Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

A novel end-to-end brain tumor segmentation method using improved fully convolutional networks

Haichun Li, Ao Li*, Minghui Wang

School of Information Science and Technology, and Centers for Biomedical Engineering, University of Science and Technology of China, Hefei, AH 230027, China

ARTICLE INFO

Keywords:
Brain tumor segmentation
Fully convolutional networks
Deep learning
Glioma
Magnetic resonance imaging

ABSTRACT

Accurate brain magnetic resonance imaging (MRI) tumor segmentation continues to be an active research topic in medical image analysis since it provides doctors with meaningful and reliable quantitative information in diagnosing and monitoring neurological diseases. Successful deep learning-based proposals have been designed, and most of them are built upon image patches. In this paper, a novel end-to-end brain tumor segmentation method is developed using an improved fully convolutional network by modifying the U-Net architecture. In our network, an innovative structure referred to as an up skip connection is first proposed between the encoding path and decoding path to enhance information flow. Moreover, an inception module is adopted in each block to help our network learn richer representations, and an efficient cascade training strategy is introduced to segment brain tumor subregions sequentially. In contrast to those patchwise methods, our model can automatically generate segmentation maps slice by slice. We have validated our proposal by using imaging data from the Multimodal Brain Tumor Image Segmentation Challenge (BRATS) 2015 and BRATS 2016. Experimental results compared with U-Net suggest that our method is 2.6%, 3.9%, and 5.2% higher (by using the BRATS 2015 training dataset) as well as 2.8%, 3.7%, and 8.1% (by using the BRATS 2017 training dataset) higher in terms of complete, core and enhancing tumor regions, respectively. Quantitative and visual evaluation of our method has revealed the effectiveness of the proposed improvements and indicated that our end-to-end segmentation method can achieve a performance that can compete with state-of-the-art brain tumor segmentation approaches.

1. Introduction

As the most prevalent primary brain tumors among adults, gliomas emanate from glial cells of the brain or the spine and are pathologically categorized into high-grade gliomas (HGGs) and low-grade gliomas (LGGs) according to the well-known World Health Organization (WHO) grading system [1,2]. Grade III and IV tumors are classified as HGGs and are highly invasive and commonly lead to a worse prognosis, while LGGs (comprised of grade I and II) tend to exhibit benign tendencies with better prognoses [3]. In clinical practice, brain magnetic resonance imaging (MRI) has been commonly used to diagnose gliomas and for follow-up evaluations of the brain due to its sensitivity and significant image contrast in soft tissues [4]. However, the multiplicity and complexity of brain tumors under MRI often make tumor recognition and segmentation difficult for radiologists and other clinicians [5]. Hence, automated segmentation of heterogeneous tumors is of great importance in clinical medicine by freeing doctors from the time-consuming and laborious manual delineation of tumors [4]. Moreover, excellent brain tumor segmentation algorithms with

quantitative measurements of tumor delineation will significantly contribute to cancer diagnosis, treatment and prognosis.

In the past few decades, massive research efforts have been devoted to exploring accurate segmentation approaches in silico to generate reliable depictions of brain tumors, and many successful methods have been proposed in this field [6–15]. These methods generally fall into two categories, namely, generative models and discriminative models. Generative models usually demand prior information generated from probabilistic atlases for image registration [6–8], while image registration is not reliable for very large tumors. On the other hand, discriminative models exploit various image features to segment target tumors by categorizing each image voxel [9–15]. Discriminative models usually adopt discriminative learning approaches, including support vector machines (SVMs) and random forests. These discriminative methods are usually costly and time consuming because they rely heavily on various handcrafted image features generated by experts, such as texture features [16] and local histograms [9].

Recently, deep learning has become the favored approach for various complicated tasks due to the considerable capacity of the models to

---

automatically learning task-adaptive image features, which often outperform manually designed features [17]. Implementing deep learning techniques in brain tumor segmentation has attracted increasing attention. For instance, Dvorak and Menze [18] adopted three convolutional neural networks (CNNs) to solve binary segmentation subtasks in the multiclass brain tumor segmentation task. Pereira et al. [19] modeled deep CNNs using 3×3 filters to achieve automatic brain tumor segmentation. Havaei et al. [20] introduced a novel brain tumor segmentation method with a two-pathway structure to capture multiscale features and a cascaded architecture where the probability outputs of prior CNNs are used as extra inputs to subsequent CNNs. Although these deep learning-based methods efficiently adopt CNNs to attain promising performance, most of them are pointwise convolutions, which model the segmentation task as numerous classification tasks for each voxel by using local regions centered on the target voxel, namely, image patches. The limitation of pointwise segmentation methods is that they normally suffer from high computational cost and the segmentation performance is sensitive to the size of image patches [21,22].

Instead of modeling a powerful classification network to categorize numerous image patches, a more elegant architecture called fully convolutional networks (FCNs) [23] was proposed to automatically generate the label map for whole input images, which solved the limitation of high computational cost and difficulty in determining the patch size for pointwise methods. Based on this architecture, a symmetric fully convolutional network named U-Net was proposed by Ronneberger et al. [21], which achieved convincing performance on medical images and has been extensively adopted in various tasks. Some researchers have employed U-Net in automated brain tumor segmentation [24–26]. For example, Beers et al. [24] proposed a 3D glioma segmentation method by adopting sequential U-Nets, and Dong et al. [25] proposed a 2D fully automatic brain tumor segmentation network built upon the U-Net architecture. However, on account of the low memory capacity, U-Net presents limited capacity in effectively learning the feature information of the images in complicated tasks [27]. This limitation leads to a need for optimizing the network structure to enlarge the parameter space, which allows the network to learn more representative features [28].

In this paper, a novel end-to-end brain tumor segmentation method using an improved U-Net is developed for 2D brain MRI slices. We modify and improve the U-Net architecture to strengthen model capacity for yielding more accurate segmentations; see Fig. 1. One important extension in our architecture is that we improve the network connectivity using an innovative structure referred to as an up skip connection. Different from previous fully convolutional architectures, the up skip connection further enhances information flow between the encoding portion and decoding portion to ensure more low-level features are used for optimizing the segmentation results. Moreover, the inception module [29] is adopted in each block to increase the network's capacity for learning richer representations. Based on our improved fully convolutional network, a cascade of models is proposed to sequentially segment brain tumor subregions, and an effective cascaded training strategy is introduced to generate more precise segmentations for small tumor regions. Above all, in both the training and testing phases, we performed a simplified end-to-end brain tumor segmentation by using image slices. We evaluated the proposed method based on publicly available MR images obtained from the Multimodal Brain Tumor Image Segmentation Challenge (BRATS) datasets from 2015 to 2017. The experimental results have indicated the efficiency of the proposed improvements and suggest that our approach could acquire competitive performance as state-of-the-art brain tumor segmentation methods.

2. Material and Methods

2.1. Datasets

The proposed method is primarily validated on the image database from the Multimodal Brain Tumor Image Segmentation Challenge (BRATS) of 2015 [30]. The BRATS 2015 dataset comprises data from (1) BRATS 2012, (2) BRATS 2013, and (3) The Cancer Imaging Archive (TCIA) of the NIH. There are 220 high-grade glioma (HGG) patients and 54 low-grade glioma (LGG) patients in the training dataset with professional segmentations provided as ground truth. The testing dataset consists of 110 unlabeled patients, which can be evaluated only via the challenge website. Each patient had different MRI pulses (i.e., T1-weighted (T1), T1 contrast-enhanced (T1c), T2-weighted (T2) and FLAIR), each of which is composed of 155 brain slices, for a total of 620 images per patient. All MRI images were skull stripped and coregistered to the identical anatomical template at 1 mm isotropic resolution [30]. There are five labels for the imaging data in the BRATS 2015 dataset, including necrosis (labeled as 1), edema (labeled as 2), nonenhancing tumor (labeled as 3), enhancing tumor (labeled as 4) and everything else (labeled as 0). According to practical clinical applications, the evaluation of the multiclass brain tumor segmentation task is performed for different tumor subregions:

1) **Complete tumor.** Contains all four intratumor regions (i.e., label 1, label 2, label 3 and label 4).
2) **Core tumor.** Consists of three intratumor regions (i.e., label 1, label 3 and label 4).
3) **Enhancing tumor.** Includes an enhancing tumor (i.e., label 4)

The BRATS 2017 [30,31] training dataset, which contains 210 HGGs and 75 LGGs with manual segmentations, is also included to examine the performance of our proposed method. In contrast to the BRATS 2015 dataset, annotations of imaging data in BRATS 2017 comprise the GD-enhancing tumor (labeled as 4), the peritumoral edema (labeled as 2), the necrotic and nonenhancing tumor (labeled 1) and everything else (labeled as 0). Accordingly, the tumor subregions are defined as (1) complete tumor (i.e., labels 1, 2, and 4) (2) core tumor (i.e., 1, 4), and (3) enhancing tumor (i.e., 4).

