IRDNU-Net: Inception residual dense nested u-net for brain tumor segmentation

Nagwa M. AboElenein[1,2] · Piao Songhao[1] · Ahmed Afifi[2,3]

Abstract

Accurate segmentation of brain tumors is an essential stage in treatment planning. Fully convolutional neural networks, specifically the encoder-decoder architectures such as U-net, have proven successful in medical image segmentation. However, segmenting brain tumors with complex structure requires building a deeper and wider model which increases the computational complexity and may also cause the gradient vanishing problem. Therefore, in this work, we propose a novel encoder-decoder architecture, called Inception Residual Dense Nested U-Net (IRDNU-Net). In this model carefully designed Residual and Inception modules are used in place of standard U-Net convolutional layers to increase the width of the model without increasing the computational complexity. Additionally, in the proposed architecture, the encoder and decoder are connected via a sequence of Inception-Residual densely nested paths to extract more information and increase the depth of the network while reducing the number of network parameters. The proposed segmentation architecture was evaluated on two large brain tumor segmentation benchmark datasets; the BraTS'2019 and BraTS'2020. It achieved a mean Dice similarity coefficient of 0.888 for the whole tumor region, 0.876 for the core region, and 0.819 for the enhancement region. Experimental results illuminate that IRDNU-Net outperforms U-Net by 1.8%, 11.4%, and 11.7% in the whole tumor, core tumor, and enhancing tumor, respectively. Moreover, the IRDNU-Net enables a great improvement on the accuracy compared to comparative approaches, and its ability in the face of challenging problems, such as small tumor regions, with fewer parameters.

✉ Piao Songhao
  piaosh@hit.edu.cn

  Nagwa M. AboElenein
  nagwa.salim1@hit.edu.cn

  Ahmed Afifi
  aafifi@kfu.edu.sa

1   School of Computer Science and Technology, Harbin Institute of Technology,
    Harbin, 150001, China

2   Faculty of Computers and Information, Menoufia University, Menoufia, 32511, Egypt

3   Department of Computer Science, College of Computer Science and Information Technology,
    King Faisal University, P.O. Box 400, Al-Ahsa, 31982, Saudi Arabia

1 Introduction

Brain tumors are among the most deadly cancers worldwide [30]. The most prominent brain tumor [4] are Gliomas. Gliomas are graded into High-Grade (HGG) and Low-Grade (LGG) gliomas. Based on the tumor's pathological evaluation gliomas, comprise numerous sub-areas of heterogeneous histology, namely enhancing, edema, and necrotic core tumors. Magnetic Resonance Imaging (MRI) is one of the most important tools for assessing gliomas as it can provide a lot of information about the tumor structure. The typical phases of MRI screening [21] include Fluid-attenuated reversal, T1-weighted, T1-weighted contrast-enhanced, and T2-weighted. Accordingly, brain tumor segmentation from multi-modality MRI is essential to evaluate the tumor aggressions, and responsiveness to glioma treatment, and it has beneficial applications for brain tumor diagnosis, tracking, and treatment. However, manual segmentation of certain brain tumors is vulnerable to human mistakes and it is a time-consuming task. There is also a lack of reproducibility, which harms patient management's effectiveness and can result in ineffective therapy [31]. On the other hand, automatic brain tumor segmentation can provide a more effective solution [29]. The progress of deep learning in this area in recent years [10, 23] is immense, as it is considered the state-of-the-art in classification, segmentation, and detection applications [25]. Convolutional neural networks (CNNs) [3] specifically are one of the most popular techniques for building efficient segmentation approaches as they can automatically learn the most useful and relevant features. However, accurate tumor segmentation remains a difficult task due to the heterogeneous appearance and multiple types of brain tumors, as well as the high variability of brain tumor size, shape, position, and intensity, and contrast in different imaging modalities [34]. Therefor, in this paper we introduce a novel Inception Residual Dense Nested U-Net (IRDNU-Net) for solving the insufficient precision of small-scale tumors with fewer numbers of parameters. Experiments on two brain tumor segmentation datasets demonstrate that IRDNU-Net method enables a significant improvement in accuracy compared to other models, especially for small tumors, with less computational complexity. The main contributions of this study can be summarized as follows:

- Based on the U-Net architecture, we propose an efficient brain tumor segmentation approach called IRDNU-Net. It can extract more representative features from brain tumors which enhances the segmentation accuracy specially for small tumors.
- In IRDNU-Net, to make the network structure wider, the standard convolutional layers used in the U-Net architecture are replaced by carefully designed Residual Inception modules. The IRDNU-Net encoder and decoder sub-networks are linked by many nested dense paths to increase the depth of the network as well.
- We evaluate the proposed architecture using two brain tumor segmentation datasets, Brats 2019 and Brats 2020. The experimental results indicate that the proposed automated segmentation approach is accurate and computationally efficient.

Following this introduction, Section 2 presents a brief survey of the related work. Section 3 includes the details of the proposed architecture. In Section 4, experiments and results are introduced in detail. Discussion is provided in Section 5, and conclusion is summarized in Section 6.

