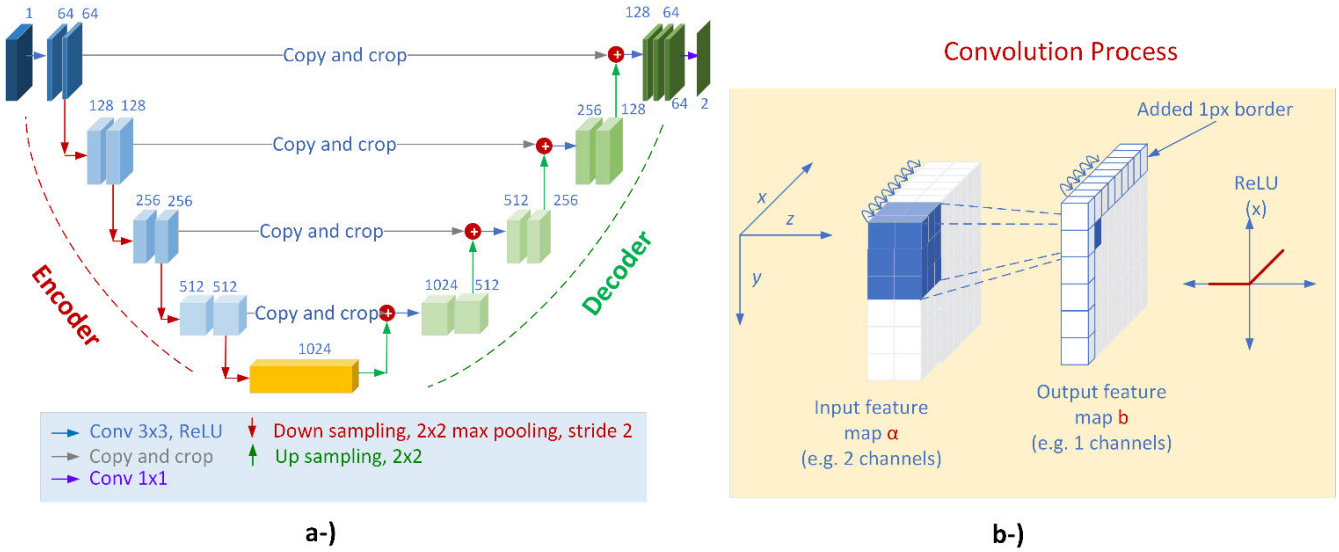

FIGURE 5. a-) UNet architecture developed by Ronneberger et al. [19]. b-) The input images are transferred to the encoder structure by convolutions as shown in Figure 5 (b) . After applying different convolution operations in the encoder structure, it is transferred both directly to the decoder structure at the same level itself and a lower layer by downsampling. In the decoder structure of UNet, both the data from the lower layer by upsampling and the data from the encoder are processed to produce the result.

The UNet segmentation model, as shown in Figure 5 (a), consists of an expanding and contracting path called encoder and decoder. The overall structure of the network resembles a classical convolution-based architecture but without a fully connected layer. Starting from the first layer, the 3×3 convolution process is repeatedly applied to the incoming image. It is then passed to the ReLU activation function. This process is shown in Figure 5 (b) and (3).

$$b_{x,y,l} = \text{ReLU}\left(\sum_{\substack{i\in\{-1,0,1\}\\ j\in\{-1,0,1\}\\ c\in\{1,\ldots,C\}}} w_{i,j,k,l}.\alpha_{x+i,y+j,k} + c_l\right) \quad (3)$$

In equation (3), w is the weight of the matrix used in the convolution process, α is the input feature matrix, c is the bias value, and b is the newly created convolution matrix. The 2×2 max pooling process with a stride value of 2 is applied to the value obtained as a result of (3). This process is called downsampling. Repeating the downsampling process down to the lowest layer is called the encoder process. In the encoder stage, for each downsampling operation, the image size is halved and the feature channels are doubled. In the decoder block, the compressed image obtained from the encoder stage is converted back to its original size by 2×2 upsampling operations. In the upsampling process, the features from the bottom layer are generally combined with the features from the layers shown in Figure 5 (b). The decoder block and encoder block are similar in structure but differ in that only attribute fusion and max pooling is not applied to the decoder layer. In addition, in the output layer, 1×1 convolution is performed to map the feature vector to the desired number of classes. The UNet architecture created in this way consists of 23 convolution layers in total [19]. The UNet model, while providing great innovations in segmentation, is computationally costly and inefficient due to overlapping patches in the image [44]. Therefore, there is a need to develop new segmentation models to improve the performance of the UNet model in challenging segmentation tasks.

B. PROPOSED MODEL

Although the UNet model is used in the literature for brain segmentation, it has some limitations. One of the main ones is the gradient vanishing problem that occurs as the network gets deeper. Because as the network gets deeper, the gradient becomes almost zero in the lower layers [45]. In addition, when creating segmentation masks in the UNet model, low-level features and high-level features that contain information about the boundary, edges, and location of the tumor region are equally important. In this case, the influence of important features may decrease whereas the value of unimportant features may increase. This has a negative impact on segmentation success. In order to improve performance metrics in brain segmentation, the gradient vanishing problem arising from the classical UNet structure needs to be addressed.

In general, whereas skipping connections between the encoder and decoder is important for data transfer, downsampling needs to be done more deeply in order to access more discriminative features. As a result, local details from previous layers may be lost due to convolutional operations.

To address these issues, the ResUNet+ model is proposed. The ResUNet+ model differs from other UNet-based models in that some modifications are made in the encoder stage of the model to preserve low-level features and to solve the gradient vanishing problem. For this, residual blocks are used. Residual blocks are one of the structures that can be used to preserve low-level features and solve the gradient vanishing problem. For this reason, residual blocks are preferred in

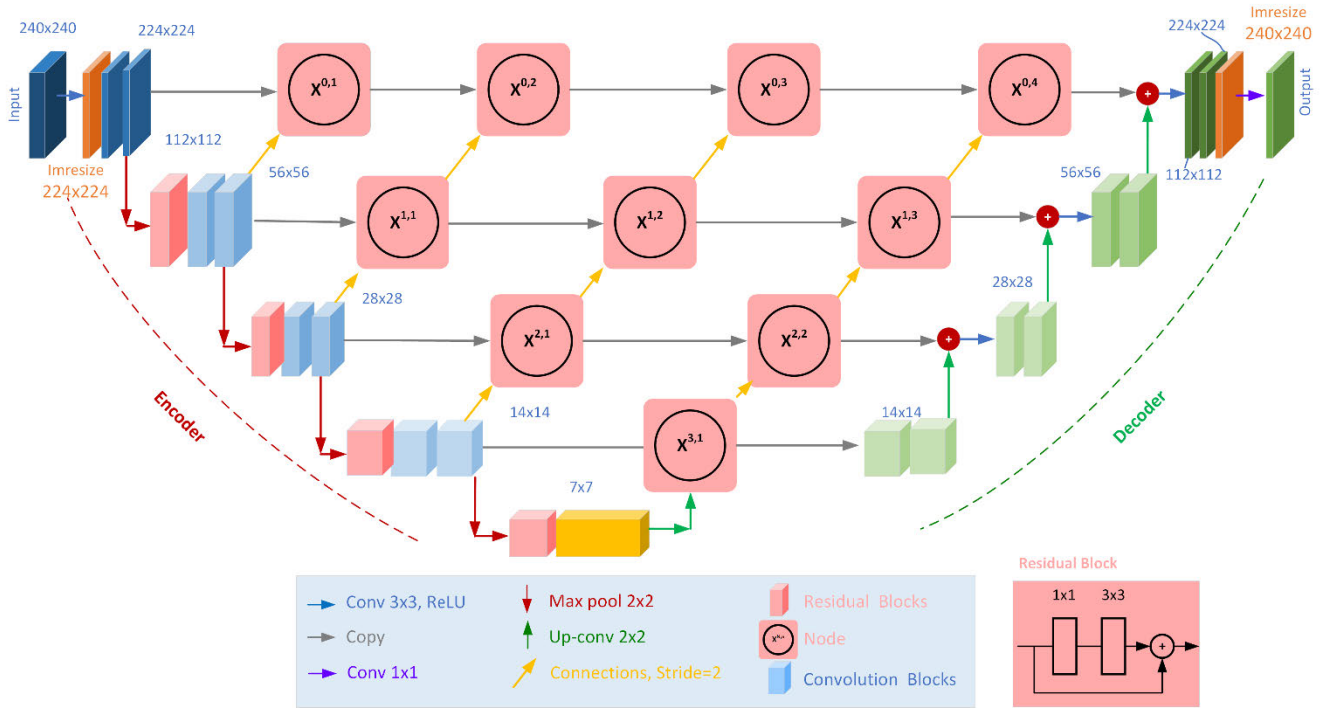

FIGURE 6. Images are first resized and then transferred to the input of the encoder structure. The proposed model structure that produces the output data with the help of residual, node, and attention block structures in the encoder-decoder.

the encoder structure of the proposed model presented in Figure 6. In addition, there are semantic gaps between the encoder and decoder in the classical UNet structure [46]. To solve this problem, unlike other UNet models, new connection nodes are added between the encoder and decoder.