In this study, each 2D brain slice is normalized to mitigate the large intensity variance among MR images by using the z-score defined as follows:

$$I_{norm} = \frac{I - \mu(I)}{\sigma(I)} \tag{1}$$

where μ and σ represent the average value and the standard deviation of image I, respectively.

2.2. Methods

In this study, the proposed end-to-end brain tumor segmentation model is built upon the U-Net architecture, which represents one of the most well-known fully convolutional network architectures for medical image segmentation [28]. We extend and modify the U-Net structure by introducing a new cross-layer architecture, namely, an up skip connection, and incorporating inception modules (see Fig. 1). In the training phase, we further propose a cascaded training strategy for sequentially segmenting complete, core and enhancing tumors (see Fig. 2). The proposed method segments brain MR images slice by slice using all four sequences (i.e., T1, T2, T1c and FLAIR).

2.2.1. Up skip connection

As depicted in Fig. 1(a), we enhance network connectivity through an innovative up skip connection between the downsampling path and the upsampling path to facilitate the model's capacity to learn multi-level features for accurate brain tumor segmentation. On the one hand,
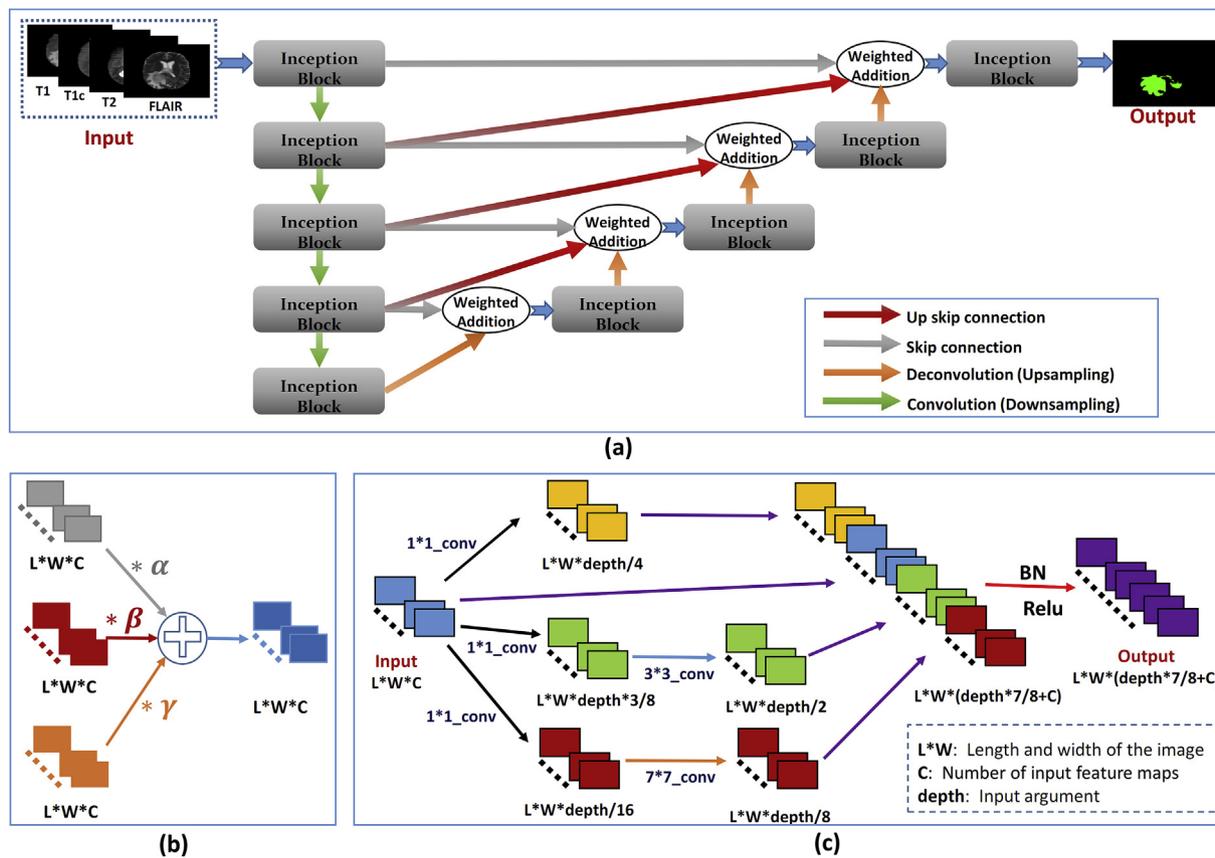
Fig. 1. (a) The network structure of our proposed network. This figure shows an illustration of the encoding path (left: top to bottom) and the decoding path (right: bottom to top). (b) Details of the weighted addition. (c) Details of the inception block. L and W represent the length and width of input images, respectively, where C is the number of input feature maps and depth denotes the manually defined parameter.
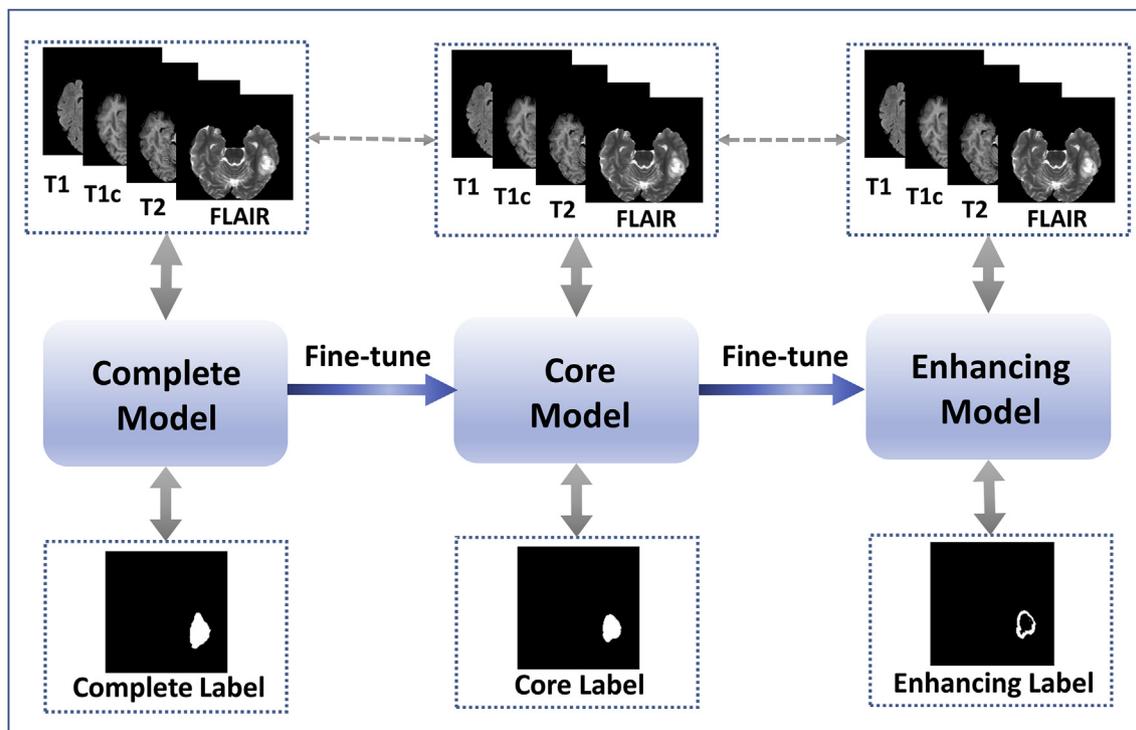