2 Related work

Utilizing deep learning for brain tumor segmentation has attracted increasing attention. Perira et al. [32] suggested a 2D CNN network with a deeper layer structure using small 3 x 3 kernels to develop an automated brain tumor segmentation approach, which ranked in the second position in the BraTS 2015 Challenge. Havaei et al. [16] constructed a cascaded architecture that combines local and global paths. However, their architectures lose spatial continuity and requires a large storage space which leads to low segmentation performance. Zhao et al. [39] presented their solution to the segmentation of brain tumors by combining fully convolutional (FCN) and Conditional Random Fields (CRFs), which achieved competitive performance with only three imaging modalities, rather than four imaging modalities. Using FCN architecture, Ronneberger et al. [33] constructed an asymmetric, fully convolutional network called U-Net, comprising a contracting path that extracts spatial image features and an expanding path that generates a segmentation map from the encoded features. U-Net was commonly used for medical image segmentation tasks [17, 28], and brain tumor segmentation is not an exception. Dong et al. [13] establish a 2D U-Net approach for automatic tumor segmentation, which utilizes the Dice loss function, they obtained equivalent results for the complete tumor region and better results for the core tumor regions. Cahall et al. [5] proposed a new image segmentation framework using a U-Net architecture with Inception modules to perform multi-scale feature extraction. They proposed a new loss function based on the modified Dice Similarity coefficient. Haichun Li et al. [26] presented a novel end-to-end approach for brain tumor segmentation, they proposed Inception module in each block to learn network, and an efficient cascade training method for segmenting sub regions however, the proposed approach suffers from a data imbalance. Cheng et al. [7] introduced a new Memory-Efficient Cascade 3D U-Net which uses fewer down-sampling channels to achieve high segmentation precision with less memory complexity. Ibtehaz et al. [20] proposed MultiRes blocks to build a more robust and reliable approach, their proposed approach are not perfect, but in most of the cases, it outperforms the classical U-Net by a moderate margin. Zhou et al. [41] presented the deep architecture supervision with re-designed skip pathways. However, the number of parameters, as well as the time to train the network, is significantly higher. Brain tumors are known to have complex forms and various sizes that contribute to the presence of small tumors. U-Net continuously reduces the image dimension during the down-sampling process, resulting in low segmentation accuracy for small-scale tumors. To further enhance its segmentation performance, several modules, such as the Multi residual module, Dense module [12], and Inception module [36] are added to the baseline model, which has facilitated the development of methods for brain tumor segmentation. Parvez et al. [2] introduced approach of residual-dense connections based on U-Net model, however, the final scores not improved enough. In general, although some recent models have achieved some level of competitive success, they still have limitations. For example, 1) they are ineffective at identifying smaller tumors, 2) the most advanced models need enormous computational resources to achieve high segmentation accuracy. To address these shortcomings, we developed a novel Inception Residual Dense Nested U-Net. The Inception-Residual block allows us to make the proposed network substantially wider, while residual connection makes the network easier to train. Meanwhile, nested dense paths can increase the depth of the network, optimize the network results, and minimize the computational complexity. In addition, compared to other state-of-the-art methods, our IRDNU-Net significantly can achieve comparable results for the whole tumor regions, and superior results for the core tumor regions, and enhancing tumor regions, with fewer parameters.

3 Material and methods

3.1 Datasets

The proposed segmentation approach is evaluated using two benchmark datasets; the BraTS 2019, and BraTS 2020 brain tumor MRI datasets which were published by the Multi-Modal Brain Tumor Segmentation Challenge (BraTS) [11]. BraTS 2019 train dataset contains 335 cases, 259 cases of HGG, and 76 cases of LGG while the BraTS 2019 validation dataset has 125 unlabeled cases. The BraTS2020 dataset is larger than the BraTS2019 dataset with a training set of 367 cases. The datasets are co-registered, re-sampled, and skull-stripped to 1 mm^3. The size of each MRI image is 240 × 240 × 155, and each case has FLAIR, T1-, T1-enhanced (T1c), and T2 volumes. Where the whole tumor (WT) region includes all intra-tumor regions, i.e., necrosis non-enhancing tumor, edema, and enhancing tumor, Tumor core (TC) region that incorporates non-enhancing tumor necrosis and enhancing tumor (ET) region. For training, 25,000 2D patches of size 128 × 128 × 4 are randomly sampled from each case that represents the four modalities. From the training dataset of BraTS 2019, 205 cases are used to form the 5,125,000 patches training set, and from BraTS 2020 we use 160 cases from the training dataset to form the 4,000,000 patches as shown in Table 1.

3.2 Methods

The proposed IRDNU-Net model for brain tumor segmentation depends on the same encoder-decoder architecture like the U-Net model. However, as shown in Fig. 1, we deepen and widen the encoder-decoder architecture by combining Inception-Residual modules and re-designing the skip pathways, namely dense skip connections are presented. This architecture will be used for brain tumor segmentation from four MRI modalities, namely; FLAIR, T1, T2, and T1c.

3.2.1 Inception Residual Dense Nested U-Net(IRDNU-Net)

U-Net structures are Encoder-Decoder architecture. The Encoder tries to gradually lower the spatial dimension of feature maps while capturing more high-level semantic characteristics. While The goal of the decoder is to restore object characteristics and spatial dimensions. Therefore, capturing more high-level characteristics in the encoder is important, and preserving more spatial information in the decoder to increase segmentation performance.

Table 1 Training data partitioning

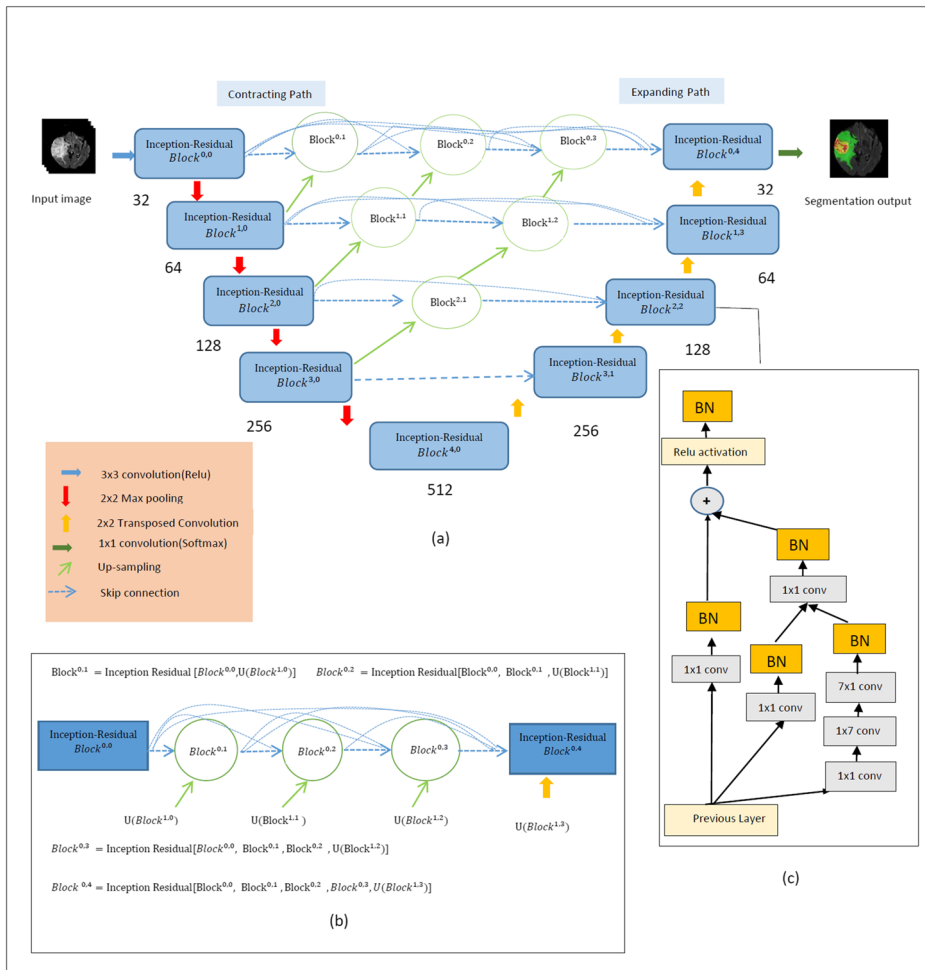|	BraTS 2019	BraTS 2020
Training	164	128
Validation	41	32
Testing	–	56
Number of training cases	205	160
Numbers of patches for training	4,100,000	3,200,000
Numbers of patches for validation	1,025,000	800,000
Numbers of patches for testing	–	–
Total numbers of patches	5,125,000	4,000,000

