Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

Automatic Brain Tumor Segmentation Using 3D Architecture Based on ROI Extraction

Jing Huang and Minhua Zheng
School of Mechanical, Electronic and Control Engineering
Beijing Jiaotong University
No. 3 Shangyuancun, Haidian District, Beijing 100044, P. R. China
{*17121265, mhzheng*}*@bjtu.edu.cn*

Peter X. Liu
Department of Systems and Computer Engineering
Carleton University
Ottawa, ON Canada
xpliu@sce.carleton.ca

***Abstract*— MRI (magnetic resonance image) analysis is crucial for diagnosis, monitoring, and treatment of brain tumors. Manual MRI segmentation of brain tumors requires professional knowledge and costs huge amount of time. Automatic segmentation for multimodal 3D MR images is thus very desirable for clinical applications. In this paper, we present an automatic brain tumor segmentation method based on the U-Net architecture, which is composed of ROI (region of interest) extraction and 3D segmentation networks with long and short skip connections. The introduced method is able to capture detailed features and improves the problem of the imbalanced classification of tumors' sub-regions. The method was evaluated on the BRATS 2017 dataset and the results were promising.**

***Index Terms*— Brain tumor segmentation, ROI extraction, 3D CNN.**

I. Introduction

Gliomas are the primary brain tumors, accounting for roughly 28% of all central nervous system tumors and 80% of all malignant brain tumors [1]. According to their histology and molecular, gliomas can be graded into low-grade gliomas (LGGs) and high-grade gliomas (HGGs), of which the LGGs carry a better prognosis while the HGGs almost certainly lead to death of patients [2]. Surgery, radiation therapy and chemotherapy are the three main treatment options for gliomas [3].

The treatment plans need to be devised under the prediction of growth rate, which can be obtained by segmentation of brain tumor using MRIs (magnetic resonance images) [4]. Although the segmentation of gliomas can be of great help in computer-aided diagnosis (CAD), it is tedious and inefficient to segment tumor images manually. Hence fully automatic and semi-automatic methods based on deep learning are being developed for providing more objective and reproducible measurements of tumors in clinical applications. There are several challenges for brain tumor segmentation tasks: 1) the brain tumors vary across patients, including their shapes, sizes and locations; 2) the original images are complex consisting of multiple modalities of MRIs; 3) the unbalanced distribution of tumor labels increases the difficulty of accurate segmentation.

Convolutional Neural Networks (CNNs) have achieved excellent performance in biomedical image segmentation. In [5], a convolutional network based on U-Net [6] architecture was proposed to achieve fully automatic brain tumor segmentation, and it obtained outstanding results for regions of whole tumors and tumor cores. Havaei et al. [7] created a two-pathway architecture which can capture detailed local and global features, and established the dependence of labels of pixels by stacking two CNNs. Pereira et al. [8] proposed a CNN with small kernels, which requires less number of weights to fix overfitting. In addition, LReLU activation function is applied to all the layers except the last, to train an architecture for LGGs and a deeper one for HGGs. Considering the connection between the MRI slices, the method based on 3D CNNs was proposed. One of the successful 3D CNNs for brain tumor segmentation was DeepMedic [9], which used a dual pathway to capture features for getting local and contextual information. A fully connected conditional random field (CRF) was used as post-process to get more accurate results. Based on encoder-decoder architecture, Myronenko [10] proposed a segmentation approach by adding an additional variational autoencoder (VAE) branch for obtaining extra guidance and regularizing the encoder part. In [11], a hybrid architecture, which combined the two-stream parallel network with three-path network, was proposed to segment the entire brain structure by classifying central pixel of individual patches considering both local and global features. Obviously, since the MRI data provides three-dimensional information, 3D CNNs are more suitable for brain tumor segmentation. However, 3D network inevitably has a higher requirement for computational resources and stronger constrains both on model complexity and the size of receptive field.

In this paper, an architecture that combines ROI (region of interest) extraction and 3D segmentation is proposed. The main contributions of this work are as follows:

This work was supported in part by the Fundamental Research Funds for the Central Universities (2017JBZ003, 2019RC028), National Science Foundation of China (61773051, 61761166011, 51705016) and Beijing Natural Science Foundation (4172048).