Fig. 2. The proposed cascaded framework for brain tumor segmentation. The three models, from left to right, are proposed to hierarchically segment the complete tumor, the core tumor and the enhancing tumor.

the up skip connection provides a new pathway between the encoding portion and the decoding portion during the forward propagation process, which allows the upsampling layers to extract more low-level features and thus helps recover spatial information lost during downsampling. On the other hand, the up skip connection improves the gradient flow in the backward propagation process, offering additional guidance for the learning of earlier layers and thus alleviating the difficulty of training the deep network. As shown in the figure, given the total number of layers N, a skip connection is typically employed to combine layer i and layer $N-i+1$ following conventional U-Net architectures, and our up skip connection further shares information between layer i and layer $N-i+2$ accordingly. A weighted addition structure is applied to automatically incorporate low-level feature maps in upsampling layers (see Fig. 1(a)). To ensure that different feature maps input into the weighted addition structure possess identical size, the proposed up skip connection and the upsampling operation in the decoding path both adopt 3×3 deconvolution layers with 2×2 strides to acquire uniform-sized outputs, while the skip connection utilizes a 1×1 convolution layer with a 1×1 stride to balance feature size. The weighted addition structure is graphically presented in Fig. 1(b) and can be formally expressed as:

$$x_o = \alpha \cdot x_{skip} + \beta \cdot x_{up-skip} + \gamma \cdot x_{up-sampling} \qquad (2)$$

where x_o is the output filter of the weighted addition structure and x_{skip}, $x_{up-skip}$ and $x_{up-sampling}$ represents the output feature maps of the skip connection, the up skip connection and the corresponding upsampling layer, respectively. Parameters α, β and γ are learned parameters, which denote the corresponding constraint coefficients of x_{skip}, $x_{up-skip}$ and $x_{up-sampling}$. The detailed implementation is given in Fig. s1.

2.2.2. Inception module

To improve the representation capacity of the segmentation network and to optimize the segmentation performance, we modify the U-Net architecture with an inception module, which has been experimentally proven to enhance the capturing of more visual information under constrained computational complexity [29]. As illustrated in Fig. 1(c), the inception module implemented in our network is redesigned where a 7×7 convolution layer to enlarge the receptive field and the max pooling layer is replaced by a short path to directly incorporate the input filters. The output filters generated from the 1×1, 3×3 and 7×7 convolutional layers are concatenated with the input feature map to achieve feature fusion. At the same time, to speed up the convergence and further overcome the disadvantages of deep neural networks that are difficult to train, we adopt BN (batch normalization) operations to normalize the feature maps [32].

2.2.3. Cascaded training strategy

In this study, a cascade of binary segmentation networks was trained for three subtasks in segmenting brain tumor subregions (complete, core and enhancing tumor) sequentially. However, the separated process of the three tumor regions often leads to less adequate performance for core and enhancing tumor segmentation as the number of pixels for core and enhancing tumor is fairly small, which may impede the network to learn abundant feature representations. To partially solve this problem and to further improve the segmentation performance for the core and the enhancing tumor, we proposed an efficient cascaded training strategy inspired by the success of knowledge transfer in boosting performance [33]. By the use of the proposed cascaded training strategy, the knowledge learned from the complete segmentation task is transferred to the core segmentation task, and the information learned from the core segmentation task is shared with the enhancing segmentation task. As depicted in Fig. 2, the cascaded training strategy progressively optimizes the segmentation results for small tumor regions considering the inclusion relations of topologies between complete, core and enhancing tumors. In detail, the well-trained complete model is utilized as the pretrained model to initialize

the core model, and the enhancing model is fine-tuned based on the well-trained core model.

2.3. Evaluation metrics

In this work, we evaluate the performance of the proposed method for complete, core and enhancing tumor regions using three publicly available metrics: the Dice score, the positive predictive value (PPV) and the Sensitivity. Another widely adopted measure, the so-called Jaccard score [34], is also calculated to assess the results of our proposed method.

The Dice score measures the similarity of two segmented maps. It ranges from 0 to 1, and a higher Dice value indicates a better match. With regard to each tumor subregion, given the binary segmentation map $P \, \varepsilon \, (0,1)$ obtained from predictions of segmentation models and the corresponding ground truth $T \, \varepsilon \, (0,1)$, the Dice score is defined as follows:

$$Dice(P, T) = \frac{2 * |P_1 \cap T_1|}{(|P_1| + |T_1|)} \qquad (3)$$

Where $|P_1 \cap T_1|$ calculates the amount of elements found in both sets and $|P_1|$ and $|T_1|$ are the cardinalities of the two sets (i.e., the number of pixels where $P = 1$ and $T = 1$, respectively).

The Sensitivity and the PPV are employed to evaluate the number of positive segmentations. The PPV measures the number of false positives and true positives, while the Sensitivity represents a useful measure to evaluate the amount of true positives and false negatives. Accordingly, they are defined as Eq. (4) and Eq. (5):

$$Sensitivity(P, T) = \frac{|P_1 \cap T_1|}{|T_1|} \qquad (4)$$

$$PPV(P, T) = \frac{|P_1 \cap T_1|}{|P_1|} \qquad (5)$$

The Jaccard score is one of the most popular evaluation measures in medical image segmentation [35], which can be calculated as follows:

$$Jaccard(P, T) = \frac{|P_1 \cap T_1|}{|P_1 \cup T_1|} \qquad (6)$$

3. Experiments and results

We validated our method based on the BRATS 2015 dataset by simultaneously using two GPUs and one CPU on a computing workstation equipped with 4 NVIDIA 1080Ti GPUs and 2 Intel Xeon E5-2630 CPUs. Our proposed network was implemented using the Keras framework with a TensorFlow backend, and the detailed parameters are exhibited in Table 1.

Table 1
Hyperparameters in our proposed network.

Stage	Hyperparameter	Value
Initializer of convolution layers	Bias	0
	Weight	He_normal
Adam optimizer	β1	0.9
	β2	0.99
	epsilon	10–8
Training	learning rate	0.01
	batch size	64
	epoch	100
	early stopping	20
	loss function	$-\frac{2*|P_1 \cap T_1|}{(|P_1|+|T_1|)}$

Table 2
Experiment data partitioning.

	Training set	Validation set	Testing set	Total
HGG	132	44	44	220
LGG	33	11	10	54
Total	**165**	**55**	**54**	**274**

3.1. Experiments on training Data

3.1.1. Evaluating the effectiveness of the proposed improvements

Different experiments based on the BRATS 2015 training dataset were carried out to investigate how our proposed improvements influence tumor segmentation results with respect to the inception module, the up skip connection and the cascaded training strategy. In this study, we randomly split 274 labeled cases into a training set, validation set and testing set at a ratio of 6:2:2 [36]; that is, 165 samples were chosen as the training set, and 55 cases and 54 cases were chosen for the validation set and the testing set, respectively (see Table 2). We also compare different implementations of our proposed network (the inception-based U-Net, inception-based U-Net + up skip connection and inception-based U-Net + up skip connection + cascaded training strategy) with the U-Net, which is the most relevant work in our approach and achieves state-of-the-art results for brain tumor segmentation.

In this section, all models were evaluated on the testing set with 54 patients, and the quantitative results of our models along with the baseline U-Net are illustrated in Table 3, in which the Dice, PPV and Sensitivity metrics are presented as average values with standard deviations. These results are analyzed from different aspects:

(1) The baseline U-Net and ourinception-based U-Net. With respect to most measurements, ourinception-based U-Net achieves superior segmentation performance compared to the baseline U-Net for the three tumor regions. In particular, the performance of our model in terms of the Dice score is 1.3% (complete tumor), 0.8% (core tumor) and 3.6% (enhancing tumor) higher than that of U-Net, which indicates that the proposedinception module can help improve segmentation performance.
(2) Theinception-based U-Net and theinception-based U-Net + up skip connection. As we improve theinception-based U-Net with innovative up skip connection, the Dice value for the three tumors increased by 0.3%, 1.6% and 1.5%, respectively. The PPV of theinception-based U-Net with up skip connections also outperforms theinception-based U-Net. The comparison results suggest that enhancing information flow through up skip connections is efficient in generating more accurate segmentations.
(3) Theinception-based U-Net + up skip connection and theinception-based U-Net + up skip connection + cascaded training strategy. The cascaded training strategy is proposed to further optimize segmentation performance for small tumor regions. The results showing that both the Dice score and theSensitivity increased for core and enhancing tumors as we implemented a cascade training strategy demonstrate the efficiency of such a modification,

suggesting the potential of a cascaded training strategy in brain tumor segmentation.