At each node between the encoder-decoder layers in Figure 6, the data at the same level as itself and the values from a lower layer are combined. Also, if Figure 6 is examined carefully, besides the node structures having a tight connection with each other, the value in a node structure goes to also other nodes through different nodes. As a result, the semantic gap between the encoder and decoder is filled significantly. Thus, the created structure is transformed into a hybrid model.

In the proposed model, residual blocks containing 1×1 and 3×3 convolution operations, which are presented in Figure 6, respectively, are added to each encoder layer in order to prevent the features that can be lost as a result of convolutional operations that are generally used in UNet models. With this process, the features are strengthened a little more before the convolution process is applied in each layer.

However, there may still be lost features. In order to prevent this, nodes were added between the encoder and decoder, and these nodes were evaluated with the features coming from a lower layer. By means of the nodes used in this way, the features have been moved from the lowest layer to the top layer. In the notation of the nodes presented in Figure 6, the first digit represents the number of layers, and the second digit represents the node row number in that layer. As it can be understood from here, the number of nodes in the upper layers is more than in the lower layers. The reason for this is that while stronger features are obtained towards the lower layers, there are predominantly features in the upper layers, especially in the border regions of the image, which are highly distinctive but can be lost by convolution operations due to their sensitive structure. More node structures are used in the upper layers to prevent the loss of features that can be useful among these highly distinctive sensitive features. Another reason is that the features already obtained from the lower layers are moved to the upper layer with the decoder structure. With the presented model, especially the gradient vanishing problem, the gap and feature loss between the encoder and decoder have been prevented. In this case, it positively affected the performance of the ResUNet+ model.

C. IMPLEMENTATION DETAIL

The study was performed using Matlab 2020b on a workstation with Xeon processor (2X E5-2650), 64 GB RAM, 64 MB Cache, CUDA above 11.0, CuDNN 8.8.1, 6 GB GPU (1060 NVIDIA) and 64-bit operating system. BraTS 2018, 2019, and 2020 datasets were used in the training phase of the proposed model and UNet model. Training of both UNet and ResUNet+ models was started with random weights without using pre-trained weights.

The data was divided into two parts 80% training and 20% testing according to the cross-validation 5 value. The images in each modality are $240 \times 240 \times 155$ in size. The third dimension here refers to the slice value. When the images used in the proposed model are evaluated, segmenting the brain tumor structure with only single-modality images is not

TABLE 2. Some basic parameters of the proposed model.

Hyperparameters	Empirical ranges	Best value
Input size	-	240 × 240 × 155 × 4
Normalize initialization weight	Random, 1	Random
Learning rate	0.001 – 0.0000001	0.0001
Batch size	2-10	4
Activation function	ReLU, Tanh	ReLU
Optimizer	RMSProp, SGD, Adamax, Nadam, AdaGrad, Adam	Adam
Loss function	Dice, Jaccard	Dice
No. of epochs	1000 - 8000	4000
Output size	-	240 × 240 × 155 × 1

successful enough since the boundary between WT, ET, TC, and edema regions and healthy tissue is unclear [45]. For this reason, in this study, firstly, pre-processing was performed on the images to identify the relevant areas from all images and then images with four different modalities were used. It is important to emphasize that each modality has a significant effect on brain tumor segmentation. When these effects are examined, the T2 modality is used to identify edematous tumor regions, the T1 modality is used to distinguish healthy tissues, and the T1ce modality is used to make tumor borders visible. In addition, the FLAIR modality is also useful in distinguishing edema regions from cerebrospinal fluid [47].

Therefore, how to effectively integrate the feature information of multimodal images is a hot topic in tumor segmentation [48]. In this respect, the proposed model fills an important gap in the pre-processing process in the literature. This is because as shown in Figure 7, images with different modalities are considered as separate inputs. These images are then subjected to normalization. At the end of the normalization process, the pre-processing process described in Section III-B was performed. The pre-processing obtains the ROI of the region where the tumor is located. After obtaining the ROI, the rest of the processing was performed in five steps as shown in Figure 7.

In the first step, the ROI region of the T1ce modality is used. In the second step, the FLAIR and T2 modalities are multiplied element-wise and then normalized again and used as input for the proposed model. In the third step, the ROI region of the T1 modality is used as input to the proposed model. There is one attention module in the fourth step. The task of this module is to combine the images from different modalities in steps first, second, and third. Thus, instead of extracting feature information for all tumor regions through a single modality, feature information is extracted by using the T1 modality for non-enhancing tumors, FLAIR and T2 modality for enhancing tumors, and T1 modality for edema, and these features are combined.

In the fifth step, the three-segmented regions are found together on the resulting image. The segmentation success was then measured by using evaluation metrics on the images.

In order to determine the best combination of hyperparameters used in the proposed model, a series of experimental studies, were carried out. At the beginning of the experimental work, small filters were used to collect a lot of local information. Then the size of these filters was gradually increased. The learning rate was gradually decreased, starting with the largest value. This empirical approach aims to approach the global minimum point. The hyperparameter values used during the training of the model are generally presented in Table 2.

D. EVALUATION MEASURE

Brain tumor segmentation is commonly used in the literature to segment the WT region [34]. However, brain tumor segmentation is concerned with separating the WT region from the sub-regions of the tumor, namely the TC and ET regions. There are only a limited number of studies in the literature on this topic. Therefore, the study focuses on this topic.

As seen in the literature [49], [53] articles use dice, jaccard, sensitivity, specificity, and accuracy performance metrics. On the other hand, accuracy, dice, jaccard, precision, and recall metrics are used in [54] and [56] articles. In both groups of studies, the authors emphasize the importance of dice, jaccard, and accuracy metrics. As seen in both groups, similar and different evaluation metrics were used. For this reason, dice, jaccard, sensitivity, specificity, accuracy, and precision evaluation metrics were used in the study.

Dice is obtained by dividing the similar pixels between the model prediction and the actual GT by the total pixels in both images. Dice represents the spatial overlap ratio between the binary images, which consists of values between 0 and 1 as a result of comparing the segmented result with the GT result [57]. In these values, 0 represents no match and 1 represents a perfect match. A and B are defined as the region predicted by the proposed system as tumor and GT region, respectively. According to these definitions, the formulas of dice and jaccard metrics are given in (4-5) and their graphical representation is presented in Figure 8 [50].

$$\text{Dice} = \frac{2\,|A \cap B|}{|A| + |B|} \tag{4}$$

$$\text{Jaccard} = \frac{|A \cap B|}{|A \cup B|} \tag{5}$$

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \tag{6}$$

$$\text{Specificity} = \frac{T_n}{T_n + F_p} \tag{7}$$

$$\text{Sensitivity} = \frac{T_p}{T_p + F_n} \tag{8}$$

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{9}$$

Represent the true positive (Tp), true negative (Tn), false positive (Fp), and false negative (Fn) values in (6-9). The performance of the proposed model was measured in two steps. In the first step, the evaluation values of the TC, ET,