Fig. 1 (a) The proposed IRDNU-Net architecture. (b) Illustration of the first skip pathway with Details. (c) The schema for the Inception- Residual module

Based on the encoder-decoder architecture, Inception-Residual module, and the Dense model, Fig. 1a demonstrates the proposed architecture denoted as IRDNU-Net. The proposed network structure is composed of an encoding branch and a decoding branch. In each encoder block we substitute the two convolution layers in the original U-Net model with the proposed Inception-Residual block, as it will be described in the next section, followed by a 2 × 2 max-pooling operation for down-sampling. At each down-sampling step we double the number of feature channels. Correspondingly, the same number of up-sampling processes is performed in the decoding branch to restore the spatial size of the segmented output. Each up-sampling is implemented by a 2 × 2 transposed convolution, and the number of feature channels is halved. In U-Net, the feature maps of the encoder are directly received in the decoder; however, in our proposed model dense nested paths in addition to Inception-Residual blocks connect the encoder to the decoder. We allocate a parameter P for each of the Inception-Residual blocks to control the number of convolution layer filters

in these blocks to ensure a clear relation between the number of parameters of the baseline U-Net model and the proposed one. The value of P is calculated as follows:

$$P = F \times \alpha \tag{1}$$

Where F represents the number of the filters in the U-Net layer, and α represents a scalar coefficient.

Similar to the baseline U-Net, we set the number of filters, F, to 32, 64, 128, 256, and 512. To minimize the number of parameters in our model, we assign $\alpha = 1.8$ inside the Inception-Residual block, we assign $\left[\frac{P}{12}\right], \left[\frac{P}{6}\right], \left[\frac{P}{4}\right]$, and $\left[\frac{P}{2}\right]$ to the four consecutive convolutional layers, respectively. In our experiments, we found that this mixture produced good results: to balance between the parameters numbers reduction and improving segmentation accuracy. Like the U-Net architecture, it is also noted that P doubles or halves after each pooling or deconvolution process.

The Rectified Linear Unit (ReLU) function [15] is used for all covolutional layers in this network. The output layer is activated with the same sigmoid activation function as with the U-Net model. In Fig. 1, we show a diagram of the proposed Inception-Residual Dense Nested U-Net model. The details of our architecture are listed in Table 2.

Inception-Residual Block The Inception module [37] can experimentally enhance the visual representation. To increase the segmentation network's ability, we use a residual connection, due to its effectiveness in the segmentation of medical images [14]. The goal is to add feature maps from various kernels of different sizes that can expand the network and allow it to learn multi-scale features. Based on Inception-Residual module [36] the modified Inception-Residual U-Net denoted as (IResU-Net) is proposed in this work. Figure 1c shows the modified Inception-Residual block. The modified Inception-Residual module includes multiple sets of 1×1 convolutions, 7×1 convolutions, and 1×7 convolutions. Compared with the original Inception-Residual module. We first add a Batch normalization (BN) [22] layer after each convolutional layer to alleviate the gradient vanishing problem. Consequently, 1×1 convolutions are applied on the identity skip connections to preserve a similar relationship between the base U-Net number of filters and our suggested model. As it can be noticed from Fig. 1c, the output filters generated from the convolution layers in the right branch, 1×1, 1×7, 7×1, are concatenated with convolution layers in the middle branch, 1×1, then this output added with the convolution layers in the left branch, 1×1. Suppose that y_l is the l_{th} layer output. The $h_{n \times n}(\cdot)$ is a $n \times n$ kernel convolutional layer and $h_b(\cdot)$ indicates the BN layer. Where $f_r(\cdot)$ denotes the ReLU activation function, concatenation function is denoted by $\circ$, the output of each Inception-Residual module can be calculated as in (2):

$$y_{l+1} = h_b(f_r(h_b(h_{1 \times 1}(h_b(h_{1 \times 1}(y_l) \circ h_b(h_{7 \times 1}(h_{1 \times 7}(h_{1 \times 1}(y_l)))))) + h_b(h_{1 \times 1}(y_l)))) \tag{2}$$

Dense Nested Paths The encoder feature maps are directly obtained from the decoder in the original U-Net. While the main purpose of dense connection is making the network deeper and wider without causing the gradient vanishing problem. So, in IRDNU-Net, we also re-modeled skip pathways to enhance encoder-decoder network connectivity by using dense nested paths denoted as (DNU-Net) with Inception-Residual blocks. To facilitate and increase the model capacity for accurate brain tumor segmentation. We use three, two, and one Inception-Residual blocks in the four Dense Nesting paths, respectively; in the blocks of the four Nested paths, we use 32, 64, 128, and 256 filters, respectively as seen in Fig. 1a. Figure 1b shows the skip pathway between Inception-Residual Block^{0,0} and Inception-Residual Block^{0,4} which comprises a dense block of three Inception-Residual blocks. A