As an overall analysis, Fig. 3 displays the boxplot of the calculated Dice score for all testing cases. The model tends to achieve better overall segmentation performance for the three tumor regions as we gradually implement the inception module, the up skip connection and the cascaded training strategy in U-Net. In summary, weighed against all Dice, PPV and Sensitivity metrics, the comparison results demonstrate that the proposed improvements contribute to accurate brain tumor segmentation.

The widely used Jaccard score is also used to assess and compare the segmentation results for the three tumor regions of our model and U-Net. Fig. 4 depicts bar plots of the average Jaccard score for the three tumor regions based on 54 cases in the testing set. These segmentation results further indicate that (1) both the up skip connection and the inception module implemented to build an improved U-Net architecture contribute to generating accurate brain tumor segmentations and that (2) the cascaded training strategy that learns from the transferred knowledge from the large tumor regions can help improve the performance for core and enhancing tumor.

To exemplify the effect of the proposed up skip connection, segmentation results of two representative cases are plotted in Fig. 5. Segmentations generated from the inception-based U-Net, the inception-based U-Net with up skip connections and the ground truth are illustrated in the FLAIR images. Both our models display satisfactory conformity to the manual delineation performed by experts. However, some confusing areas are erroneously identified as target tumors by the inception-based U-Net. Implementing up skip connections mitigates these problems and improves segmentation performance, which corroborates the efficiency of the proposed up skip connection in our network.

3.1.2. Evaluation of the proposed method

To investigate the impact of different network parameters, we trained the proposed method (the inception-based U-Net + up skip connection + cascaded training strategy) by using the BRATS 2015 training data with various batch sizes and learning rates. The segmentation results for different batch sizes (i.e., 8, 16, 32, and 64; see Table 4) suggest that the batch size of 64 helps our method achieve the best segmentation performance with less time consumption. Analogously, the proposed model with a *learning rate = 0.01* presents superior performance to that with a *learning rate = 0.1* and a *learning rate = 0.001* according to the comparison results illustrated in Table 5.

We present the results of six representative testing slices in Fig. 6. Segmentations generated by the proposed method and the ground truth segmentations for complete, core and enhancing tumors are depicted in FLAIR, T2 and T1c, respectively. When the segmentation results of our model (green) are compared against the physician's manual segmentation results (red), it is evident that our improved U-Net with cascaded training strategy has the capacity to obtain accurate tumor segmentations from multimodal brain MR slices. The multiplicity and complexity of brain tumors make segmentation unusually challenging, yet our method still provides satisfactory segmentations, which demonstrates
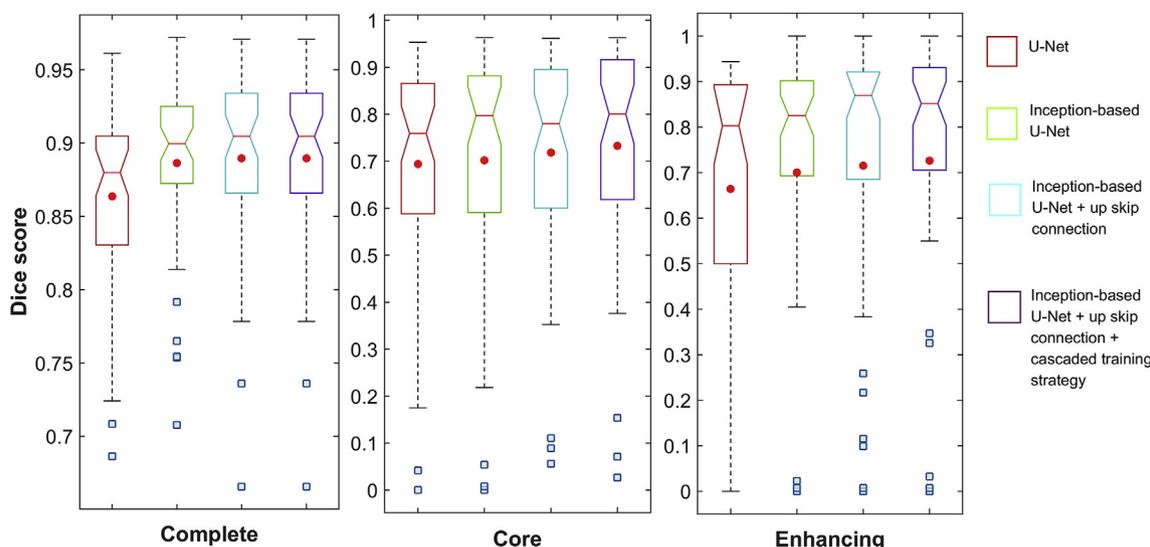
Table 3
The segmentation performance based on the Dice score, PPV and Sensitivity between our proposed methods and the baseline U-Net.

Method	Dice score			PPV			Sensitivity		
	Complete	Core	Enhancing	Complete	Core	Enhancing	Complete	Core	Enhancing
U-Net (Baseline)	0.864(0.07)	0.694(0.24)	0.664(0.32)	0.828(0.08)	0.764(0.22)	0.711(0.31)	0.911(0.08)	0.688(0.27)	0.702(0.32)
Inception-based U-Net	0.887(0.06)	0.702(0.24)	0.700(0.32)	0.868(0.07)	0.756(0.20)	0.765(0.23)	**0.912**(0.08)	0.713(0.29)	**0.754**(0.33)
Inception-based U-Net + up skip connection	0.890(0.06)	0.718(0.23)	0.715(0.31)	0.892(0.07)	**0.791**(0.20)	**0.806**(0.23)	0.895(0.09)	0.714(0.27)	0.733(0.33)
Inception-based U-Net + up skip connection + cascaded training strategy	**0.890**(0.06)	**0.733**(0.22)	**0.726**(0.31)	**0.892**(0.07)	0.785(0.19)	0.797(0.26)	**0.895**(0.09)	**0.750**(0.25)	**0.743**(0.32)

Fig. 3. The boxplot of Dice scores based on the 54 patients in the testing set. For each tumor region, from left to right: U-Net, the inception-based U-Net, the inception-based U-Net + up skip connection and inception-based U-Net + up skip connection + cascaded training strategy. The circle represents the mean values.
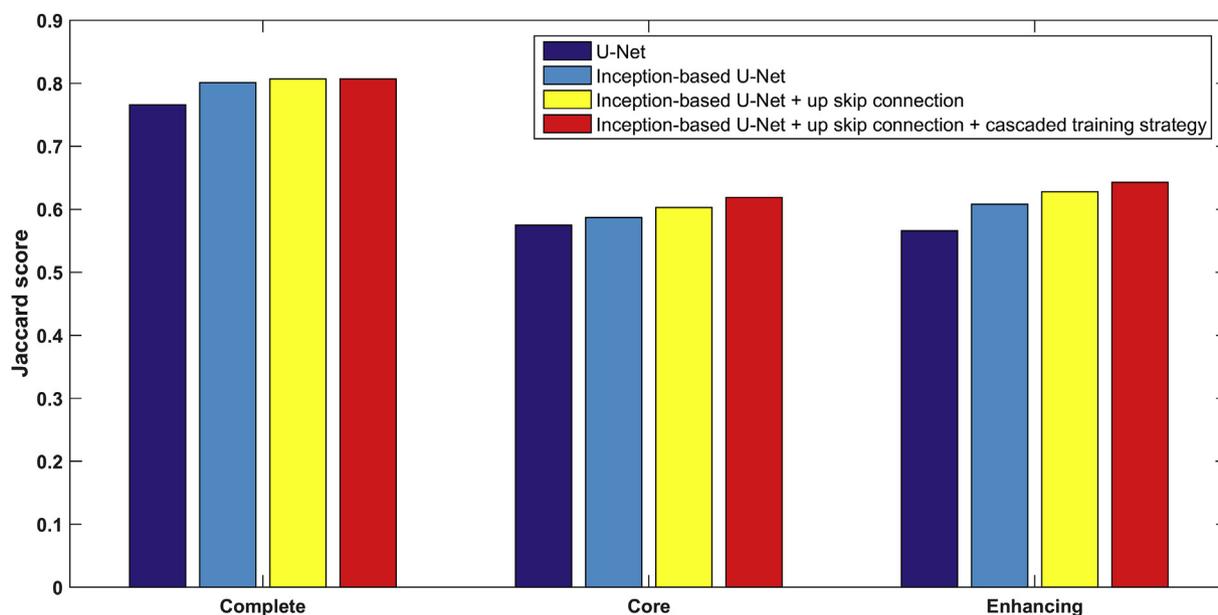

Fig. 4. Comparison of Jaccard scores based on 54 patients in the testing set. The mean value is displayed as an unfilled circle. For each tumor region, from left to right: U-Net, inception-based U-Net, the inception-based U-Net + up skip connection and the inception-based U-Net + up skip connection + cascaded training strategy.

the efficiency of the proposed end-to-end method.