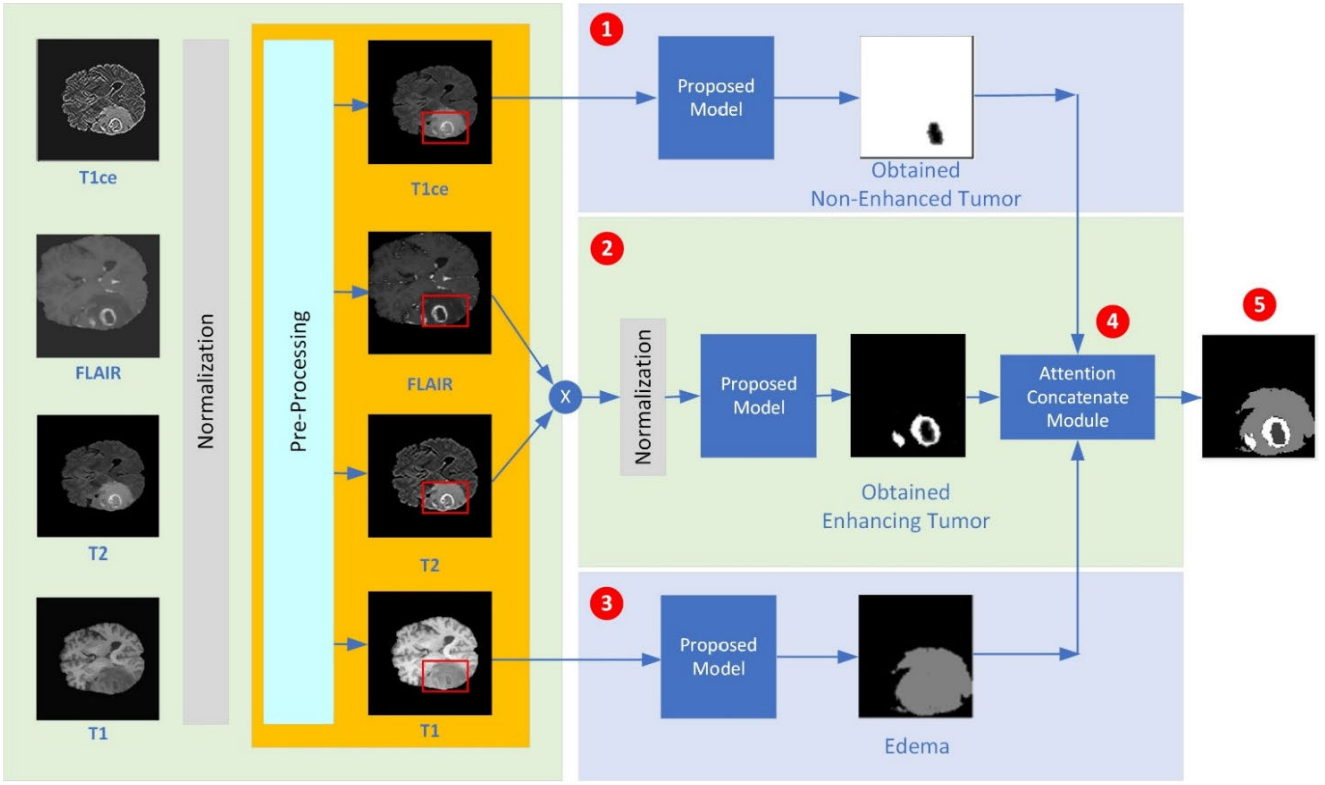

FIGURE 7. Images in four different modalities were first subjected to normalization. Then the ROI regions were detected by applying the pre-processing steps in Section III-B. The ROI regions in FLAIR and T2 modalities were multiplied on a pixel basis and normalization was applied. The proposed model is applied separately to the ROI values in T1 and T1ce modalities together with the ROI values obtained from here. Finally, the obtained values are transferred to the attention module, and the system generates output.

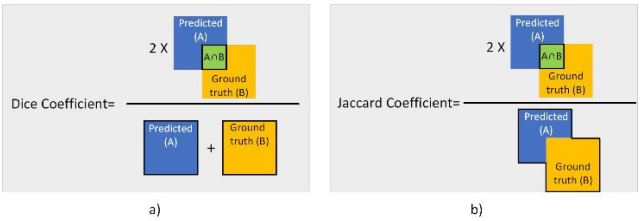

FIGURE 8. Dice and jaccard metrics.

and WT regions separately for each slice value are calculated with (4-5) as shown in Figure 8.

In the second step, the average dice, jaccard, and accuracy values of each slice were calculated using (7-9). In the second step, the performance results obtained in the first step were averaged.

V. RESULT AND DISCUSSION

In this study, the ResUNet+ model is proposed, inspired by the UNet architecture. This section, in which the results and analyses of the proposed model are presented, basically consists of three subsections. In the first subsection, the general performance results of the study carried out to determine the TC, ET, and WT regions in the brain with the segmentation process were shared in the BraTS 2020, BraTS 2019, and BraTS 2018 datasets. In the title of the ablation experiment in the second subsection, the effect of residual blocks, which will eliminate the semantic gaps between the encoder and decoder blocks, is discussed in detail. In the third subsection, the performance results were compared with state-of-the-art brain tumor segmentation approaches using the same dataset.

A. OVERALL PERFORMANCE ANALYSIS

Although the UNet segmentation model is generally used in the literature, residual blocks, and node structures are used to achieve higher segmentation performance in the study. Especially with the special node structure added between the encoder and the decoder, semantic gaps are minimized and interconnections are strengthened. In this approach the proposed model was created by taking the strengths of different segmentation models. The accuracy success for BraTS 2020 train images of the ResUNet+ model, in which four different modalities are used, is also presented in Figure 9. The overall loss value of the ResUNet+ model is also shared in Figure 9.

The success results of the proposed model and UNet on the same images allocated for the test in different datasets are presented in Figure 10. When the ground truth in Figure 10 and the images obtained from the two models are compared, the success of ResUNet+ on the same image is clearly seen.

The segmentation results of the proposed model on different BraTS datasets are presented visually in Figure 11. Evaluation metrics shared in Section IV-D were used in the

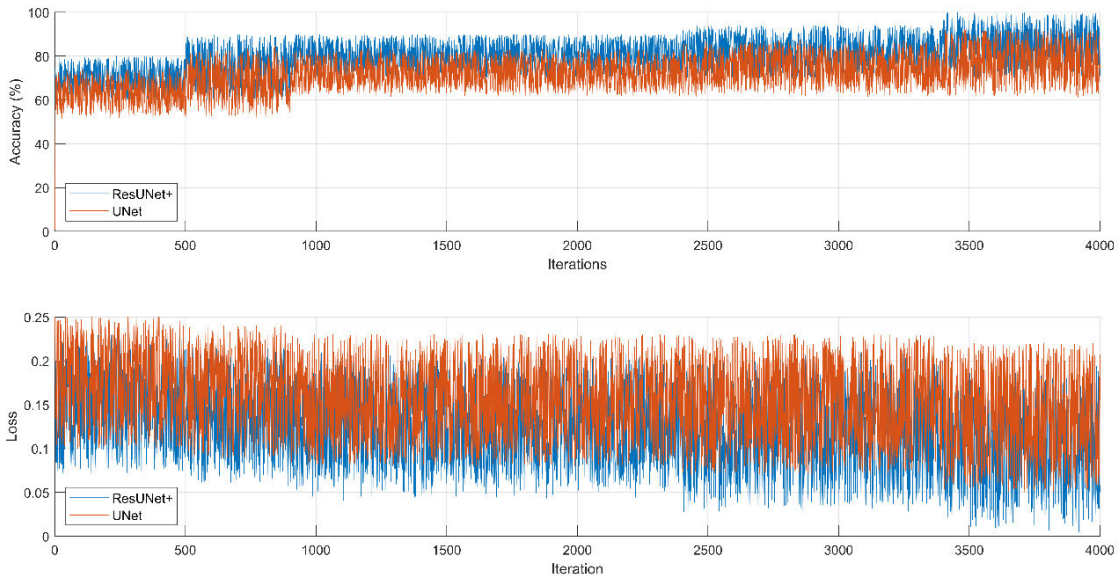

FIGURE 9. Accuracy and loss graph for the train data of the proposed model.

performance evaluation of the model. The training and testing performance of the inspired model were measured with the same evaluation metrics. The training performance results obtained are presented in Table 3. As can be seen from the results in Table 3, the training performance results of the proposed ResUNet+ model are higher than the other model in three different datasets used in the study.

The BraTS 2020 training dataset also has some additional data that is not available in BraTS 2019. For this reason, the dice, jaccard, and accuracy test results of TC, WT, and ET, especially for BraTS 2020, are shared in Figure 10. When dice values were examined, 91.90, 93.10, and 92.80, performance results were obtained for TC, ET, and WT for the BraTS 2020 dataset, respectively.

In the same graph, the accuracy and loss values of the underlying UNet model are also presented. As can be seen in Figure 9, the accuracy and loss values of the ResUNet+ model are higher than the inspired model. If the dynamics of the curves of both models are analyzed in general, the local minimum and maximum points can change during iteration. However, for the UNet model, even though the iteration value increases, the accuracy and loss values are bound to exhibit behavior in a constant range after approximately the 800th iteration. The ResUNet+ model starts to exhibit similar behavior in a fixed range after the 3400th iteration.

In Figure 11, non-enhanced tumor, enhanced tumor, and edema are denoted in blue, green, and red, respectively. The TC region is represented by the collection of enhanced and non-enhanced regions. The entire tumor region is obtained by summing the tumor core and edema regions [58].

As a result of the approach proposed in the BraTS 2020 dataset, in Table 3, details of the dice coefficient values of the training sets for TC, WT, and ET are given. Here, the

TABLE 3. Training performance results of the segmentation models used in the study on the dataset.

Methods	Datasets	Dice			Jaccard		
	BraTS	WT	ET	TC	WT	ET	TC
UNet	2020	90.3	88.3	88.9	88.7	86.1	87.4
	2019	88.9	87.5	88.6	87.8	89.3	88.5
	2018	88.3	89,1	88,1	87.9	87.4	87.1
ResUNet+	2020	95.6	94.8	94.7	95.1	93.9	94.1
	2019	95.1	94.2	94.3	94.7	93.4	93.9
	2018	94.6	93.7	93.8	94.1	92.9	93.1

BraTS 2020 training dataset contains some additional data that is not available in the BraTS 2019 and BraTS 2018 training sets [59].