Table 2 Inception-residual dense nested U-Net architectural details

Inception-Residual Dense Nested U-Net		
Block	Filter size	Number of filters
Inception-Residual Block^{0,0}	Conv2D(1,1)	4
	Conv2D(1,1)	9
	Conv2D(1,7)	14
Inception-Residual Block^{0,4}	Conv2D(7,1)	28
	Conv2D(1,1)	55
Inception-Residual Block^{1,0}	Conv2D(1,1)	9
	Conv2D(1,1)	19
	Conv2D(1,7)	28
Inception-Residual Block^{1,3}	Conv2D(7,1)	57
	Conv2D(1,1)	113
Inception-Residual Block^{2,0}	Conv2D(1,1)	19
	Conv2D(1,1)	38
	Conv2D(1,7)	57
Inception-Residual Block^{2,2}	Conv2D(7,1)	115
	Conv2D(1,1)	229
Inception-Residual Block^{3,0}	Conv2D(1,1)	38
	Conv2D(1,1)	76
	Conv2D(1,7)	115
Inception-Residual Block^{3,1}	Conv2D(7,1)	230
	Conv2D(1,1)	459
	Conv2D(1,1)	76
	Conv2D(1,1)	153
Inception-Residual Block^{4,0}	Conv2D(1,7)	230
	Conv2D(7,1)	460
	Conv2D(1,1)	919

concatenation process merges the preceding Inception-Residual block production with the equivalent up-sampled output of the lower dense block in the same dense block before each Inception-Residual block. The skip pathway is formulated as follows: $b^{i,j}$ represent the Inception-Residual output. Considering i as the down-sampling layer on the encoder, and j as the dense block's Inception-Residual layer along the skip pathway, the $b^{i,j}$ is calculated according to:

$$b^{i,j} = \begin{cases} IR(b^{i-1,j}) & , j = 0. \\ IR\left(\left[[b^{i,k}]_{k=0}^{j-1}, U(b^{i+1,j-1})\right]\right) & , j > 0. \end{cases} \tag{3}$$

Where the function $IR(\cdot)$ is the Inception-Residual operation, $U(\cdot)$ represents the layer of up-sampling, and [ ] represents the layer of concatenation. The blocks at $j = 0$ only obtain one entry from the last encoder layer. Blocks receive two inputs at level $j = 1$, both from the two-succession level encoder network and blocks at $j > 1$ obtain j+1 inputs, which obtained from the previous j blocks on the same skip path, and the last input is the up-sampled output from the bottom skip path. Figure 1b illustrates (3) by displaying how the feature maps move along the top skip path of our proposed architecture.

3.3 Combined loss function

In a deep learning system, the loss function is crucial when we are dealing with extremely imbalanced data. A proper selection of the loss function may also improve the model accuracy. In this study, we utilize a combined loss function, defined in (4), that integrates the weighted Cross-Entropy loss (WCE), defined in (5) and Generalized Dice Loss (GDL) [35], defined in (6), to alleviate the effect of the class imbalance problem.

$$CL = WCE + GDL \tag{4}$$

$$WEC = \frac{-1}{k} \sum_{k} \sum_{i}^{L} w_i g_{ik} log(p_{ik}) \tag{5}$$

$$GDL = 1 - 2 \frac{\left(\sum_{i}^{L} w_i \sum_{i} g_{i_k} \ p_{i_k} \right)}{\sum_{i}^{L} w_i \sum_{i} (g_{i_k} + \ p_{i_k})} \tag{6}$$

Where L represents the total number of labels, k is the size of a batch, w_i is the weight of the label i^{th}. For the generalized dice loss, g_i and p_i represent the value of the pixel of the binary ground truth image and the binary segmented image.

4 Experiments and results

4.1 Experimental setup

All experiments were performed using the Keras framework with the TensorFlow back-end. We used the stochastic gradient-based (SGD) [9] as optimizer. A batch of size 8 is used. For 20 epochs, we have trained our model because the loss of validity did not change afterward. The momentum = 0.8, the initial learning rate = 0.0001 decay with a decay factor of 0.1. The training was carried out on an Intel Corei7 3.5GHz machine using NVIDIA GeForce GTX 1070.

4.2 Evaluation metrics

In this work, we utilized the Dice score, Sensitivity, and Specificity metrics to assess the segmentation results. The Dice similarity score primarily measures the overlap region between the segmented lesion and the ground truth segmentation. Sensitivity is often referred to as a true positive rate, and the Specificity is utilized to define the true negative rate, which can be calculated using the next Equations, respectively:

$$DSC = \frac{2TP}{(FP + 2TP + FN)} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TP}{TP + FP} \tag{9}$$

Where TP, FN, and FP denote true positive, false negative, and false positive.

Table 3 Segmentation results on Training BraTS 2019 Dataset

	DSC			Sensitivity			Specificity		
	Whole	Core	Enhancing	Whole	Core	Enhancing	Whole	Core	Enhancing
Mean	0.888	0.876	0.819	0. 883	0.869	0.857	0.994	0.998	0.997
Std.Dev.	0.091	0.154	0.171	0.130	0.169	0.152	0.005	0.003	0.005
Median	0.92	0.918	0.869	0.926	0.920	0.903	0.997	0.999	0.998
25 quantile	0.877	0.874	0.788	0.864	0.861	0.814	0.994	0.997	0.997
75quantile	0.947	0.939	0.911	0.960	0.957	0.943	0.999	0.998	0.963

4.3 Experimental results

Our experiments are divided into three parts, which are carried out on the BraTS 2019 training dataset, the BraTS 2019 validation dataset, and the BraTS 2020 testing dataset. Evaluation results of BraTS2019 training and validation dataset are disseminated on the challenge leaderboard website.[1] Meanwhile, the BraTS2020 testing dataset is individual runs.

4.3.1 Evaluation results on BraTS2019 training dataset

205 cases from the BraTS 2019 datasets are used in this experiment. 80% of the dataset (164 subjects) are used for model training and the remaining 20% (41 subjects) are used for validation as mentioned in Table 1. The evaluation results of the proposed IRDNU-Net on the BraTS 2019 training dataset are presented in Table 3. Quantitatively the proposed network achieved a DSC of 0.888 for the whole tumor, 0.876 for core tumor, and 0.819 for the enhancing tumor. The mean, standard deviation, the median, 25th and 75th percentiles for all metrics are shown in the Table 3. The proposed approach has been evaluated by DSC, Specificity, and Sensitivity, which are measured using the online evaluation system on the leaderboard BraTS 2019 online website.[2]

Ablation Study To study the effect of different modules and enhanced architectures, we perform an ablation study for the DNU-Net, IRU-Net, and IRDNU-Net models. The ablation study results are summarized in Table 4 for the dice similarity coefficient (DSC), where (DNU-Net), (IRU-Net), and (IRDNU-Net) are our enhanced models on U-Net. As can be noticed from Table 4, IRDNU-Net produces the most accurate segmentation results among the four models with an improvement of 1.8% on the whole tumor, 11.4% for core tumor, and 11.7% on enhancing tumor over the standard U-Net. It outperforms the DNU-Net by a ratio of 1.2% for the whole tumor, 4.3% for the core tumor, and 7.8% for enhancing tumor. Compared with IRU-Net, IRDNU-Net outperforms IRU-Net with gains of 0.6% on the whole tumor, 6.7%, and 5% accuracy improvement on the core tumor, and enhancing tumor, respectively.