The imaging data provided by the BRATS 2017 training dataset were further implemented to evaluate the performance of the proposed method. We randomly split the 285 labeled cases in the BRATS 2017 training dataset into a training set (171 cases), a validation set (57 cases) and a testing set (57 cases) with a ratio of 3:1:1 [36]. The learning curves of the proposed method based on the BRATS 2015 and 2017 imaging data are provided in Fig. 7. The model that achieved the best performance on the validation set was used to evaluate segmentation performance on the testing dataset. Table s1 presents the learned parameters in the weighted addition structure associated with the up skip connection of well-trained models.

To further validate the efficiency of our method, we compared the proposed method with different deep learning-based segmentation networks, including SegNet [37], U-Net [21], PSPNet [38] and 3D U-Net [39]. Fig. 8 shows the boxplots for the different methods with

various benchmarks based on the BRATS 2015 and 2017 datasets. It is evident from the results that our method achieves the best segmentation performance with regard to various tumor subregions and measurements. Table 6 gives the mean Dice scores with standard deviations for the experiments in Fig. 8, in which the statistical significance of differences was analyzed by using the Wilcoxon signed-rank test [40,41]. In summary, the proposed method presents significant improvements compared with existing deep learning-based methods, giving p-values smaller than 0.01. The detailed segmentation performance for each case and their corresponding tumor size is provided in Table s2 and Table s3.

3.2. Experiments on testing Data

In this section, we compare the proposed models (the inception-based U-Net, inception-based U-Net + up skip connection and
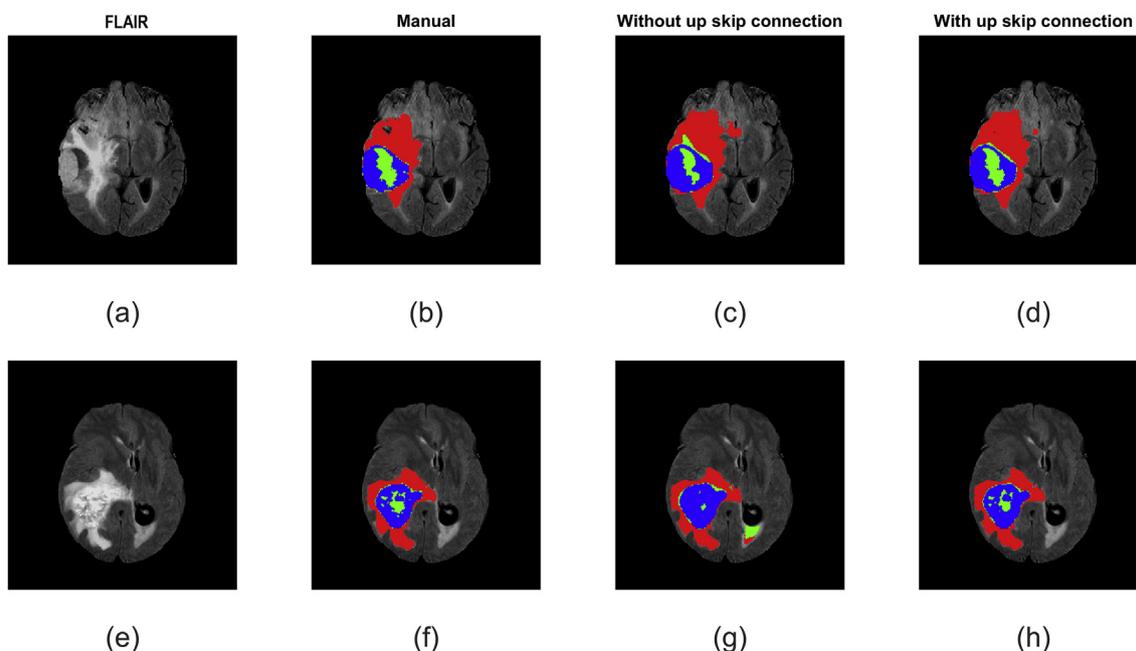
Fig. 5. Segmentation comparison between the inception-based U-Net (without up skip connection) and inception-based U-Net with up skip connection for two different cases: a) FLAIR slice, b) manual segmentation, where the blue region indicates an enhancing tumor, the red region indicates edema and the green region indicates necrosis and a nonenhancing tumor, c) segmentation generated through the inception-based U-Net, d) segmentation using the inception-based U-Net + up skip connection.

Table 4
Experimental results for different batch sizes.

Batch size		8	16	32	64
Dice score	Complete	0.869(0.08)	0.876(0.06)	0.878(0.06)	**0.890**(0.06)
	Core	0.721(0.23)	0.717(0.21)	0.725(0.22)	**0.733**(0.22)
	Enhancing	0.698(0.31)	0.709(0.31)	0.720(0.30)	**0.726**(0.31)
Time (s/epoch)		970	850	770	**740**

Table 5
Experimental results for different learning rates.

Learning rate		0.1	0.01	0.001
Dice score	Complete	0.875(0.07)	**0.890**(0.06)	0.890(0.06)
	Core	0.677(0.27)	**0.733**(0.22)	0.730(0.22)
	Enhancing	0.682(0.34)	**0.726**(0.31)	0.707(0.30)

inception-based U-Net + up skip connection + cascaded training strategy) with existing brain tumor segmentation methods presented by Refs. [17,42] on the BRATS 2015 testing data. To make full use of all 274 training cases, we implement a 5-fold cross validation procedure followed the method of Kamnitses et al. and calculate the segmentation results through the BRATS online system. To better evaluate the model efficiency, all the comparison results presented in Table 7 are generated without postprocessing steps. As shown in Table 7, implementing the up skip connections in the inception-based U-Net increases the Dice scores by 0.6% (complete tumor), 1.7% (core tumor) and 0.9% (enhancing tumor), and further application of the cascaded training strategy helps our model achieve the best Dice score performances of 84.5%, 69.8% and 60.0% for the three tumor regions. In addition to the Dice score, the comparison results based on the PPV and Sensitivity also show that the segmentation performance increases as we progressively implement the up skip connection and the cascaded training strategy. All these results are consistent with those in section 3.1 and further suggest that our proposed modifications and improvements are effective in generating accurate brain tumor segmentations.

Quantitative results of state-of-the-art brain tumor segmentation

approaches are also exhibited in this table. It is evident that the proposed method (the inception-based U-Net + up skip connection + cascaded training strategy) demonstrates superior performance on most benchmarks compared with Zhao et al.'s method. Kamnitses et al.'s method achieves remarkable segmentation performance with Dice scores of 0.836 (complete tumor), 0.674 (core tumor) and 0.629 (enhancing tumor), which shows advantages for enhancing tumors, while our method performs better in complete and core tumors. Weighted against all evaluation metrics, the proposed method outperforms Kamnitses et al.'s method. In particular, Kamnitses et al.'s method is pointwise, while our method can automatically yield segmentation maps slice by slice. In summary, the proposed end-to-end brain tumor segmentation method using improved fully convolutional networks can achieve a performance that can compete with existing brain tumor segmentation methods.