B. ABLATION EXPERIMENT

The proposed ResUNet+ model has a depth level of four. There are residual blocks in the encoder part of the model, and there are nodes between the encoder and decoder. While experimenting with the ResUNet+ architecture, first of all, arrangements were made on the convolutional layers in the encoder section of the UNet model. In these arrangements, in addition to the convolution blocks used in the UNet model, the residual blocks in Figure 12 were added. In the study, three different residual block structures were tested. In particular, the block structure in Figure 12 (c) adversely affected both the processing time and the success.

When the block structure in Figure 12 (b) was changed, no remarkable success was achieved. An increase in performance was observed when Figure 12 (a) was used. After the residual block structure to be used was determined, it was applied in all layers. As a result of this process, it has been

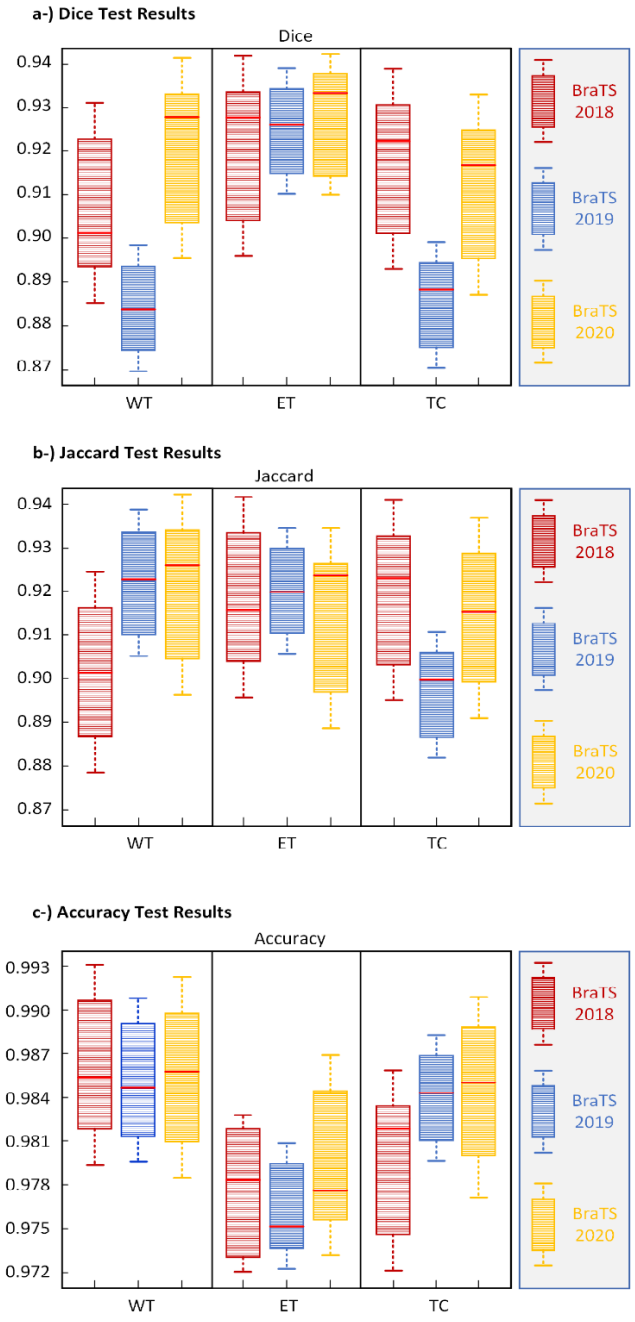

FIGURE 10. Shown the box plot of dice, jaccard, and accuracy for TC, ET, and WT of ResUNet+ model in the BraTS 2020, BraTS 2019, and BraTS 2018 test datasets. When analyzed for jaccard values, performance values of 91.32, 92.15, and 92.42 were obtained for TC, ET, and WT, respectively. Accuracy performance values were obtained as 98.46, 97.75, and 98.58, respectively. If a general evaluation is made in the BraTS 2020 dataset, it is seen that ET and NET values are determined with close accuracy.

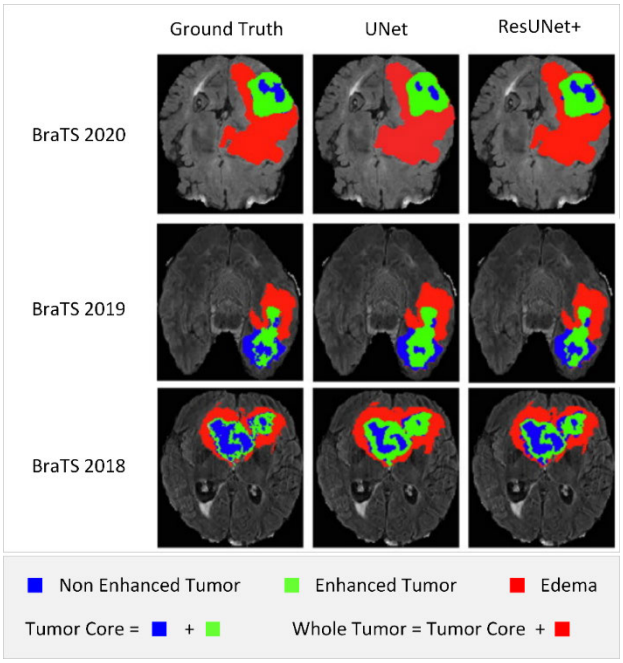

FIGURE 11. Four different MRI modalities and ground truth with corresponding manual segmentation annotation.

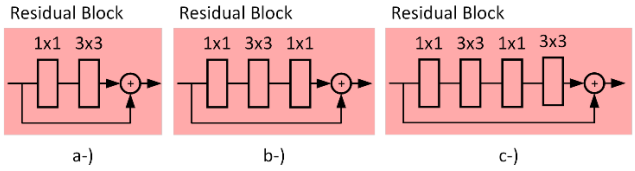

FIGURE 12. The different residual block structures for the ablation experiment.

determined that semantic gaps continue between the encoder-decoder layers. To solve this problem, a special node structure has been added between the encoder-decoder. In this node structure, the value from the convolution layer or node at the same level as itself and the value from the convolution layer or node in a lower layer is collected. At this stage, the collection of different node structures has also been tested, and the node structure that gives the highest performance is presented in the proposed model in Figure 6.

C. COMPARISON WITH STATE-OF-THE-ART STUDIES

In the literature review on brain tumor segmentation, it was determined that different evaluation metrics were used. In this section of the paper, only the evolution metrics used by the authors in the evaluation of SOTA studies are used. Since three different BraTS datasets were used in the study, the studies performed with the same datasets are presented in detail in Table 4 -VI. The (-) expression used in Tables 4-6 indicates that it does not contain a value for that metric.

The overall performance of the proposed architecture was measured with all evolution metrics presented in Section IV-D on the BraTS 2018 dataset. Thus, a more realistic comparison can be made with the studies in Table 4 using the same dataset in the literature. Some of the studies compared are;

Akbar et al. [53] developed a new model called MRAB for brain tumor segmentation by using atrous convolution and attention mechanisms to improve the segmentation performance of the UNet architecture. In the structure of the model, they proposed in their study, two atrous convolution

TABLE 4. Comparison of ResUNet+ model with state-of-the-art brain tumor segmentation methods on the BraTS 2018 datasets.

Author, Architecture	Evaluation Metrics	WT	ET	TC	Author, Architecture	Evaluation Metrics	WT	ET	TC
Akbar *et al.* [53], Multipath Residual Attention Block (MRAB)	Dice	79.77	89.59	77.71	Mehta and Arbel [61], 3D UNet	Dice	77.10	87.10	70.60
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	99.27	99.79	99.76		Specificity	-	-	-
	Sensitivity	92.57	81.37	80.45		Sensitivity	-	-	-
	Precision	-	-	-		Precision	-	-	-
Sun *et al.* [60], The multi-pathway architecture method	Dice	89.98	77.12	79.46	Wang *et al.* [30], 3D UNet	Dice	86.4	73.400	76.60
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	99.48	99.83	99.84		Specificity	-	-	-
	Sensitivity	90.39	76.87	75.09		Sensitivity	-	-	-
	Precision	-	-	-		Precision	-	-	-
Ranjbarzadeh *et al.* [34], Multipath Residual Attention Block	Dice	87.26	92.03	91.13	Rezai *et al.* [62], voxel-Gan	Dice	84.00	63.00	79.00
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	-	-	-		Specificity	99.00	99.00	99.00
	Sensitivity	93.86	92.17	97.12		Sensitivity	86.00	74.00	78.00
	Precision	-	-	-		Precision	-	-	-
Ours, UNet	Dice	88.10	88.30	87.20	**ResUNet+**	Dice	**90.10**	**92.70**	**92.30**
	Jaccard	87.72	87.23	86.83		Jaccard	90.01	91.24	92.12
	Accuracy	92.32	92.57	93.25		Accuracy	98.56	97.86	98.37
	Specificity	94.25	93.91	94.84		Specificity	99.86	99.95	99.42
	Sensitivity	93.92	93.25	92.12		Sensitivity	98.42	97.53	97.65
	Precision	93.41	92.35	92.87		Precision	98.85	96.27	97.12

sequence blocks are combined with an attention block. In the ResUNet+ model proposed in this paper, unlike Akbar's work, residual blocks and skip connection structures are used to solve the semantic gap and gradient loss between the encoder and decoder.

