The Multimodal Brain Tumor Image Segmentation Based On Convolutional Neural Networks

Wang Mengqiao,Yang Jie, Chen Yilei ,Wang Hao
Key Laboratory of Fiber Optic Sensing Technology and Information Processing
Ministry of Education, Wuhan University of Technology
Wuhan, China
e-mail: 2457254782@qq.com, 466802324@qq.com

***Abstract*—In this paper, we propose an automatic brain tumor segmentation algorithm based on a 22-layers deep, three-dimensional Convolutional Neural Network (CNN) for the challenging problem of gliomas segmentation. To correct the bias field distortion of MRI images, we have added N4ITK method before intensity normalization. The use of several cascaded convolution layers with small kernels allows building a deeper CNN. During training, we add dropout to reduce overfitting, and adapt the batch normalization technique to speed up training. Our proposed method has been validated in the BRATS 2015 databases, the performance for the complete, core and enhancing regions was evaluated by the online evaluation platform with Dice matric of 0.84, 0.79, 0.75, Positive Predictive Value matric of 0.88, 0.86, 0.70 and Sensitivity matric of 0.82, 0.75, 0.86. Due to the GPU acceleration and improvement of the algorithm, our total training time is about 140 min. And through the experiments and results, we can prove that our method is time-saving and efficient.**

Keywords-brain tumor segmentation; convolutional neural Network; intensity normalization; evaluation matrics

I. INTRODUCTION

Gliomas are the most common primary brain tumors with a high mortality rate [1]. These gliomas can be divided into low-grade gliomas and high-grade gliomas. More than half of gliomas belong to the highest level of malignant brain glioblastoma, even under treatment, the median survival time is less than 15 months [2].

The brain tumor image segmentation provides important information for treatment planning and follow-up evaluation. Manual segmentation method is time-consuming and introduces inter-observer variability, thus, automatic segmentation algorithm has become a research focus. However, due to the large variability in location, size and shape, automatic segmentation method is still a big challenge. At present, many methods have been proposed. We can divide most current tumor segmentation methods into two categories. Generative approaches [3], [4] use the prior knowledge about the shape and spatial distribution of the tumor images. When dealing with binary classification problem, they are better than the discriminant models. But their learning and computing process is more complex. The common generative models mainly include: Bayesian Networks, Conditional Random Fields (CRF) [5] and Markov Random Fields (MRF) [4], etc. Discriminative approaches directly learn from annotated images without any prior knowledge. Discriminative models are more suitable for the recognition of multiple categories, but the relationship between variables is not clear. The common discriminative models mainly include: support vector machines (SVM) and random forest (RF), etc.

Images are more than pictures, they are data [6]. They contain a lot of information that our eyes cannot see. Methods known as Convolutional Neural Network [7]-[14] automatically learn complex features directly from images, so we don't need to extract features manually. Our approach is just based on convolutional neural network. [14] employs a two-pathway architecture, in which each pathway is responsible for learning about either the local details or the global feature of the image.

In this paper, we present a fully automatic approach for brain tumor segmentation in multimodal brain MRI image based on a CNN with the following main contributions:

1. We propose an automatic brain tumor segmentation algorithm based on a 22-layers deep, three-dimensional Convolutional Neural Network.
2. We add N4ITK [15] method to correct the bias field distortion of MRI images. In post processing, we add a condition judgment before threshold processing.
3. We make full use of the information that all modal sequences contained as well as the difference between them.

The paper is organized as follows: the proposed method is presented in Section II. The experiment and results are presented in Section III. Finally, the conclusions are presented in Section IV.

II. METHOD

A. Image Datasets

The paper uses the BRATS 2015 databases. The training image data consists of 20 MR scans from glioma patients, and then was tested on 100 MR. The training data set is equipped with ground truth but the test data set is blind. There are four MRI sequences available for every patient: T1-weighted (T1), T1-weighted and contrast-enhanced (T1c), T2-weighted (T2) and T2-weighted FLAIR (FLAIR). All images in the BRATS 2015 databases have the same

978-1-5386-2030-4/17/$31.00 ©2017 IEEE

dimension order and isotropic resolution, and they are all skull stripping and rigid registration.

B. Pre-processing

MRI suffers from the bias field or intensity inhomogeneity, and it impacts the analysis of the MRI data, so the correction of it is the first step. In this article, we use the N4ITK method. Then we normalize the data patches by subtracting the mean value and dividing the standard deviation.

C. Analysis of Convolutional neural network

Our purpose of this paper is to classify the pixel in the image into five classes accurately, so we can regard brain tumor image segmentation as a multi-class classification problem.

Convolution neural network extracts the image features of different layers from the shallower to the deeper by convolution operation. The weights of the convolution kernels are adjusted during the training process in order to produce characteristics which are the most suitable for classification. The main characteristics of CNN are that the weights of the kernels are shared and the connection is sparse. These characteristics make CNN easier to train and less prone to overfitting.

In this paper, we build a 22-layers deep, 3D CNN. The main architectures of CNN are in Table I.

TABLE I. THE MAIN ARCHITECTURES OF CNN

Layer(type)	Output	Filter size	Filter number
Conv3D	28×28×4	3×3×2	60
Conv3D	28×28×4	3×3×2	60
Conv3D	28×28×4	3×3×2	60
average-pooling	14×14×2	2×2×2	-
Conv3D	14×14×2	3×3×2	120
Conv3D	14×14×2	3×3×2	120
Conv3D	14×14×2	3×3×2	120
average-pooling	7×7×1	2×2×2	-
fully-connected	120	-	-
fully-connected	120	-	-
fully-connected	5	-	-

convolutional layers are separated by average-pooling with stride 2, and we use three 3×3×2 cascaded convolutional layers to replace one 7×7×4 convolutional layer, because they have the same effect but fewer weights. Each convolutional layer is followed by one activation layer and one batch normalization layer. After all convolutional layers are the fully-connected layers, and each fully-connected layer is followed by one activation layer and one dropout layer.