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

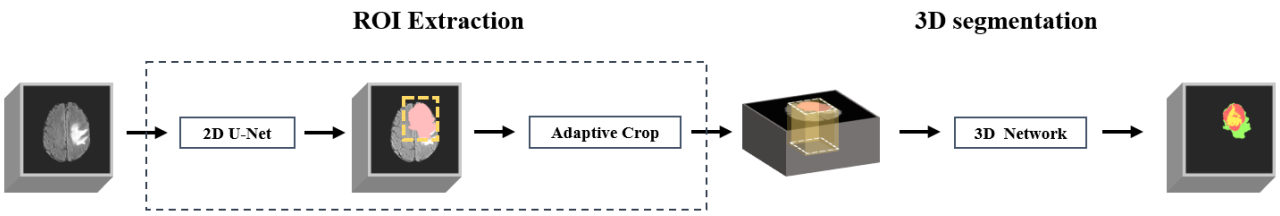

Fig. 1. The proposed framework for brain tumor segmentation. The ROI extraction takes a pair of MR images as input and outputs the 3D volumes that are obtained by 2D U-Net and adaptive crop. The 3D segmentation takes the results of ROI extraction as input and outputs the segmentation results of sub-strctures of brain tumors.

1) ROI extraction is applied by using 2D network and adaptive crop to reduce the interference of background information.

2) Additional short skip connections are added in 3D network to compensate the missing information caused by the convolution. With the long skip connections, detailed features, including semantic information and spatial information, can be captured.

3) The weighted Dice loss of 3D network is developed to improve the attention of tumor sub-regions during the segmentation process, which can solve the problem of inherent imbalanced classes of tumors.

The rest of the paper is organized as follows. Section II presents the method of brain tumor segmentation. Section III gives the details of the experiment. Section IV shows the results and discussions. The paper is concluded in Section V.

II. Method

A. Framework of Brain Tumor Segmentation

The segmentation framework is shown in Fig. 1. The U-Net architecture was improved to build a 2D network for ROI extractions and a 3D network for more accurate sub-region segmentation. Using a binary segmentation by 2D network, the whole tumor was obtained from each slice of multimodel MRI data. The three-dimensional boundary values of the whole tumor were obtained, which was effective to extract the 3D volume of the tumors from the original volume. The results of the ROI extraction were used as the input of the 3D network, to obtain the segmentation results of edema, tumor core and enhancing tumor.

B. Network Architecture

The 2D U-Net network was utilized with a large receptive and adaptive crop for ROI extraction. In 3D patch-based network we combined the long and short skip connections. The motivation of using ROI extraction was to reduce the interference of other brain tissue and improve the proportion of tumor area in the inputs of the 3D network.

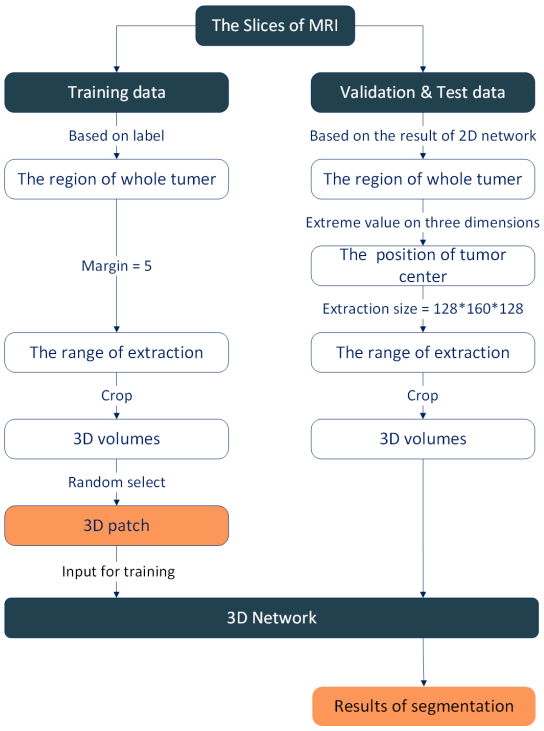

Fig. 2. ROI extraction for training data and validation & test data. For training, the input of 3D network is obtained by extracting the MRIs with the range of tumor label. For validation and test, 3D volumes are prepared to get the results via 3D network.

1) ROI Extraction: A 2D U-Net was used to obtain the regions of whole tumors. The U-Net is an encoder-decoder network that has good performance in the field of medical image segmentation. Combining the feature maps from encoder and decoder through skip connections, U-Net can capture semantic and spatial information. Different from the original U-Net, we applied the same padding for all convolutional layers, and used Dice loss [12] as loss function. The size of feature map is from 240×240 to 15×15. Based on the boundary value of the segmentation results of each slice, the tumor volume was extracted self-adaptively according to extreme value of all slices with a margin (set to be 5). The details of adaptive crop is shown in Fig. 2. The tumor volume