Sun et al. [60] determined the boundaries of the segmentation regions by using a different skip connection structure to prevent the loss of features during pooling layers and convolution operations. The layered architecture used has a fusion part in its structure. The extracted features are combined in this step. In the ResUNet+ model, fusion is applied to the segmentation results after pre-processing as shown in Figure 7. Thus, different regions on different modality images are combined and presented to the user.

Ranjbarzadeh et al. [34] focused on ROI regions in their studies. This aspect also inspired the ResUNet+ model. They state that they use a strong pre-processing method in their work and therefore develop a less complex architecture. Although a similar strategy is used in the ResUNet+ model, a stronger segmentation method is proposed. In this case, TC has a high impact on the success of classification.

Mehta and Arbel [61] classified multimodal MR images with a simple CNN model in their 3DUNet model. However, it achieved low performance in ET and TC. The main reason for this is the architecture used. Since the ResUNet+ model proposed in this study is a more comprehensive model, there is not a big difference between training and test results and it shows superior performance.

Wang et al. [30] developed a simpler network structure for each sub-task instead of using a single network structure for multiple classifications. In this way, they claim that they have solved the overfitting problem by facilitating the training of the model they developed. On the other hand, their model can confuse the WT and ET regions as they themselves state. The fact that they use different network structures in their studies has been a point of inspiration during the development of the ResUNet+ model. The ResUNet+ structure was used in three different arms in parallel, as shown in the general structure of the study in Figure 7. This had a positive effect on the performance results.

Rezaei et al. [6] developed a method to solve the problem of pixel level imbalance that occurs when the majority of pixels belong to the healthy region and only a few pixels belong to the tumor region. They used a segmented network structure in their method. However, there is a significant difference between the specificity and sensitivity performance values of the model they developed. This situation definitely affected the accuracy value of their study. However, no accuracy value was presented in their study. In the ResUNet+ model, the focus is directly on the segmentation and classification of WT, ET, and TC regions. Therefore, especially the dice score of the ResUNet+ model showed a much higher performance than this study.

TABLE 5. Comparison of ResUNet+ model with state-of-the-art brain tumor segmentation methods on the BraTS 2019 datasets.

Author, Architecture	Evaluation Metrics	WT	ET	TC	Author, Architecture	Evaluation Metrics	WT	ET	TC
Akbar *et al.* [53], Multipath Residual Attention Block (MRAB)	Dice	80.98	88.48	74.91	Ahmad *et al.* [7], RD^2A	Dice	90.35	71.93	82.34
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	99.18	98.82	99.63		Specificity	99.44	99.81	99.78
	Sensitivity	92.25	77.21	82.71		Sensitivity	91.57	80.09	80.80
	Precision	-	-	-		Precision	-	-	-
Zhang *et al.* [8], AGResU-Net	Dice	87.00	70.90	77.70	Sun *et al.* [60], Fully Convolutional Networks (FCN)	Dice	88.99	76.11	77.87
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	-	-	-		Specificity	99.51	99.81	99.68
	Sensitivity	-	-	-		Sensitivity	88.31	76.70	76.24
	Precision	-	-	-		Precision	-	-	-
Ours, UNet	Dice	88.60	88.90	87.50	**ResUNet+**	Dice	**88.30**	**92.50**	**88.90**
	Jaccard	88.32	88.91	87.33		Jaccard	92.38	92.01	90.29
	Accuracy	94.34	94.21	94.71		Accuracy	98.51	97.55	98.41
	Specificity	94.71	95.99	95.92		Specificity	99.90	99.87	99.90
	Sensitivity	93.57	93.48	92.63		Sensitivity	98.37	97.21	97.23
	Precision	94.13	92.54	93.82		Precision	98.25	96.28	97.44

The ResUNet+ model gives the highest performance in terms of dice values, which is the common performance measure in all studies, as can be seen from these comparisons and from Table 4 for the BraTS 2018 dataset. Similarly, the sensitivity values are higher than in other studies. However, according to Rezeai et al. [62], the specificity performance of the ResUNet+ model is competitive. As seen in Table 4, evaluation metric values such as jaccard, accuracy, and precision are not given in most of the studies given for comparison.

In this study, SOTA studies using the same dataset with the values obtained from the BraTS 2019 dataset are presented in Table 5. The first study presented in Table 5 is also again by Akbar et al. [53]. They obtained similar results with the BraTS 2018 dataset. Along with Akbar et al. [53], Sun et al. [60] also obtained similar results on the BraTS 2018 and BraTS 2019 datasets.

Zhang et al. [63] developed the model called AGResU-Net in their study. In their model, they examine the effectiveness of the attention gate on a classical UNet model. They also added an attention mechanism to this structure. As they stated, the model they developed can lose some context information and local details between different slices. In the ResUNet+ model, the context information is further strengthened thanks to the connection used. As a result, the detection of local details can be achieved with high success. This is especially reflected in the classification results of ET and TC.

Ahmad et al. [64] developed a model called Residual-Dilated Dense Atrous-Spatial Pyramid Pooling(RD^2A) to segment relatively small brain tumor regions. They aimed to preserve the contextual information of small tumors in the encoder stage.

They used dense connections to reduce the parameters. They had problems in training with large patch sizes on the BraTS dataset. Therefore, they used relatively smaller patch sizes. No large patch problem was encountered in the training of the BraTS dataset with the ResUNet+ model. As a result, higher performance was obtained especially in ET and TC classification compared to Ahmad et al. [64].

The proposed ResUNet+ model provides higher success in the detection of ET and TC regions than other models on the BraTS 2019 dataset. In addition, WT detection produces results at a level that can compete with Sun et al. [60].

In Table 6 presents the SOTA studies using the BraTS 2020 dataset and the performance evaluation metrics of these studies. Raza et al. [45] tested the DresUNet model they used for the BraTS 2018 dataset on the BraTS 2020 dataset and obtained similar results.

Ullah et al. [51] developed a new automatic model for the segmentation of brain tumor regions using multiscale residual attention-UNet (MRA-UNet). The model they developed also, were used three consecutive slices as input. Although it is claimed that progress has been made in the detection of TC and ET regions, the results of the ResUNet+ model provide higher performance values. The main reason for this is that the ResUNet+ model uses slice values in four modalities at the same time.

Akbar et al. [53] tested the MRAB model on BraTS 2018 and BraTS 2019 datasets and also tested it on BraTS 2020 dataset. The proposed ResUNet+ model gave better results than MRAB on all three datasets, especially in the dice evaluation performance metric.

TABLE 6. Comparison of ResUNet+ model with state-of-the-art brain tumor segmentation methods on the BraTS 2020 datasets.

Author, Architecture	Evaluation Metrics	WT	ET	TC	Author, Architecture	Evaluation Metrics	WT	ET	TC
Raza *et al.* [45], Deep Residual UNet (DresUNet)	Dice	83.57	86.6	80.04	Ma *et al.* [66], Deep Supervision CNN	Dice	78.79	87.26	75.93
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	99.36	97.91	98.62		Specificity	-	-	-
	Sensitivity	97.86	96.48	96.95		Sensitivity	-	-	-
	Precision	-	-	-		Precision	-	-	-
Ullah. *et al.* [51], Multiscale Residual Attention UNet (MRA-UNet)	Dice	87.22	90.18	86.74	Soltaninejad *et al.* [67], MR Encoder-Decoder	Dice	78.0	66.0	70.0
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	93.19	92.61	95.64		Specificity	100	100	100
	Sensitivity	92.71	93.54	91.88		Sensitivity	78.0	70.0	74.0
	Precision	-	-	-		Precision	-	-	-
Akbar *et al.* [53], Multipath Residual Attention Block (MRAB)	Dice	80.19	88.58	72.91	Savadikar *et al.* [65], Probabilistic UNet	Dice	79.88	72.49	77.71
	Jaccard	-	-	-		Jaccard	-	-	-
	Accuracy	-	-	-		Accuracy	-	-	-
	Specificity	99.86	99.97	99.94		Specificity	99.79	99.96	99.96
	Sensitivity	92.76	73.71	81.75		Sensitivity	87.19	75.13	78.13
	Precision	-	-	-		Precision	-	-	-
Ours, UNet	Dice	88.90	90.30	88.30	**ResUNet+**	Dice	**92.80**	**93.10**	**91.90**
	Jaccard	88.53	89.13	88.27		Jaccard	92.42	92.15	91.32
	Accuracy	94.54	93.65	94.53		Accuracy	98.58	97.75	98.46
	Specificity	94.74	95.96	95.84		Specificity	99.96	99.94	99.72
	Sensitivity	93.63	93.25	92.32		Sensitivity	98.46	97.42	96.93
	Precision	94.21	92.23	93.51		Precision	98.17	96.33	97.41

Savadikar et al. [65] developed a probabilistic UNET model by utilizing different numbers of attention block structures to improve the segmentation quality. They claim that increasing the number of attentions in their model gives better segmentation results. Instead of the attention block structure used in their proposed model, the node and attention structures used in the ResUNet+ structure can provide much higher success on the same dataset, especially for the dice evaluation metric.