Comparative Study The proposed IRDNU-Net is also compared to other related brain tumor segmentation approaches to assess its efficiency. This comparison is presented in

---

[1]https://ipp.cbica.upenn.edu

[2]https://www.cbica.upenn.edu/BraTS19/lboardTraining.html

Table 4 Ablation study on the training BraTS 2019 Dataset and comparison segmentation results with baselines

Methods	Whole	Core	Enhancing
$U - Net^*$	0.87	0.762	0.703
DNU-Net(our)	0.876	0.833	0.742
IRU-Net(our)	0.882	0.809	0.77
IRDNU-Net (our)	0.888	0.876	0.82

Table 5, IRDNU-Net outperforms other top networks in the DSC value for the core tumor and enhancing tumor but it is slightly lower than the approach proposed by li et al. [26], and Hu et al. [18] for the whole tumor. In li et al. [26], the network structure is optimized by designing and refining the U-Net architecture. K.Hu et al. [18] authors apply fully connected conditional random fields and multi-cascaded. In comparison to Zhang et al. [38] and Chen et al. [8] methods, the proposed IRDNU-Net model achieves enhanced segmentation efficiency. Zhao et al. [39] utilized conditional random field to increase efficiency. However, our IRDNU-Net achieves 0.8% on the whole tumor, 2.7%, on core tumor, and 4.9% for enhancing tumor achieving gains over them , without applying any post-processing strategy. Compared to the network developed in Memory Efficient Cascade 3D U-Net [7], Our IRDNU-Net outperforms this network for core tumor and enhanced tumor by a large margin of 5.6% and 5.4%.

By comparing the sensitivity metric, our IRDN U-Net achieves a sensitivity score of 0.883 for the whole tumor, 0.869 for the core tumor, and 0.857 for enhancing tumor segmentation. In particular, an optimum sensitivity score for core tumor segmentation is obtained. While the best sensitivity score for the whole and Enhancing tumor obtained by Hu et al. [18], the comparative sensitivity results, to a certain extent, indicate the efficacy of IRDNU-Net in the segmentation of small tumors. By comparing the specificity score, our IRDNU-Net achieved a specificity of 0.994 for the whole tumor, 0.998 for the core tumor, and 0.997 for the enhancing tumor. In general, the IRDNU-Net model can achieve competitive efficiency and outperforming other state-of-the-art techniques.

Table 5 Comparison of segmentation results on the BraTS 2019 Training Dataset with typical methods

Methods	DSC			Sensitivity			Specificity		
	Whole	Core	Enhancing	Whole	Core	Enhancing	Whole	Core	Enhancing
Cheng et al. [7]	0.89	0.811	0.765	0.898	0.816	0.769	0.994	0.994	0.997
Li et al. [26]	0.89	0.733	0.726	0.895	0.75	0.743	–	–	–
K. Hu et al. [18]	0.89	0.82	0.777	0.90	0.84	0.86	–	–	–
Chen et al. [8]	0.888	0.844	0.739	0.880	0.832	0.786	0.994	997	0.997
Zhao et al. [39]	0.88	0.84	0.77	0.86	0.82	0.80	–	–	–
Zhang et al. [38]	0.876	0.772	0.72	–	–	–	–	–	–
IRDNU-Net(our)	0.888	0.867	0.819	0.883	0.869	0.857	0.994	0.998	0.997

Table 6 Segmentation results on BraTS 2019 validation dataset

	DSC			Sensitivity			Specificity		
	Whole	Core	Enhancing	Whole	Core	Enhancing	Whole	Core	Enhancing
Mean	0.865	0.864	0.806	0.885	0.884	0.842	0.992	0.996	0.996
Std.Dev.	0.104	0.119	0.132	0.101	0.093	0.105	0.009	0.007	0.008
Median	0.897	0.907	0.842	0.923	0.906	0.855	0.994	0.999	0.998
25 quantile	0.865	0.843	0.772	0.851	0.849	0.80	0.991	0.997	0.997
75quantile	0.922	0.934	0.89	0.951	0.95	0.917	0.997	0.999	0.999

4.3.2 Evaluation results on BraTS2019 validation dataset

We use 66 validation cases from the validation to take part in the BraTS 2019 competition. Our algorithm's segmentation efficiency was calculated by using the online evaluation system for DSC, Specificity, and Sensitivity in the challenge leaderboard Web site.[3] These results are available on the leader-board section of these challenges under the title "Nagwasalim". The experimental results are shown in Table 6. Quantitatively, DSC is 0.865 for the whole tumor, 0.864 for the core tumor, and 0.806 for the enhancing tumor. The mean, standard deviation, the median, 25th and 75th percentiles of all metrics are also shown in Table 6.

Ablation Study Table 7 demonstrates a comparison of the segmentation results with baselines. Also, Table 8 demonstrates the comparative results with other standard approaches. The comparison results between U-Net, DNU-Net, IRU-Net, and IRDNU-Net has shown in Table 7 are identical to those in Table 4. In comparison, IRDNU-Net achieves higher performance than U-Net. While 0.1% and 11.8% increase on U-Net for the whole tumor, and core tumor segmentation, it increases the enhancing tumor by 11.2%, demonstrating its good effect on little tumor segmentation. IRDNU-Net outperforms DNU-Net 0.3%, 4.3%, and 6.5% on the whole, core, and enhancing tumors. It exceeds IRU-Net by 6.1% margin in core tumor segmentation. These comparative results define the ability of IRDNU-Net to segment brain tumors.

Comparative Study Table 8 demonstrates the suggested technique's effectiveness and other advanced techniques to the 66 validation datasets; IRDNU-Net provides highly competitive performance relative to other advanced brain tumor segmentation approaches. IRDNU-Net achieves DSC values of 86.5% on the whole tumor, 86.4% on the core tumor, and 80.6% on the enhancing tumor. In specific, Our approach achieves the highest DSC values for (core and enhancing) tumor and core tumor sensitivity. Hu et al. [18] achieved slightly higher on whole tumor segmentation; their approach suggested multi cascaded convolutional neural networks. Still, their models cannot achieve good segmentation results for each view. Our IRDNU-Net achieves superior segmentation efficiency on DSC and Sensitivity metric than some recent approaches by Hu et al. [19] and Abouelenien et al. [1].