4. Discussion

As an important component of diagnosing tumors, treatment planning and subsequent evaluations, accurate segmentation of gliomas have attracted enormous attention from medical doctors and researchers. Since manual depiction of tumor regions is laborious and time consuming, it is important to develop efficient computational methods for automatic brain tumor segmentation. Nevertheless, most existing deep learning methods proposed for brain tumor segmentation are voxelwise, which demands a high computational cost, and the segmentation performance is normally sensitive to the patch size [18–20]. Since fully convolutional networks have been proposed to overcome this problem by automatically generating segmentations for whole input images, incorporating this framework in brain tumor segmentation may solve the aforementioned limitations efficiently. Although efforts have been made to implement such a network architecture, especially U-Net, the segmentation performance remains to be improved due to low memory capacity [27]. Therefore, in this study, we propose a novel end-to-end brain tumor segmentation method built upon an improved U-Net for 2D brain MRI slices. Experiments were performed on the BRATS 2015 and 2017 datasets. The proposed
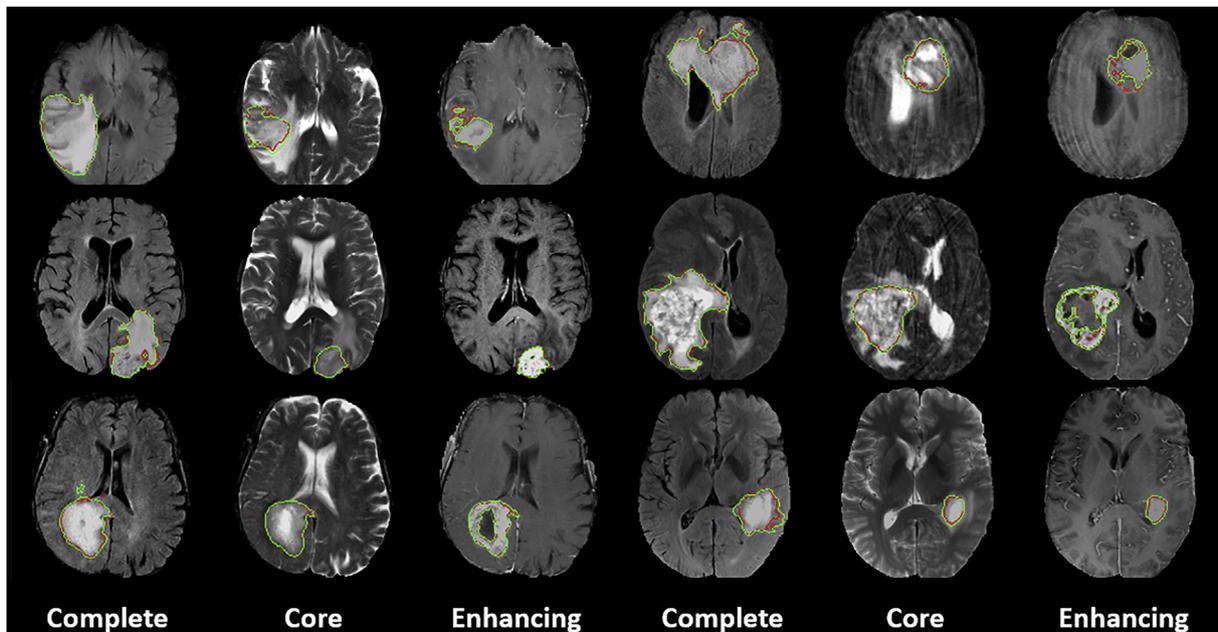
Fig. 6. Example of consensus expert annotation (red) and automatic segmentation of our method (green) applied to the testing set. Two cases are shown in each row. From left to right: segmentation of complete tumor (shown in FLAIR), core tumor (shown in T2) and enhancing tumor (shown in T1c).
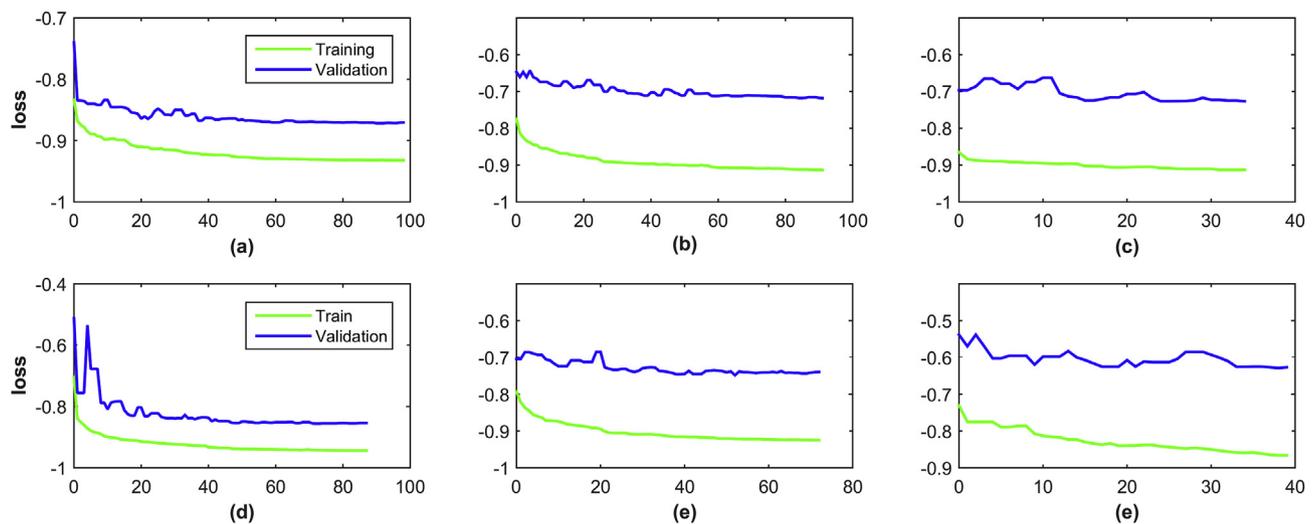

Fig. 7. The learning curves for our method based on the BRATS 2015 and the BRATS 2017 training datasets. (a) Complete tumor.

method presents more accurate segmentation performance than most previous methods.

One important contribution of our work is that we improve the network connectivity using an innovative structure referred to as an up skip connection, which enhances the information flow between the encoding portion and the decoding portion with an associated weighted addition structure. We evaluated the segmentation performance of such architecture based on the BRATS 2015 dataset (see Table 3, Table 7, Figs. 3, Fig. 4, Fig. 5). Quantitative and visual results have demonstrated that the up skip connection could help the network extract more low-level features, optimizing brain tumor segmentation performance. Moreover, the inception module [29] is adopted in each block to increase the network's capacity for learning richer representations. Experimental results (see Table 3, Figs. 3, Fig. 4) also indicate that our proposed inception-based U-Net could improve the segmentation performance in contrast to the baseline U-Net. The last contribution of this work is the cascaded training strategy, which effectively makes use of transferred knowledge among the three tumor regions to optimize the

segmentation performance of core and enhancing tumors. Accordingly, the comparison results (see Table 3, Table 7, Figs. 3, Fig. 4) indicate the potential of utilizing the cascaded training strategy to enhance the segmentation performance for small tumor regions.