Ma et al. [66] developed a deep supervision-based 2D residual model (a deep supervision-based 2D residual UNet) to alleviate the gradient distribution caused by the network depth. They claim that their model improves training stability. However, the encoder-decoder is directly connected to each other. As a result of any no module in between the encoder-decoder. Since this type of connection can cause some problems, a special node structure is used in the ResUNet+ model. The effect of using the node structure is directly reflected in the evaluation values of WT, ET, and TC regions.

Soltaninejad et al. [67] developed a model based on the encoder-decoder structure to combine low and high-resolution contextual information in MRIs. In their model, they obtained the system output by combining the features from two different encoder-decoder structures. Although two encoder-decoder structures were used in their model, an attention or node structure was not used to eliminate the semantic gaps in the encoder-decoder structures as stated in the literature. In this case, especially compared to the ResUNet+ model, lower performance was obtained in the detection of the TC region.

The ResUNet+ model proposed in the study outperformed its closest competitor with approximately 5.58 points in WT detection, approximately 5.16 points in TC detection, and approximately 2.92 points in ET detection for the dice metric in the BraTS 2020 dataset. In addition, the specificity and sensitivity values obtained from the ResUNet+ model are capable of the level competitive with the studies in the literature.

Unlike the SOTA studies presented above, there are some also studies that have been carried out to optimize the features extracted by CNN-based methods. One of these studies was performed by Khan et al. [68]. In the first step of their study, they developed a fusion-based technique to increase the contrast of tumors. In the second step, they used an active contour-based method to enhance the segmentation region. In their study, they extracted features by fine-tuning them EfficientNetB0. The extracted features were selected as the best features using the dragonfly optimization algorithm and classified using an extreme learning machine (ELM). In their study, they obtained 94.89%, 95.14%, and 95.94% success rates on the modalities in BraTS 2018, BraTS 2019, and BraTS 2020 datasets, respectively. In their study, unlike the proposed ResUNet+ model, they evaluated the classes as T1,

T1ce, T2, and FLAIR. In the ResUNet+ model, T1, T1ce, T2, and FLAIR images were evaluated as MRI modalities in parallel with the literature [45], [51], [69]. WT, ET, and TC regions were used as tumor classes in the ResUNet model.

In SOTA studies, automatic feature extraction features of CNN-based algorithms are generally used extensively [44], [59], [70], [71]. On the other hand, there are applications where feature extraction is performed by traditional manual systems. Rajinikanth et al. [72] resized the images obtained in step one. They performed tumor segmentation with VGG-UNet in the second step, deep features extraction with VGG-16 in the third step, handcrafted feature extraction in the fourth step, and selected the best features. In the fifth step, they used the firefly algorithm. In the last step, the serial feature has completed the application with concatenation and binary classification. If the proposed ResUNet+ model and this study are evaluated together, the ResUNet+ model first segments and then classifies WT, ET, and TC regions over images in four different modalities. Rajinikanth et al. [72], on the other hand, classify healthy or unhealthy based on T2 modality images.

The weights of all the studies presented in Tables 4, 5, and 6 in the literature are initialized random values without any pre-training as stated by the authors. The proposed model differs in its architecture and evaluation metrics, as can be seen from the SOTA studies. Tested on BraTS 2020, 2019, and 2018 datasets, the ResUNet+ model performed better than a lot of pre-trained models.

The accuracy or jaccard values of many studies in the literature are not shared as evaluation metrics. However, the metric values of the dice have been shared in many studies in the literature. In some studies, in addition, only specificity and sensitivity values were shared. In order to make a realistic comparison with the obtained values, the accuracy value must be given in the studies. In addition, as seen in Table 4, very different dice values can be obtained even when two very similar UNet models are tested on the same dataset by different authors ([30], [61], OursUNet). The main reason for this is the pre-processing methods used to obtain TC, WT, and ET regions. This directly shows the effect of pre-processing on performance. Due to the correct pre-processing method used in the study, even the classical UNet model gives better results than many SOTA methods derived from it.

VI. CONCLUSION

The main topic of this study is the development of a new segmentation model with high-performance criteria on MRI. In order to improve the segmentation performance of the existing segmentation models in the literature, a new segmentation model called ResUNet+ is proposed, inspired by residual blocks and node structures.

In the proposed model, the first step is focused on the extraction of the ROI region. At this stage, the ROI region was determined by using four different modalities in a hybrid manner. This method used in the extraction of the ROI region differs from many studies in the literature. Thus, unnecessary convolutions on the entire image were avoided, and the focus was only on the ROI region in the study. The impact of this procedure on success is also shown for the classical UNet model on different BraTS datasets in Tables 4, 5, and 6. As can be seen here, the success of the UNet model, which uses or does not use the proposed pre-processing, is also demonstrated by different metrics. The model proposed in this study requires a much longer training time than the inspired UNet model. However, this is tolerable due to the superior performance metrics and the importance of human health. However, gradient vanishing and exploding problems are frequently encountered in UNet-based deep learning segmentation applications. In the study, the gradient exploding problem was solved with the nodes placed between the encoder and decoder. With the residual blocks added to the encoder block, a structure that can prevent the gradient vanishing has been created. Thus, each layer of the model is connected to the next layer, and the encoder–decoder blocks at the same level are more strongly connected to each other. The model proposed in the study and the compared UNet model were analyzed using the cross-validation 5 value on the BraTS 2020, 2019, and 2018 datasets. Unlike the experimental test results of the ResUNet+ model, the training performance results are also presented in detail in Table 3. The ResUNet+ model outperformed both the underlying UNet model and many studies in the literature using the same dataset.

In the BraTS 2020 dataset, it is about 5.58 points ahead of its closest competitor in the dice metric for WT detection, about 5.16 points for TC detection, and about 2.92 points for ET detection. Similar results were obtained in other datasets used in the study. Although the study focuses specifically on the dice metric, other evaluation performance metrics are also shared so that they can be easily compared with other studies in the future.

In the future, the proposed method can be applied to different segmentation datasets and can be also used in different decision support systems to assist experts in the health sector. At the same time, RestfulAPIs can be designed that can be used on any platform by expert health personnel. In future studies, new ResUNet+ based algorithms that can perform segmentation on different biomedical datasets will continue to be created.

The ResUNet+ model also has a few minor limitations. The most important of these is the training time and the number of epochs. For this, the model parameters can be optimized in the future. Another limitation of the study is that the ResUNet+ model and the pre-processing steps used are T1, T1ce, T2, and FLAIR modalities. Different modalities models can be used in the future.

ACKNOWLEDGMENT

The authors would like to thank IEEE Access journal and their respective universities for their support.

REFERENCES

[1] H. Çetiner, "Citrus disease detection and classification using based on convolution deep neural network," *Microprocess. Microsyst.*, vol. 95, Nov. 2022, Art. no. 104687, doi: 10.1016/j.micpro.2022.104687.

[2] A. Tiwari, S. Srivastava, and M. Pant, "Brain tumor segmentation and classification from magnetic resonance images: Review of selected methods from 2014 to 2019," *Pattern Recognit. Lett.*, vol. 131, pp. 244–260, Mar. 2020, doi: 10.1016/j.patrec.2019.11.020.

[3] E.-E. M. Azhari, M. M. M. Hatta, Z. Z. Htike, and S. L. Win, "Tumor detection in medical imaging: A survey," *Int. J. Adv. Inf. Technol.*, vol. 4, no. 1, p. 21, Feb. 2014.

[4] S. Al-Qazzaz, "Deep learning-based brain tumour image segmentation and its extension to stroke lesion segmentation," Ph.D. dissertation, School Eng., Cardiff Univ., Cardiff, U.K., 2020. [Online]. Available: https://orca.cardiff.ac.uk/id/eprint/134897/

[5] R. Guldager, P. V. Hansen, and M. Ziebell, "Past, present and future, the experience of time during examination for malignant brain tumor: A qualitative observational study," *Acta Neurochirurgica*, vol. 163, no. 4, pp. 959–967, Apr. 2021, doi: 10.1007/s00701-020-04693-z.