Figure 2 shows a sample results of the standard U-net model compared to the proposed one. The red regions in this figure refer to necrosis, edema is shown in green areas, and enhancing tumor is the yellow areas. In the meantime, Flair image, ground truth, U-Net, and

---

[3]https://www.cbica.upenn.edu/BraTS19/lboardValidation.html

Table 7 Ablation study on the Validation BraTS 2019 Dataset and comparison segmentation results with baselines

Methods	Whole	Core	Enhancing
$U - Net^*$	0.864	0.746	0.694
DNU-Net(our)	0.862	0.821	0.741
IRU-Net(our)	0.865	0.803	0.763
IRDNU-Net (our)	0.865	0.864	0.806

IRDNU-Net segmentation results are shown from left to right, respectively.It can be noticed from Fig. 2 that, IRDNU-Net evident produces the best performance in the segmentation of brain tumors. Figure 3 shows the results for Dice, Sensitivity, and Specificity for Validation data. The boxplots show the minimum, median, maximum, lower, and upper quartile. Points outside of the interquartile are referred to as outliers. From the boxplots, it was evident that our algorithm achieves considerably high segmentation accuracy in most cases.

4.3.3 Evaluation results on BraTS 2020 training dataset

We also execute an experiment on the BraTS 2020 training database to demonstrate our approach's effectiveness further. Here, to train our brain tumor segmentation models, we use 160 cases from the training dataset and 56 cases for the test as mention in Table 1. Table 9 shows the ablation study and comparison results with baselines. Besides, Fig. 4 Represents bar plots of the average DSC for the BraTS 2020 test dataset for the three tumor regions. In this experiment, IRDNU-Net achieves higher segmentation efficiency in the three tumor regions relative to its baseline U-Net. Meanwhile, IRDNU-Net exceeds the U-Net by 0.9% in the whole, 4.7% in the core tumor, and 5% in the enhancing tumor, respectively. After that, compared DNU-Net with IRDNU-Net, the DSC scores increased by 0.7%, 0.8%, and 3%, respectively, on three tumor segmentation regions. In specific, our models gain accuracy improvement over the baselines; this is also because of the efficacy of multi Inception residual with dense nested U-net in improving the segmentation of small brain tumors. Besides that, Fig. 5 shows a scatter diagram of the validation set. It reveals that for most brain images, our algorithm performs well. Because the BRATS 2020 data varies widely, and the class distribution is severely imbalanced, some outliers cause a reduction in the average score.

Table 8 Comparison of segmentation results on the BraTS 2019 Validation Dataset with typical methods

| Methods	DSC			Sensitivity			Specificity		
	Whole	Core	Enhancing	Whole	Core	Enhancing	Whole	Core	Enhancing
K. Hu et al. [18]	0.882	0.748	0.718	0.907	0.76	0.868	0.991	0.996	0.994
Zhang et al. [38]	0.865	0.80	0.745	–	–	–	–	–	–
Hu et al. [19]	0.850	0.70	0.65	0.83	0.79	0.65	–	–	–
Abouelenien et al. [1]	0.865	0.80	0.745	0.883	0.80	0.78	0.999	0.998	0.999
Chandra et al. [6]	0.872	0.795	0.741	0.829	0.788	0.795	0.994	0.997	0.998
Pereira et al. [32]	0.84	0.72	0.62	0.86	0.76	0.68	0.85	0.82	0.60
IRDNU-Net(our)	0.865	0.864	0.806	0.885	0.884	0.842	0.992	0.966	0.966
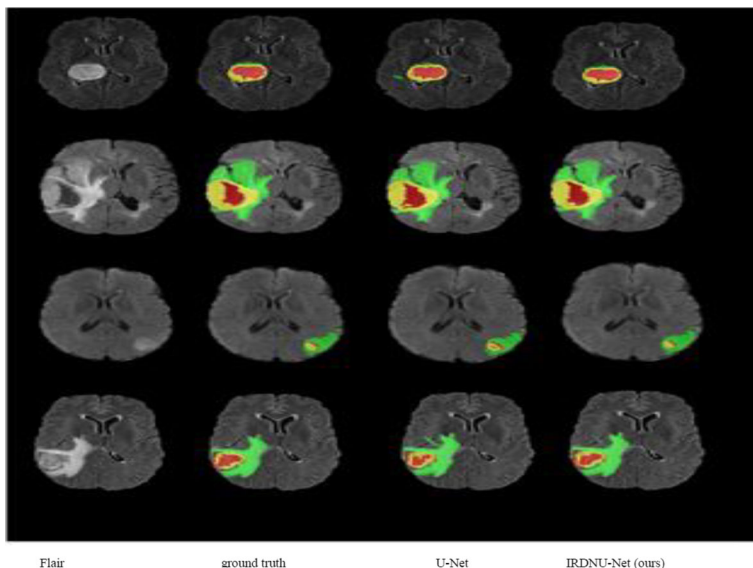
Fig. 2 Samples of segmentation results for the BraTS 2019 training dataset. Flair image, ground truth, U-Net, and IRDNU-Net, respectively, from left to right. Each color describes the class of tumor Red for necrosis, non-enhancing, green for edema, and yellow for enhancing tumor
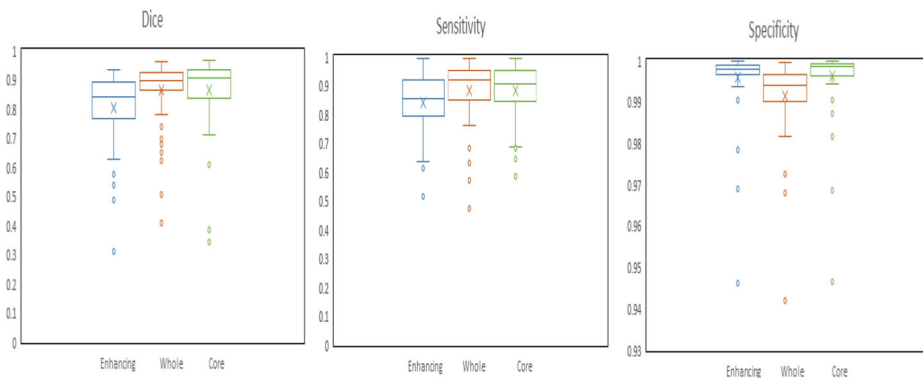