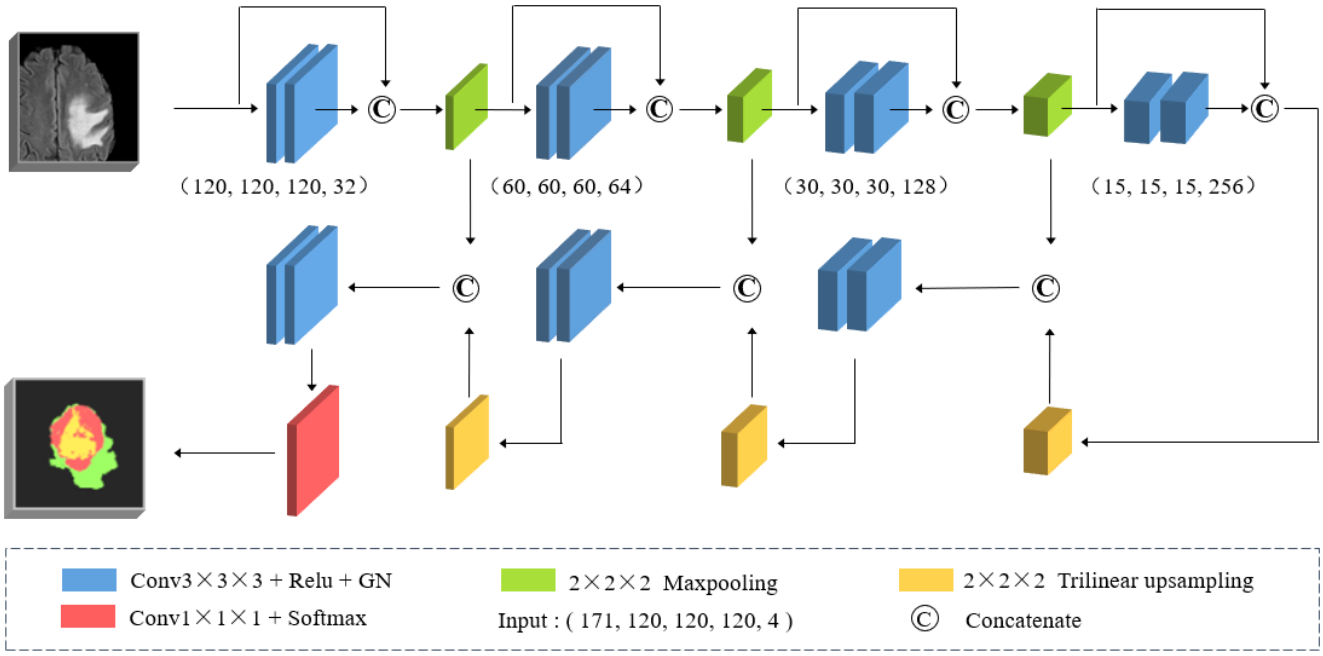

Fig. 3. Our 3D segmentation network with long and short skip connections, Group Normalization (GN) and weighted Dice loss. Input is a series of moltimodal 3D MRI crops of four channels with 32 filters initially. The output of the network is a three channel image with the same size as input, demonstrating the class of segmentation for three tumor sub-regions (edema, enhancing tumor, and tumor core).

is then used for 3D segmentation of sub-structures.

2) 3D Network Architecture: The architecture of the 3D segmentation network is shown in Fig. 3. Due to the problem of large memory consumption of 3D convolution, we used 3D patches which were randomly selected from the results of ROI extraction, and we reduced the filters' initial number to 32. The size of patches is $120 \times 120 \times 120$, and three blocks were applied in the encoder part. The encoder endpoint had size $15 \times 15 \times 15 \times 256$, which is four times smaller than the input image. Each block consists of two convolutional layers with 3×3 filter, using the same padding. During down sampling, max pooling with a step size of two was applied, and the size of filters is from 120×120 to 15×15. We added additional short skip connection to merge the feature before and after convolution of each block in order to compensate the information loss caused by the convolution. The long skip connection could concatenate the feature maps of encoder part to decoder part, so we could obtain more overall features with long and short skip connections. Instead of using Batch Normalization (BN), Group Normalization (GN) was applied considering the small batch size used in our network (batch size = 2).