[6] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: A summary," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, Jun. 2016, doi: 10.1007/s00401-016-1545-1.

[7] K. Niu, Z. Guo, X. Peng, and S. Pei, "P-ResUnet: Segmentation of brain tissue with purified residual Unet," *Comput. Biol. Med.*, vol. 151, Dec. 2022, Art. no. 106294, doi: 10.1016/j.compbiomed.2022.106294.

[8] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016, doi: 10.1109/TMI.2016.2548501.

[9] V. S. Khoo, D. P. Dearnaley, D. J. Finnigan, A. Padhani, S. F. Tanner, and M. O. Leach, "Magnetic resonance imaging (MRI): Considerations and applications in radiotherapy treatment planning," *Radiotherapy Oncol.*, vol. 42, no. 1, pp. 1–15, Jan. 1997, doi: 10.1016/S0167-8140(96)01866-X.

[10] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.

[11] L. G. Nyul, J. K. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 143–150, Feb. 2000, doi: 10.1109/42.836373.

[12] L. P. Clarke, R. P. Velthuizen, M. Clark, J. Gaviria, L. Hall, D. Goldgof, R. Murtagh, S. Phuphanich, and S. Brem, "MRI measurement of brain tumor response: Comparison of visual metric and automatic segmentation," *Magn. Reson. Imag.*, vol. 16, no. 3, pp. 271–279, 1998, doi: 10.1016/S0730-725X(97)00302-0.

[13] W. Zhang, G. Yang, H. Huang, W. Yang, X. Xu, Y. Liu, and X. Lai, "ME-Net: Multi-encoder net framework for brain tumor segmentation," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 4, pp. 1834–1848, Dec. 2021, doi: 10.1002/ima.22571.

[14] H. Mzoughi, I. Njeh, A. Wali, M. B. Slima, A. BenHamida, C. Mhiri, and K. B. Mahfoudhe, "Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification," *J. Digit. Imag.*, vol. 33, no. 4, pp. 903–915, Aug. 2020, doi: 10.1007/s10278-020-00347-9.

[15] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, vol. 58, no. 13, pp. R97–R129, Jul. 2013, doi: 10.1088/0031-9155/58/13/R97.

[16] X. Liu, A. Yu, X. Wei, Z. Pan, and J. Tang, "Multimodal MR image synthesis using gradient prior and adversarial learning," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1176–1188, Oct. 2020, doi: 10.1109/JSTSP.2020.3013418.

[17] H. Lu, A. H. Kashani, K. Arfanakis, A. Caprihan, C. DeCarli, B. T. Gold, Y. Li, P. Maillard, C. L. Satizabal, L. Stables, D. J. J. Wang, R. A. Corriveau, H. Singh, E. E. Smith, B. Fischl, A. Kouwe, K. Schwab, K. G. Helmer, and S. M. Greenberg, "MarkVCID cerebral small vessel consortium: II. Neuroimaging protocols," *Alzheimer's Dementia*, vol. 17, no. 4, pp. 716–725, Apr. 2021, doi: 10.1002/alz.12216.

[18] S. Metlek, "A new proposal for the prediction of an aircraft engine fuel consumption: A novel CNN-BiLSTM deep neural network model," *Aircr. Eng. Aerosp. Technol.*, vol. 95, no. 5, pp. 838–848, Mar. 2023, doi: 10.1108/AEAT-05-2022-0132.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[20] J. Walsh, A. Othmani, M. Jain, and S. Dev, "Using U-Net network for efficient brain tumor segmentation in MRI images," *Healthcare Anal.*, vol. 2, Nov. 2022, Art. no. 100098, doi: 10.1016/j.health.2022.100098.

[21] S.-T. Tran, C.-H. Cheng, T.-T. Nguyen, M.-H. Le, and D.-G. Liu, "TMD-Unet: Triple-Unet with multi-scale input features and dense skip connection for medical image segmentation," *Healthcare*, vol. 9, no. 1, p. 54, 2021, doi: 10.3390/healthcare9010054.

[22] Y. Weng, T. Zhou, Y. Li, and X. Qiu, "NAS-Unet: Neural architecture search for medical image segmentation," *IEEE Access*, vol. 7, pp. 44247–44257, 2019, doi: 10.1109/ACCESS.2019.2908991.

[23] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, and A. Madabhushi, Eds. Cham, Switzerland: Springer, 2018, pp. 3–11, doi: 10.1007/978-3-030-00889-5_1.

[24] Y. Cao, S. Liu, Y. Peng, and J. Li, "DenseUNet: Densely connected UNet for electron microscopy image segmentation," *IET Image Process.*, vol. 14, no. 12, pp. 2682–2689, Oct. 2020, doi: 10.1049/iet-ipr.2019.1527.

[25] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2663–2674, Dec. 2018.

[26] S. Qamar, H. Jin, R. Zheng, P. Ahmad, and M. Usama, "A variant form of 3D-UNet for infant brain segmentation," *Future Gener. Comput. Syst.*, vol. 108, pp. 613–623, Jul. 2020, doi: 10.1016/j.future.2019.11.021.

[27] J. Zhang, X. Lv, H. Zhang, and B. Liu, "AResU-Net: Attention residual U-Net for brain tumor segmentation," *Symmetry*, vol. 12, no. 5, p. 721, 2020, doi: 10.3390/sym12050721.

[28] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Lwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A full-scale connected Unet for medical image segmentation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 1055–1059, doi: 10.1109/ICASSP40776.2020.9053405.

[29] J. Zhang, Q. Qin, Q. Ye, and T. Ruan, "ST-Unet: Swin Transformer boosted U-Net with cross-layer feature enhancement for medical image segmentation," *Comput. Biol. Med.*, vol. 153, Feb. 2023, Art. no. 106516, doi: 10.1016/j.compbiomed.2022.106516.

[30] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation," *Frontiers Comput. Neurosci.*, vol. 13, p. 56, Aug. 2019.

[31] J. Zhuang, "LadderNet: Multi-path networks based on U-Net for medical image segmentation," 2018, *arXiv:1810.07810*.

[32] B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.

[33] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features," *Sci. Data*, vol. 4, no. 1, Sep. 2017, Art. no. 170117, doi: 10.1038/sdata.2017.117.

[34] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghoushchi, S. Anari, M. Naseri, and M. Bendechache, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Sci. Rep.*, vol. 11, no. 1, p. 10930, May 2021, doi: 10.1038/s41598-021-90428-8.

[35] J. Mohan, V. Krishnaveni, and Y. Guo, "A survey on the magnetic resonance image denoising methods," *Biomed. Signal Process. Control*, vol. 9, pp. 56–69, Jan. 2014, doi: 10.1016/j.bspc.2013.10.007.

[36] S. Amutha, R. D. R. Babu, R. Shankar, and H. N. Kumar, "MRI denoising and enhancement based on optimized single-stage principle component analysis," *Int. J. Adv. Eng. Technol.*, vol. 5, no. 2, p. 224, 2013.

[37] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: Speed and effectiveness," in *Proc. 1st Conf. Visualizat. Biomed. Comput.*, Atlanta, Georgia, May 1990, pp. 337–345.

[38] M. F. Stollenga, W. Byeon, M. Liwicki, and J. Schmidhuber, "Parallel multi-dimensional LSTM, with application to fast biomedical, volumetric image segmentation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, no. 1, 2015, pp. 1–9.

[39] W. Chen, B. Liu, S. Peng, J. Sun, and X. Qiao, "Computer-aided grading of gliomas combining automatic segmentation and radiomics," *Int. J. Biomed. Imag.*, vol. 2018, Oct. 2018, Art. no. 2512037.

[40] L. Novamizanti, A. L. Prasasti, and I. F. Noor Kiranda, "Comparison of discrete cosine transform and dual-tree complex wavelet transform based on arithmetic coding in medical image compression," *J. Phys., Ser.*, vol. 1367, no. 1, Nov. 2019, Art. no. 012021, doi: 10.1088/1742-6596/1367/1/012021.

[41] W. Liu, S. Jiang, and H. Li, "Experimental study of liquid-carrying by swirling flow in a U-shaped tube," *Experim. Thermal Fluid Sci.*, vol. 130, Jan. 2022, Art. no. 110479, doi: 10.1016/j.expthermflusci.2021.110479.

[42] G. Chen, F. Li, J. Geng, P. Jing, and Z. Si, "Identification, generation of autoclaved aerated concrete pore structure and simulation of its influence on thermal conductivity," *Construct. Building Mater.*, vol. 294, Aug. 2021, Art. no. 123572, doi: 10.1016/j.conbuildmat.2021.123572.

[43] X. Hu, X. Zhu, and Z. Sun, "Development of an SEM image analysis method to characterize intumescent fire retardant char layer," *Prog. Organic Coat.*, vol. 139, Feb. 2020, Art. no. 105461, doi: 10.1016/j.porgcoat.2019.105461.