Fig. 3 Boxplots of DSC Sensitivity and Specificity for the BraTS'2019 validation dataset. The 'x' signifies the mean score, "○" shows outliers

Table 9 Ablation study on the BraTS 2020 test dataset and segmentation results with baselines

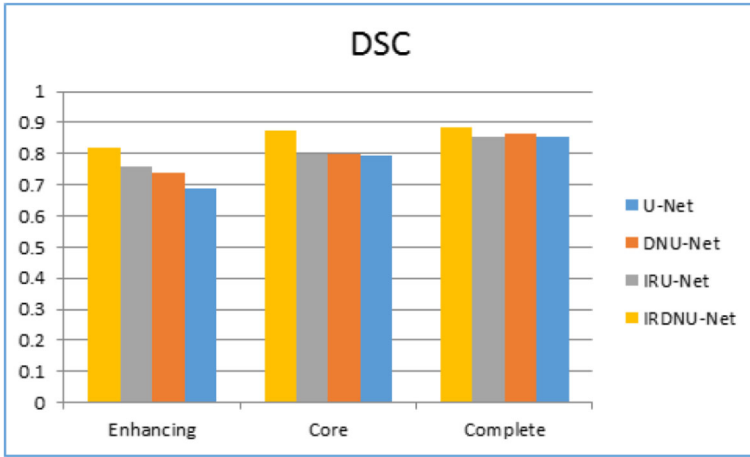| Methods	Whole	Core	Enhancing
$U - Net^*$	0.867	0.793	0.751
DNU-Net(our)	0.869	0.832	0.771
IRU-Net(our)	0.87	0.821	0.786
IRDNU-Net (our)	0.876	0.84	0.801

Fig. 4 Comparison of DSC score in the BraTS 2020 test dataset

Moreover, the proposed model decreases the number of trainable parameters, reducing the computational cost. Table 10 displays a comparison of the trainable parameters and times for one epoch for both the proposed model, and related models, which demonstrate that IRDNU-Net has 5.91M trainable parameters only, which is the least parameters among all competitive methods. Lin et al. [27] have 24.62M parameters, which are almost four times the proposed method. Zhou exceeds our proposed method by little margin in the training times for one epoch. These results showed that IRDNU-Net requires the lower computational resources with only 5.91M parameters. Generally, our suggested approach obtains a good balance between brain tumor segmentation accuracy and the number of training parameters of the BRATS2019 and BRATS2020 datasets.
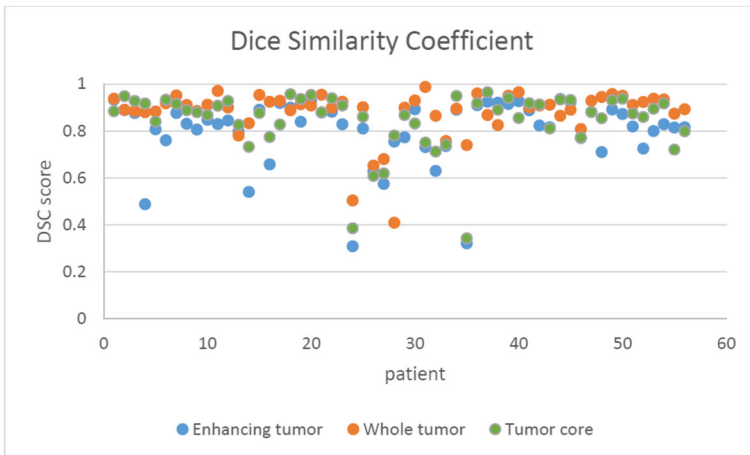

Fig. 5 Scatter plot on the BRATS 2020 dataset using the proposed approach

Table 10 The number of trainable parameters and the average time for one epoch

Methods	Parameters	Time/epoch(seconds)
U-Net(baseline)	7.76 M	381
Zhou et al. [41]	9.04 M	280
Ibtehaz et al. [20]	7.26 M	295
Kermi et al. [24]	10.15 M	446
Wang et al. [40]	13.81 M	–
Lin et al. [27]	24.62 M	–
Lou et al. [28]	10.07M	–
IRDNU-Net(our)	5.91M	300

5 Discussion

Accurate segmentation of gliomas has garnered considerable interest from medical doctors and researchers as a critical component of tumor diagnosis, treatment preparation, and subsequent assessment. Since manual tumor segmentation is laborious and time-consuming, developing an accurate automated segmentation is very important. Therefore, in this study, We develop an inception residual model with nested dense paths based on the U-Net to achieve high segmentation accuracy with fewer parameters. We evaluated the network on BRATS 2019 and BRATS 2020 datasets. These datasets are composed of MRI images taken from various institutions. It is gathered by regular clinical assessment of preoperative scans for glioblastoma patients. The proposed network achieved better results compared to the U-Net, and other techniques. The results in Fig. 2, indicates that the size, shape, location, and intensity of tumors in these samples are different, and also enhance the segmentation performance for small tumor regions. Generally, the results of the proposed architecture are comparable to the ground truth result. In Tables 5, and 8 it is also observed a small gaps for the evaluation metrics between training and validation dataset because we used only 66 cases in validation due to memory limitations of the current GPU. Figure 3 shows boxplots for validation of the BraTS'2019 dataset it is observed that the variance of the specificity tumor core (TC) is larger than that of specificity enhancing tumor(ET), the most likely reason for this is that the network sometimes incorrectly predicts the whole tumor as the tumor core due to the influence of the LGG tumor samples, resulting in increased variance. Due to the memory limitations of the current GPU, and the multi-modality nature of MRI, it is worth noting that in the proposed method, the time for one epoch is around 5 min.

6 Conclusion

This paper introduced an efficient IRDNU-Net model for automated brain tumor segmentation from Multi-modality MRI images. This approach is an efficient extension to the successful idea of encoder-decoder fully convolutional neural networks. First, we integrate the Inception module and residual units into U-Net in each block to enhance brain tumor segmentation performance. A series of dense nested pathways then connect the sub-networks encoder and decoder. The re-modeled skip connections aim to minimize the semantic gape between the feature maps of the encoder and decoder networks. We assessed our proposed approach using the BRATS 2019 and BRATS 2020 datasets. The experiment

results showed that IRDNU-Net surpassed the U-Net and other typical brain tumors segmentation approaches by a large margin. IRDNU-Net is capable of achieving comparable segmentation accuracy with fewer parameters. However, to build our segmentation model, we used 2D slices due to computational resources limitations. In the future, we expect to work on 3D networks while seeking a balance between high accuracy and computational resources. In addition, we will use a more powerful GPU. Also, for further evaluation, we will extend our model to other medical segmentation image tasks.