C. Loss Function

Since the two-cross entropy loss function in U-Net is highly susceptible to category imbalanced, the Dice loss was used in our 2D network. Although the ROI extraction is used to narrow the range of targets, the problem of imbalanced classes still existed in 3D segmentation. In this case, Dice loss tends to fall into a local minimum in multi-segmentation tasks. To solve this problem, a weight ω is added in the Dice loss to increase the attention for sub-regions of tumors and reduce the influence of background.

$$\mathbf{L}_{WDC} = 1 - \frac{2}{|L|} \sum_{l \in L} \frac{\omega_l \sum_n r_{ln} p_{ln}}{\omega_l \sum_n r_{ln} + p_{ln}} \tag{1}$$

where l is the number of categories, r_{ln} is the standard value of the pixel n of the class l, p_{ln} is the predicted value of the pixel n of the class l, and ω is the weight that is determined by the number of pixels in this class as follows:

$$\omega_l = \frac{1}{(\sum_{n=1}^{N} r_{ln})^2} \tag{2}$$

D. Preprocessing

Non-standardized MR image values from different institutes and scanners are detrimental to brain tumor segmentation. Based on the mean and standard deviation of non-zero voxels, Z-score normalization was applied to each modality of each case individually. At the same time, the background area was set to zero, making the background area and the brain area more contrasting.

III. Experiment

A. Data Acquisition

The dataset used is the BRATS 2017 challenge[1]. It contains MR images from 285 patients, including 210 HGG images and 75 LGG images. We divided the dataset into training, validation and test sets, which contain images from 171, 57 and 57 patients, respectively. These scan slices of patients contain four modes, including native (T1), post-contrast T1-weighted (T1Gd), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (FLAIR). The results of manual annotations were approved by experienced neuro-radiologists. Annotations are composed of the GD-enhancing tumor (ET-label 4), the peritumoral edema (ED-label 2), and the necrotic and non-enhancing tumor (NCR / NET-label 1).

B. Implementation Details

Our method was implemented in Keras and trained on NVIDIA RTX TITAN 24GB GPU. We used Adaptive Moment Estimation(Adam) optimizer for training in both 2D and 3D networks, with initial learning rate as 5e-4. For the 2D network, the batch size is 64. For the 3D network, the batch size is 2. The initialization of the weights comes from a truncated normal distribution. The parameters of L2 norm regularization on the convolutional kernel is a weight decay of 1e-7. In the training stage, the volume we extracted is based on the labels. In the validation and test stage, we extracted 3D volume of fixed size $128 \times 160 \times 160$ according to the center of 2D network results.

C. Performance Evaluation

Dice similarity coefficient and Sensitivity are used to evaluate our segmentation results of each sub-region of brain tumors. The Dice score is the overlap area (the number of voxels in the intersection of segmentation and manual labels), divided by the number of voxels in the union of segmentation and reference.

$$\textbf{Dice} = \frac{2TP}{2TP + FN + FP} \tag{3}$$

where TP, FN, FP represent true positive, false positive and false negative, respectively. For a perfect segmentation, this value is one, and in the case of no overlap between segmentation and reference, the lowest possible value is zero.

Sensitivity is also an important value to evaluate the performance in medical field, which is defined as

$$\textbf{Sensitivity} = \frac{TP}{TP + FN} \tag{4}$$

[1]Data required: https://www.med.upenn.edu/sbia/brats2017/data.html

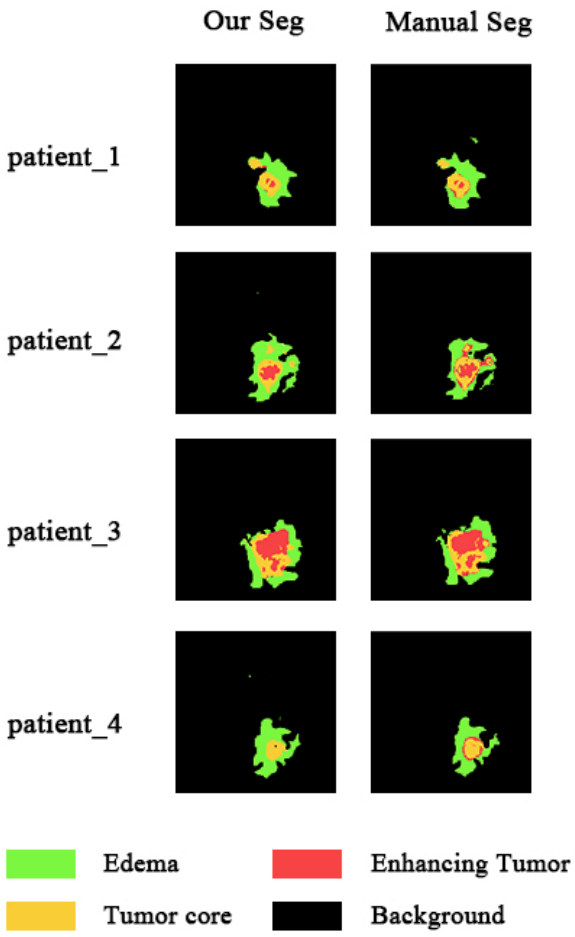

Fig. 4. The results of our segmentation method compared with manual annotation from experts

IV. Results and Discussions

Fig. 4 shows four examples of our brain tumor segmentation results, and the corresponding manual annotations. The green, yellow, red and black regions represent edema, enhancing tumor, tumor core and image background, respectively. It can be seen that our segmentation results coincide with the manual annotations quite well, which means that our segmentation method could provide supports for brain tumor diagnosis.