The proposed method, which implements all of these effective modifications and improvements, is further evaluated based on both the BRATS 2015 and 2017 datasets. First, the proposed method was validated with various batch sizes and learning rates to investigate the impact of different parameters (see Table 4, Table 5). The comparison results indicate that (1) a larger batch size leads to better segmentation performance and (2) the model with a learning rate = 0.01 achieves the best performance for our model. Second, the learning curves of the proposed method based on the BRATS 2015 and 2017 imaging data are provided in Fig. 7 to show the convergence progress. The learned parameters of the weighted addition structure in well-trained models are also depicted in Table s1, reflecting the self-learning parameter optimization results for different blocks in the decoding portion. Third, we compared the proposed method with different deep learning-based
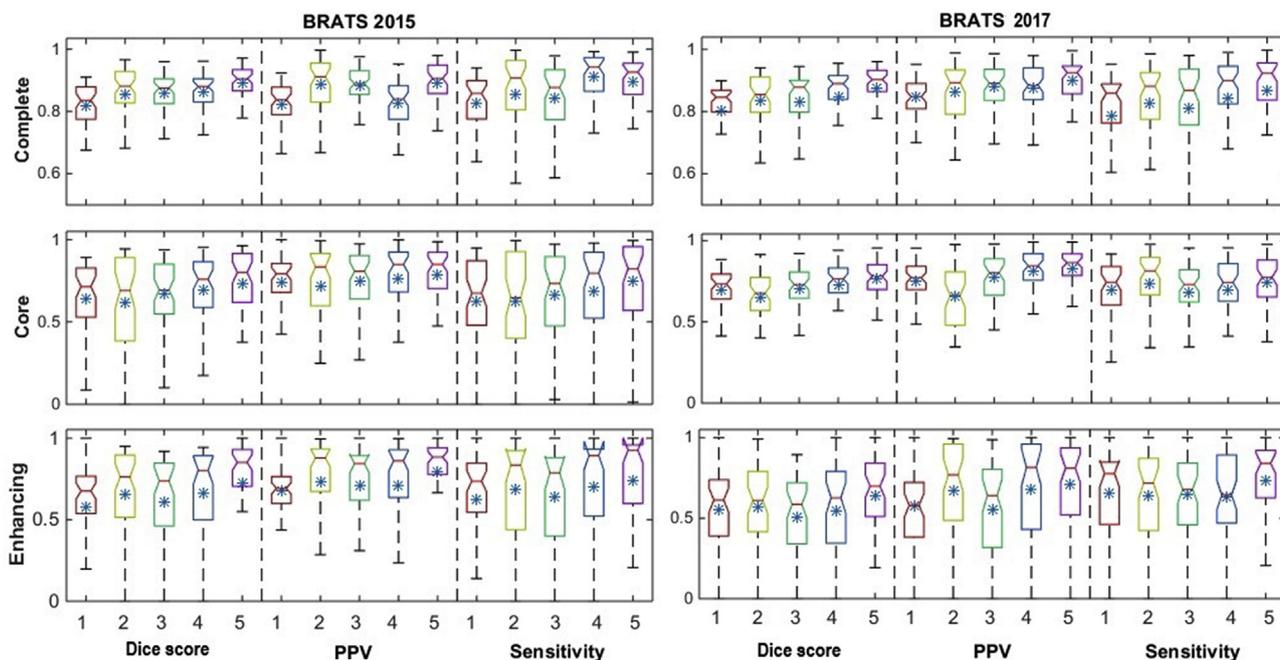
Fig. 8. The boxplot of PSPNet (shown as 1), 3D U-Net (shown as 2), SegNet (shown as 3), U-Net (shown as 4) and the proposed method (shown as 5) based on the BRATS 2015 and the BRATS 2017 training datasets with Dice score, PPV and Sensitivity benchmarks. From top to bottom: segmentation performance for complete tumors, core tumors and enhancing tumors. The asterisk represents the mean value.

Table 6

The segmentation performance based on the Dice score between our method and previous methods using the BRATS 2015 and BRATS 2017 training datasets.

Dataset	Method	Dice score		
		Complete	Core	Enhancing
BRATS 2015	SegNet	0.856(0.08)[a]	0.670(0.22)[a]	0.612(0.31)[a]
	U-Net	0.864(0.07)[a]	0.694(0.24)[a]	0.664(0.32)[a]
	3D U-Net	0.856(0.11)[a]	0.618(0.29)[a]	0.655(0.31)[a]
	PSPNet	0.820(0.07)[a]	0.639(0.24)[a]	0.575(0.27)[a]
	Our method	**0.890**(0.06)	**0.733**(0.22)	**0.726**(0.31)
BRATS 2017	SegNet	0.833(0.12)[a]	0.703(0.15)[a]	0.496(0.28)[a]
	U-Net	0.848(0.11)[a]	0.726(0.18)[a]	0.549(0.30)[a]
	3D U-Net	0.835(0.10)[a]	0.655(0.18)[a]	0.551(0.31)[a]
	PSPNet	0.809(0.11)[a]	0.701(0.15)[a]	0.554(0.26)[a]
	Our method	**0.876**(0.09)	**0.763**(0.13)	**0.642**(0.28)

[a] Dice scores that are significantly different from our method (from the paired Wilcoxon signed-rank test with p < 0.01 significance level).

segmentation networks, including SegNet [37], U-Net [21], PSPNet [38] and 3D U-Net [39] (see Table 6 and Fig. 8). The statistical test results further indicate that the proposed method presents significant improvements compared with existing deep learning-based methods. Finally, our method is evaluated by using the online system for the BRATS 2015 testing dataset (see Table 7). The proposed method

achieves 0.845 (complete tumor), 0.698 (core tumor) and 0.600 (enhancing tumor) based on Dice scores, suggesting that our method can achieve a performance that can compete with existing brain tumor segmentation methods.

Despite the success of our approach, there also exist some limitations in our method. First, the postprocessing steps [17,19,20,42], such as conditional random fields (CRF), have been proven to be efficient in optimizing the segmenting results based on spatial consistency [42]. Therefore, an improvement of in segmentation performance of the proposed end-to-end method by utilizing efficient postprocessing steps is a topic of future research. Second, the proposed method with whole brain slices suffers from a data imbalance because the number of pixels for enhancing tumors is relatively small, which may worsen the segmentation performance of the trained network [42]. However, pointwise segmentation methods are commonly trained on equally sampled image patches from each class, which may lead to a higher performance for enhancing tumors [17]. Third, to effectively incorporate domain information in the axial, coronal and sagittal views, building a 3D fully convolutional network may further enhance the model capacity's to generate more precise segmentations. Finally, with more creative and effective architectures, such as capsule networks [43] and nonlocal neural networks [44], which have been proposed and have been proven to be efficient in general image analysis tasks, a more robust brain tumor segmentation model could be constructed by incorporating these structures in future work.

Table 7

Average performance of our system on the 110 test cases of BRATS 2015, as computed on the online evaluation platform.

Method	Dice score			PPV			Sensitivity		
	Complete	Core	Enhancing	Complete	Core	Enhancing	Complete	Core	Enhancing
Zhao et al. (2017)	0.80(−)	0.66(−)	0.57(−)	0.81(−)	0.79(−)	0.50(−)	0.83(−)	0.64(−)	**0.72**(−)
Kamnitsas et al. (2016)	0.836(−)	0.674(−)	**0.629**(−)	0.823(−)	**0.846**(−)	**0.64**(−)	**0.885**(−)	0.616(−)	0.656(−)
Inception-based U-Net	0.838(0.15)	0.676(0.23)	0.586(0.32)	**0.838**(0.17)	0.784(0.23)	0.611(0.29)	0.873(0.18)	0.667(0.24)	0.637(0.33)
Inception-based U-Net + up skip connection	0.844(0.14)	0.693(0.22)	0.595(0.32)	0.835(0.16)	0.767(0.22)	0.596(0.29)	0.884(0.17)	**0.700**(0.22)	0.661(0.32)
Inception-based U-Net + up skip connection + cascaded training strategy	**0.845**(0.14)	**0.698**(0.21)	0.600(0.32)	0.837(0.17)	0.791(0.22)	0.608(0.29)	**0.885**(17)	0.690(0.19)	0.659(0.32)

5. Conclusion

In this paper, we propose a novel end-to-end brain tumor segmentation method built upon an improved fully convolutional network by modifying the U-Net architecture. First, an innovative structure named the up skip connection is first proposed between the encoding portion and the decoding portion to enhance information flow. Then, an inception module is adopted in each block to help our network learn richer representations. In addition, an efficient cascaded training strategy is introduced to segment the brain tumor subregions sequentially. Different from existing patchwise methods, our model can automatically generate segmentation maps slice by slice, simplifying the experimental procedure. The proposed method was evaluated on the BRATS 2015 and the BRATS 2017 datasets. Quantitative evaluation of our method has revealed that our end-to-end segmentation method can achieve performance that can complete with state-of-the-art brain tumor segmentation approaches, suggesting the potential of the proposed improvements in brain tumor segmentation.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.03.014.

References

[1] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T.L. Jones, T.R. Barrick, F.A. Howe, X. Ye, Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI, Int. J. Comput. Assist. Radiol. Surg. 12 (2017) 183–203, https://doi.org/10.1007/s11548-016-1483-3.

[2] M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T.L. Jones, T.R. Barrick, F.A. Howe, X. Ye, Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels, Comput. Methods Progr. Biomed. 157 (2018) 69–84, https://doi.org/10.1016/j.cmpb.2018.01.003.