Declarations

Conflict of interest The authors declare that they have no conflict of interest

References

1. Aboelenein NM, Songhao P, Koubaa A, Noor A, Afifi A (2020) HTTU-Net: hybrid two track U-Net for automatic brain tumor segmentation. IEEE Access 8:101406–101415
2. Ahmad P, Qamar S, Hashemi SR, Shen L (2019) Hybrid labels for brain tumor segmentation. In: International MICCAI brainlesion workshop, pp 158–166
3. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Eaton-Rosen Z (2018) Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint arXiv:1811.02629
4. Bauer S, Wiest R, Nolte LP, Reyes M (2013) A survey of MRI-based medical image analysis for brain tumor studies. Physics in Medicine and Biology, 58
5. Cahall DE, Rasool G, Bouaynaya NC, Fathallah-Shaykh HM (2019) Inception modules enhance brain tumor segmentation. Front Comput Neurosci 13:44
6. Chandra S, Vakalopoulou M, Fidon L, Battistella E, Estienne T, Sun R, Paragios N (2018) Context aware 3D CNNs for brain tumor segmentation. In: International MICCAI brainlesion workshop. Springer, Berlin, pp 299–310
7. Cheng X, Jiang Z, Sun Q, Zhang J (2019) Memory-efficient cascade 3D U-Net for brain tumor segmentation. In: International MICCAI brainlesion workshop. Springer, Berlin, pp 242–253
8. Chen W, Liu B, Peng S, Sun J, Qiao X (2018) S3D-UNet: Separable 3D U-Net for brain tumor segmentation. In: International MICCAI brainlesion workshop. Springer, Berlin, pp 358–368
9. Da K (2014) A method for stochastic optimization. arXiv preprint arXiv:1412.6980
10. Dargan S, Kumar M, Ayyagari MR, Kumar G (2019) A survey of deep learning and its applications: A new paradigm to machine learning. Archives of Computational Methods in Engineering,, pp 1–22
11. Dataset:CBICA. https://www.med.upenn.edu/cbica/brats2019/data.html
12. Dolz J, Gopinath K, Yuan J, Lombaert H, Desrosiers C, Ayed IB (2018) HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation. IEEE Trans Med Imaging 38(5):1116–1126
13. Dong H, Yang G, Liu F, Mo Y, Guo Y (2017) Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In: Annual conference on medical image understanding and analysis. Springer, Berlin, pp 506–517
14. Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C (2016) The importance of skip connections in biomedical image segmentation. In: Deep learning and data labeling for medical applications. Springer, Berlin, pp 179–187
15. Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning 1-2. MIT press, Cambridge
16. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Larochelle H (2017) Brain tumor segmentation with deep neural networks. Med Image Anal 35:18–31
17. He H, Zhang C, Chen J, Geng R, Chen L, Liang Y, Xu Y (2021) A hybrid-attention nested UNet for Nuclear segmentation in histopathological images. Front Mol Biosci 8:6
18. Hu K, Gan Q, Zhang Y, Deng S, Xiao F, Huang W, Gao X (2019) Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field. IEEE Access 7:92615–92629
19. Hu Y, Xia Y (2017) 3D deep neural network-based brain tumor segmentation using multimodality magnetic resonance sequences. In: International MICCAI brainlesion workshop. Springer, Berlin, pp 423–434

20. Ibtehaz N, Rahman MS (2020) MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Netw 121:74–87
21. Işın A, Direkoğlu C, Şah M (2016) Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Comput Sci 102:317–324
22. Ioffe S, Szegedy C (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning. PMLR, pp 448–456
23. Ker J, Wang L, Rao J, Lim T (2017) Deep learning applications in medical image analysis. Ieee Access 6:9375–9389
24. Kermi A, Mahmoudi I, Khadir MT (2018) Deep convolutional neural networks using U-Net for automatic brain tumor segmentation in multimodal MRI volumes. In: International MICCAI brainlesion workshop. Springer, Berlin, pp 37–48
25. Kumar M, Gupta S, Kumar K, Sachdeva M (2020) Spreading of COVID-19 in India, Italy, Japan, Spain, UK, US: A prediction using ARIMA and LSTM model. Digit Gov: Res Prac 1(4):1–9
26. Li H, Li A, Wang M (2019) A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. Comput Biol Med 108:150–160
27. Lin F, Wu Q, Liu J, Wang D, Kong X (2020) Path aggregation U-Net model for brain tumor segmentation. Multimedia Tools and Applications, pp 1–14
28. Lou A, Guan S, Loew MH (2021) DC-UNet: Rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation. In: Medical imaging 2021: image processing 11596: 115962T, international society for optics and photonics
29. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Van Leemput K (2014) The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34(10):1993–2024
30. Miranda-Filho A, Piñeros M, Soerjomataram I, Deltour I, Bray F (2017) Cancers of the brain and CNS: global patterns and trends in incidence. Neuro-oncology 19(2):270–280
31. Olabarriaga SD, Smeulders AW (2001) Interaction in the segmentation of medical images: a survey. Med Image Anal 5(2):127–142
32. Pereira S, Pinto A, Alves V, Silva CA (2016) Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging 35(5):1240–1251
33. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer, Berlin, pp 234–241
34. Saouli R, Akil M, Kachouri R (2018) Fully automatic brain tumor segmentation using end-to-end incremental deep neural networks in MRI images. Comput Methods Prog Biomed 166:39–49
35. Sudre CH, Li W, Vercauteren T, Ourselin S, Cardoso MJ (2017) Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Berlin, pp 240–248
36. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence, vol 31, p 1
37. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
38. Zhang J, Jiang Z, Dong J, Hou Y, Liu B (2020) Attention Gate ResU-Net for automatic MRI brain tumor segmentation. IEEE Access 8:58533–58545
39. Zhao X, Wu Y, Song G, Li Z, Zhang Y, Fan Y (2018) A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med Image Anal 43:98–111
40. Zhou C, Ding C, Wang X, Lu Z, Tao D (2020) One-pass multi-task networks with cross-task guided attention for brain tumor segmentation. IEEE Trans Image Process 29:4516–4529
41. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2018) Unet++: A nested u-net architecture for medical image segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Berlin, pp 3–11