[44] V.-T. Pham, T.-T. Tran, P.-C. Wang, P.-Y. Chen, and M.-T. Lo, "EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images," *Artif. Intell. Med.*, vol. 115, May 2021, Art. no. 102065, doi: 10.1016/j.artmed.2021.102065.

[45] R. Raza, U. Ijaz Bajwa, Y. Mehmood, M. Waqas Anwar, and M. Hassan Jamal, "dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 103861, doi: 10.1016/j.bspc.2022.103861.

[46] Y. Lu, X. Qin, H. Fan, T. Lai, and Z. Li, "WBC-Net: A white blood cell segmentation network based on UNet++ and ResNet," *Appl. Soft Comput.*, vol. 101, Mar. 2021, Art. no. 107006, doi: 10.1016/j.asoc.2020.107006.

[47] A. M. Hasan, F. Meziane, R. Aspin, and H. A. Jalab, "Segmentation of brain tumors in MRI images using three-dimensional active contour without edge," *Symmetry*, vol. 8, no. 11, p. 132, Nov. 2016, doi: 10.3390/sym8110132.

[48] L. Fang and X. Wang, "Multi-input Unet model based on the integrated block and the aggregation connection for MRI brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104027, doi: 10.1016/j.bspc.2022.104027.

[49] A. M. G. Allah, A. M. Sarhan, and N. M. Elshennawy, "Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 118833, doi: 10.1016/j.eswa.2022.118833.

[50] N. Cinar, A. Ozcan, and M. Kaya, "A hybrid DenseNet121-UNet model for brain tumor segmentation from MR images," *Biomed. Signal Process. Control*, vol. 76, Jul. 2022, Art. no. 103647, doi: 10.1016/j.bspc.2022.103647.

[51] Z. Ullah, M. Usman, M. Jeon, and J. Gwak, "Cascade multiscale residual attention CNNs with adaptive ROI for automatic brain tumor segmentation," *Inf. Sci.*, vol. 608, pp. 1541–1556, Aug. 2022, doi: 10.1016/j.ins.2022.07.044.

[52] M. Raja and S. Vijayachitra, "A hybrid approach to segment and detect brain abnormalities from MRI scan," *Expert Syst. Appl.*, vol. 216, Apr. 2023, Art. no. 119435, doi: 10.1016/j.eswa.2022.119435.

[53] A. S. Akbar, C. Fatichah, and N. Suciati, "Single level UNet3D with multipath residual attention block for brain tumor segmentation," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 6, pp. 3247–3258, Jun. 2022, doi: 10.1016/j.jksuci.2022.03.022.

[54] A. Iqbal, M. Sharif, M. A. Khan, W. Nisar, and M. Alhaisoni, "FF-UNet: A U-shaped deep convolutional neural network for multimodal biomedical image segmentation," *Cognit. Comput.*, vol. 14, no. 4, pp. 1287–1302, Jul. 2022, doi: 10.1007/s12559-022-10038-y.

[55] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1240–1251, May 2016, doi: 10.1109/TMI.2016.2538465.

[56] V. Groza, B. Tuchinov, E. Pavlovskiy, E. Amelina, M. Amelin, S. Golushko, and A. Letyagin, "Data preprocessing via multi-sequences MRI mixture to improve brain tumor segmentation," in *Bioinformatics and Biomedical Engineering*, I. Rojas, O. Valenzuela, F. Rojas, L. J. Herrera, and F. Ortuño, Eds. Cham, Switzerland: Springer, 2020, pp. 695–704, doi: 10.1007/978-3-030-45385-5_62.

[57] R. Almajalid, J. Shan, Y. Du, and M. Zhang, "Development of a deep-learning-based method for breast ultrasound image segmentation," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, 2018, pp. 1103–1108.

[58] Y. Liu, J. Du, C.-M. Vong, G. Yue, J. Yu, Y. Wang, B. Lei, and T. Wang, "Scale-adaptive super-feature based MetricUNet for brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103442, doi: 10.1016/j.bspc.2021.103442.

[59] Z. Xiao, K. He, J. Liu, and W. Zhang, "Multi-view hierarchical split network for brain tumor segmentation," *Biomed. Signal Process. Control*, vol. 69, Aug. 2021, Art. no. 102897, doi: 10.1016/j.bspc.2021.102897.

[60] J. Sun, Y. Peng, Y. Guo, and D. Li, "Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN," *Neurocomputing*, vol. 423, pp. 34–45, Jan. 2021, doi: 10.1016/j.neucom.2020.10.031.

[61] R. Mehta and T. Arbel, "3D U-Net for brain tumour segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11993, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2020, pp. 254–266, doi: 10.1007/978-3-030-46643-5.

[62] M. Rezaei, H. Yang, and C. Meinel, "voxel-GAN: Adversarial framework for learning imbalanced brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. van Walsum, Eds. Cham, Switzerland: Springer, 2019, pp. 321–333, doi: 10.1007/978-3-030-11726-9_29.

[63] J. Zhang, Z. Jiang, J. Dong, Y. Hou, and B. Liu, "Attention gate ResU-Net for automatic MRI brain tumor segmentation," *IEEE Access*, vol. 8, pp. 58533–58545, 2020, doi: 10.1109/ACCESS.2020.2983075.

[64] P. Ahmad, H. Jin, S. Qamar, R. Zheng, and A. Saeed, "RD2A: Densely connected residual networks using ASPP for brain tumor segmentation," *Multimedia Tools Appl.*, vol. 80, no. 18, pp. 27069–27094, Jul. 2021, doi: 10.1007/s11042-021-10915-y.

[65] C. Savadikar, R. Kulhalli, and B. Garware, "Brain tumour segmentation using probabilistic U-Net," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 255–264, doi: 10.1007/978-3-030-72087-2_22.

[66] S. Ma, Z. Zhang, J. Ding, X. Li, J. Tang, and F. Guo, "A deep supervision CNN network for brain tumor segmentation," *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 158–167, doi: 10.1007/978-3-030-72087-2_14.

[67] M. Soltaninejad, T. Pridmore, and M. Pound, "Efficient MRI brain tumor segmentation using multi-resolution encoder-decoder networks," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi and S. Bakas, Eds. Cham, Switzerland: Springer, 2021, pp. 30–39, doi: 10.1007/978-3-030-72087-2_3.

[68] M. A. Khan, A. Khan, M. Alhaisoni, A. Alqahtani, S. Alsubai, M. Alharbi, N. A. Malik, and R. Damaševičius, "Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm," *Int. J. Imag. Syst. Technol.*, vol. 33, no. 2, pp. 572–587, Mar. 2023, doi: 10.1002/ima.22831.

[69] Z. Zhu, X. He, G. Qi, Y. Li, B. Cong, and Y. Liu, "Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI," *Inf. Fusion*, vol. 91, pp. 376–387, Mar. 2023, doi: 10.1016/j.inffus.2022.10.022.

[70] Y. Peng and J. Sun, "The multimodal MRI brain tumor segmentation based on AD-Net," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104336, doi: 10.1016/j.bspc.2022.104336.

[71] J. Sun, W. Chen, S. Peng, and B. Liu, "DRRNet: Dense residual refine networks for automatic brain tumor segmentation," *J. Med. Syst.*, vol. 43, no. 7, p. 221, Jul. 2019, doi: 10.1007/s10916-019-1358-6.

[72] V. Rajinikanth, S. Kadry, and Y. Nam, "Convolutional-Neural-Network assisted segmentation and SVM classification of brain tumor in clinical MRI slices," *Inf. Technol. Control*, vol. 50, no. 2, pp. 342–356, Jun. 2021.

SEDAT METLEK received the B.S. degree in computer and control teacher from Marmara University, Istanbul, in 2005, the B.S. degree in computer engineering from Burdur Mehmet Akif Ersoy University, in 2022, the M.S. degree in electronic computer education from Süleyman Demirel University, Isparta, in 2009, and the Ph.D. degree in electronic communication engineering from Süleyman Demirel University, Isparta, in 2015.

From 2006 to 2009, he was a computer teacher at Burdur. From 2009 to 2016, he was a Lecturer at Mehmet Akif Ersoy University. Since 2016, he has been an Assistant Professor with the Mechatronic Department, Burdur Mehmet Akif Ersoy University, Vocational School of Technical Sciences.

He is the author of five books (chapters), and more than 20 articles. His research interests include deep learning, artificial intelligence and applications, image segmentation-processing, and special machine design and manufacture for specific subjects.

HALİT ÇETİNER received the B.S. degree from the Department of Computer Engineering, Selçuk University, the M.S. degree from the Department of Computer Engineering, Süleyman Demirel University, and the Ph.D. degree in computer engineering from the Department of Computer Engineering, Süleyman Demirel University. He continues his academic research at Isparta University of Applied Sciences. He has many studies and articles in the field of computer vision, natural language processing, and artificial intelligence.