[3] S. Bauer, R. Wiest, L.-P. Nolte, M. Reyes, A survey of MRI-based medical image analysis for brain tumor studies, Phys. Med. Biol. 58 (2013) R97.

[4] S. Cui, L. Mao, J. Jiang, C. Liu, S. Xiong, Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network, J. Healthc. Eng. 2018 (2018).

[5] G. Tabatabai, R. Stupp, M.J. Van Den Bent, M.E. Hegi, J.C. Tonn, W. Wick, M. Weller, Molecular diagnostics of gliomas: the clinical perspective, Acta Neuropathol. (Berl.) 120 (2010) 585–592.

[6] M.B. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J.-G. Villemure, J.-P. Thiran, Atlas-based segmentation of pathological MR brain images using a model of lesion growth, IEEE Trans. Med. Imaging 23 (2004) 1301–1314.

[7] M. Prastawa, E. Bullitt, S. Ho, G. Gerig, A brain tumor segmentation framework based on outlier detection, Med. Image Anal. 8 (2004) 275–283.

[8] B.H. Menze, K. Van Leemput, D. Lashkari, M.-A. Weber, N. Ayache, P. Golland, A generative model for brain tumor segmentation in multi-modal images, Int. Conf. Med. Image Comput. Comput.-Assist. Interv, Springer, 2010, pp. 151–159.

[9] J. Kleesiek, A. Biller, G. Urban, U. Kothe, M. Bendszus, F. Hamprecht, Ilastik for multi-modal brain tumor segmentation, Proc. MICCAI BraTS Brain Tumor Segmentation Chall (2014) 12–17.

[10] S. Bauer, L.-P. Nolte, M. Reyes, Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization, Int. Conf. Med. Image Comput. Comput.-Assist. Interv, Springer, 2011, pp. 354–361.

[11] M. Goetz, C. Weber, J. Bloecher, B. Stieltjes, H.-P. Meinzer, K. Maier-Hein, Extremely randomized trees based brain tumor segmentation, Proceeding BRATS Chall.-MICCAI, 2014, pp. 006–011.

[12] R. Meier, S. Bauer, J. Slotboom, R. Wiest, M. Reyes, Appearance-and context-sensitive features for brain tumor segmentation, Proc. MICCAI BRATS Chall (2014) 020–026.

[13] M. Schmidt, I. Levner, R. Greiner, A. Murtha, A. Bistritz, Segmenting brain tumors using alignment-based features, Mach. Learn. Appl. 2005 Proc. Fourth Int. Conf. on, IEEE, 2005, p. 6.

[14] M. Soltaninejad, L. Zhang, T. Lambrou, G. Yang, N. Allinson, X. Ye, MRI brain tumor segmentation and patient survival prediction using random forests and fully convolutional networks, in: A. Crimi, S. Bakas, H. Kuijf, B. Menze, M. Reyes (Eds.), Brainlesion Glioma Mult. Scler. Stroke Trauma. Brain Inj, Springer International Publishing, 2018, pp. 204–215.

[15] M. Soltaninejad, X. Ye, G. Yang, N. Allinson, T. Lambrou, Brain Tumour Grading in Different MRI Protocols Using SVM on Statistical Features vol. 6, (2014).

[16] N.K. Subbanna, D. Precup, D.L. Collins, T. Arbel, Hierarchical probabilistic Gabor and MRF segmentation of brain tumours in MRI volumes, Int. Conf. Med. Image Comput. Comput.-Assist. Interv, Springer, 2013, pp. 751–758.

[17] K. Kamnitsas, C. Ledig, V.F.J. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2017) 61–78, https://doi.org/10.1016/j.media.2016.10.004.

[18] P. Dvorak, B. Menze, Structured Prediction with Convolutional Neural Networks for Multimodal Brain Tumor Segmentation, (2015), pp. 13–24.

[19] S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images, IEEE Trans. Med. Imaging 35 (2016) 1240–1251, https://doi.org/10.1109/TMI.2016.2538465.

[20] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.

[21] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, Int. Conf. Med. Image Comput. Comput.-Assist. Interv, Springer, 2015, pp. 234–241.

[22] Y. Xue, T. Xu, H. Zhang, L.R. Long, X. Huang, Segan: adversarial network with multi-scale l 1 loss for medical image segmentation, Neuroinformatics (2018) 1–10.

[23] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2015, pp. 3431–3440.

[24] A. Beers, K. Chang, J. Brown, E. Sartor, C. Mammen, E. Gerstner, B. Rosen, J. Kalpathy-Cramer, Sequential 3D U-Nets for Biologically-Informed Brain Tumor Segmentation, (2017) ArXiv Prepr. ArXiv170902967.

[25] H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo, Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks, Annu. Conf. Med. Image Underst. Anal, Springer, 2017, pp. 506–517.

[26] B. Erden, N. Gamboa, S. Wood, 3D Convolutional Neural Network for Brain Tumor Segmentation, (2018).

[27] G. Liu, J. Si, Y. Hu, S. Li, Photographic image synthesis with improved U-net, Adv. Comput. Intell. ICACI 2018 Tenth Int. Conf. on, IEEE, 2018, pp. 402–407.

[28] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, D. Rueckert, DRINet for medical image segmentation, IEEE Trans. Med. Imaging 37 (11) (2018) 2453–2462.

[29] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proc. IEEE Conf. Comput. Vis. Pattern Recognit, 2015, pp. 1–9.

[30] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M.A. Weber, T. Arbel, B.B. Avants, N. Ayache, P. Buendia, D.L. Collins, N. Cordier, J.J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C.R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K.M. Iftekharuddin, R. Jena, N.M. John, E. Konukoglu, D. Lashkari, J.A. Mariz, R. Meier, S. Pereira, D. Precup, S.J. Price, T.R. Raviv, S.M.S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H.C. Shin, J. Shotton, C.A. Silva, N. Sousa, N.K. Subbanna, G. Szekely, T.J. Taylor, O.M. Thomas, N.J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D.H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, K.V. Leemput, The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imaging 34 (2015) 1993–2024, https://doi.org/10.1109/TMI.2014.2377694.

[31] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, J.B. Freymann, K. Farahani, C. Davatzikos, Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Sci. Data. 4 (2017) 170117, https://doi.org/10.1038/sdata.2017.117.

[32] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, (2015) ArXiv Prepr. ArXiv150203167.

[33] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans. Med. Imaging 35 (2016) 1299–1312.

[34] P. Jaccard, The distribution of the flora in the alpine zone. 1, New Phytol. 11 (1912) 37–50.

[35] A.A. Taha, A. Hanbury, Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, BMC Med. Imaging 15 (2015) 29.

[36] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep learning-based classification of hyperspectral data, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (2014) 2094–2107.

[37] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, (2015) ArXiv151100561 Cs http://arxiv.org/abs/1511.00561 , Accessed date: 7 January 2019.

[38] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid Scene Parsing Network, ArXiv161201105 Cs, (2016) http://arxiv.org/abs/1612.01105 , Accessed date: 7 January 2019.

[39] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, (2016) ArXiv160606650 Cs http://arxiv.org/abs/1606.06650 , Accessed date: 9 February 2017.

[40] H. Li, G. Jiang, J. Zhang, R. Wang, Z. Wang, W.-S. Zheng, B. Menze, Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images, Neuroimage 183 (2018) 650–665, https://doi.org/10.1016/j.neuroimage.2018.07.005.

[41] S. Will, P. Martirosian, C. Würslin, F. Schick, Automated segmentation and volumetric analysis of renal cortex, medulla, and pelvis based on non-contrast-enhanced T1- and T2-weighted MR images, Magn. Reson. Mater. Phys. Biol. Med. 27 (2014)

445–454, https://doi.org/10.1007/s10334-014-0429-4.

[42] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating FCNNs and CRFs for brain tumor segmentation, Med. Image Anal. 43 (2018) 98–111.

[43] S. Sabour, N. Frosst, G.E. Hinton, Dynamic routing between capsules, Adv. Neural Inf. Process. Syst. 2017, pp. 3856–3866.

[44] X. Wang, R. Girshick, A. Gupta, K. He, Non-local neural networks, IEEE Conf. Comput. Vis. Pattern Recognit. CVPR, 2018.