Table I shows the results of our method on the test set. The average Dice scores of our method for the segmentation of whole tumor, tumor core, and enhancing tumor are 0.9089, 0.7165, 0.8398, respectively. The sensitivity of our method for the segmentation of whole tumor, tumor core, and enhancing tumor are 0.9118, 0.7050, and 0.8182, respectively. For comparison, the results of the original 3D U-net and the improved 3D U-Net are also shown in Table I. The results indicate that the ROI extraction and the improvements on

TABLE I

DICE AND SENSITIVITY RESULTS OF THE ORIGINAL 3D U-NET, THE IMPROVED 3D U-NET AND THE PROPOSED METHOD ON BRATS 2017

	Dice[a]			**Sensitivity**[a]		
	Whole[b]	Core[b]	Enh.[b]	Whole	Core	Enh.
3D U-Net	0.8860	0.6716	0.8165	0.8573	0.6610	0.8035
Improved 3D U-Net	0.8926	0.6897	0.8328	0.8604	0.6705	0.7906
Proposed method	**0.9089**	**0.7165**	**0.8398**	**0.9118**	**0.7050**	**0.8182**

[a]Dice and Sensitivity: the higher the score, the better the segmentation performance.
[b]Whole stands for whole tumor, Core stands for tumor core and Enh. stands for enhancing tumor.

the 3D U-Net are effective in increasing the segmentation accuracy.

The segmentation results of edema were normally not reported in previous work, because for practical application, edema does not need to be removed if important nerve tissues are not compressed severely. The brain tissues in the edema area could slowly recover after surgery. However, the enhancing tumor area that supplies blood to the tumor and the tumor core area need to be removed during surgery. Therefore, the segmentation accuracy of the tumor core and the enhancing tumor is relatively more important. We need to distinguish these parts that must be removed from other parts so that doctors could plan the operation procedures better.

As the future work, we would improve the segmentation accuracy of the tumor core and the enhancing tumor areas by further improving the loss terms for these areas.

V. CONCLUSION

In this paper, we demonstrated a method of brain tumor segmentation for multimodal 3D MRIs. It contains two parts, i.e. the ROI extraction and the 3D segmentation network. The ROI extraction obtains the regions of whole tumors and provides the input for the 3D network. The 3D segmentation network is used for more detailed sub-region segmentation of tumors. The average Dice scores for whole tumor, tumor core, and enhancing tumor are 0.9089, 0.7165, and 0.8398, respectively. The proposed method can segment brain tumors automatically without professional physicians' interventions, and assist doctors' diagnosis and treatment planning for brain tumors.

REFERENCES

[1] S. Okuchi, A. Rojas-Garcia, A. Ulyte, I. Lopez, J. Ušinskienė, M. Lewis, S. M. Hassanein, E. Sanverdi, X. Golay, S. Thust *et al.*, "Diagnostic accuracy of dynamic contrast-enhanced perfusion mri in stratifying gliomas: A systematic review and meta-analysis," *Cancer medicine*, vol. 8, no. 12, pp. 5564–5573, 2019.

[2] D. N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, and D. W. Ellison, "The 2016 world health organization classification of tumors of the central nervous system: a summary," *Acta neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016.

[3] L. M. DeAngelis, "Brain tumors," *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, 2001.

[4] P.-Y. Kao, T. Ngo, A. Zhang, J. W. Chen, and B. Manjunath, "Brain tumor segmentation and tractographic feature extraction from structural mr images for overall survival prediction," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 128–141.

[5] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks," in *annual conference on medical image understanding and analysis*. Springer, 2017, pp. 506–517.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

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

[8] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in mri images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[9] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61–78, 2017.

[10] A. Myronenko, "3d mri brain tumor segmentation using autoencoder regularization," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 311–320.

[11] S. Sajid, S. Hussain, and A. Sarwar, "Brain tumor detection and segmentation in mr images using deep learning," *Arabian Journal for Science and Engineering*, vol. 44, no. 11, pp. 9249–9261, 2019.

[12] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016, pp. 565–571.