The paper from AlexNet has done the comparative experiment between Rectifier linear units (ReLU) and ordinary sigmoid function, the result has shown that the use of ReLU reduces the learning time greatly. Compared with ReLU, Leaky Rectifier linear units (LReLU) can avoid gradient disappeared. The LReLU function is defined as follow:

When applied the LReLU, we do the experiment again and again until find an appropriate value of α. Here, A final value of α is 0.3. For activation layer, we choose Leaky Rectifier linear units (LReLU) as the activation function, but in the last layer, we use softmax.

Pooling layers can reduce the coupling of the model by summarizing the outputs of neighboring groups of neurons in the same kernel map. Traditionally, the neighborhoods summarized by adjacent pooling units do not overlap. The most common pooling methods are max-pooling and average-pooling, in this article, we choose average-pooling method.

In the fully-connected layers, we add dropout to reduce overfitting. it disconnects some nodes of network's hidden layer with a selected probability, but this is only temporary, these nodes may work again when the next samples input.

Each layer's inputs would change during training. This might slow down the training speed because of requiring lower learning rates and careful initial parameters, we call this phenomenon as the "internal covariate shift", and it affects 3D CNN more severely than 2D CNN. In order to solve this problem, we adopt the Batch Normalization technique. In traditional deep neural networks, large learning rate may lead to the gradients explosion or disappear. To solve this problem, Batch Normalization prevents small changes in layer parameters by normalizing activations throughout the network.

CNN is trained patch by patch. A batch is processed by the network for one training iteration of optimization algorithm, we use Stochastic Gradient Descent (SGD) as the optimization algorithm. This step can update the network's parameters such as weights and biases in order to minimize the expected loss, here we choose categorical cross entropy as the loss function:

$$J(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\log(p_{c,i}) \tag{1}$$

where N is the size of batch, $p_{c,i}$ represents the probabilistic that pixel belongs to class c.

D. Post-Processing

After segmentation, some small clusters or outliers may be classified wrongly, so we use connected domain labeling method to deal with this problem. First, we calculate the area of each connected domain. Then, we choose a suitable threshold. Last, we remove the clusters that areas are smaller than threshold. For each patient image, tumor size of the layer near the tumor edge tend to be small, so sometimes they might be removed by mistake. To address it, we can add a condition judgment before threshold processing.

$$f(x) = \max(0,x) + \alpha \min(0,x) \tag{2}$$

III. EXPERIMENTS AND RESULTS

A. Experiments

Take each pixel as the center, we extracted around 100000 patches of fixed size to train the CNN. The training data contains five classes: normal tissue, necrosis, edema,

non-enhancing, and enhancing tumor. In order to balance the classes, we selected the same samples for every class.

During training, we adopt 5 fold cross-validation method to train the network. The training result is shown in Fig. 1. The CNN was implemented based on Theano in an Intel Core i7 2.8 GHz machine with a GPU NVIDIA GeForce GTX 1050. For each epoch, the training time is about 7 min. And we can see from the result, the performent reach the best when the epoch number is around 20. So the total training time is about 140 min.

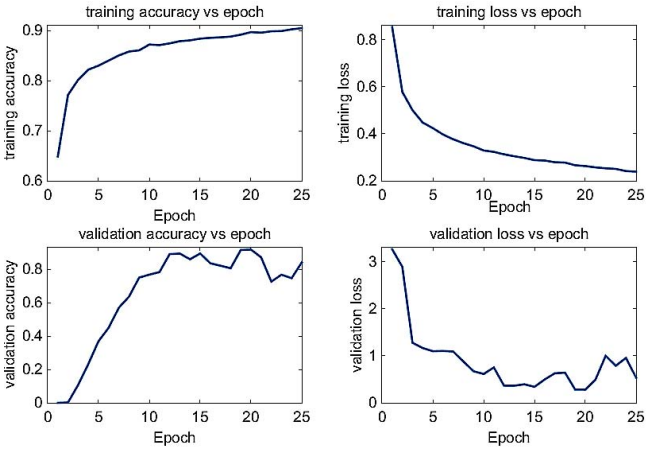

Figure 1. The training result.

B. Results

In this article, we used online evaluation platform to test the results. The website of evaluation platform is: https://www.virtualskeleton.ch/BRATS/Start2015#evaluation. The evaluation contains three metrics: Dice, Positive Predictive Value and Sensitivity. For each metric, we use three regions to calculate. These three regions are the whole tumor, the core tumor (including necrosis, non-enhancing and enhancing tumor) and the enhancing tumor. We used 100 MR scans to validate the algorithm, the segmentation result of one example from this test data set is shown in Fig. 2.

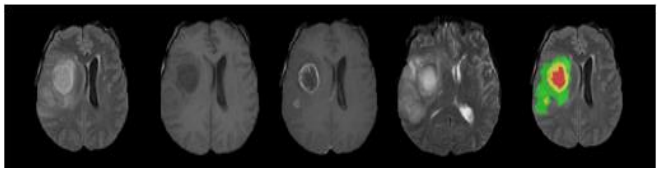

Figure 2. Segmentation result based on T1, T1C, T2 and FLAIR images.

The pictures from left to right are: FLAIR, T1, T1c, T2 and segmentation result. In the last picture, the subregions are necrosis (red), edema (blue), enhancing (yellow) and non-enhancing (green).

Then we use boxplot to count the evaluation metrics of all test data, and this is implemented by statistical software SPSS 19.0, the result is shown in Fig. 3. And the mean value of each parameter is shown in Table II.

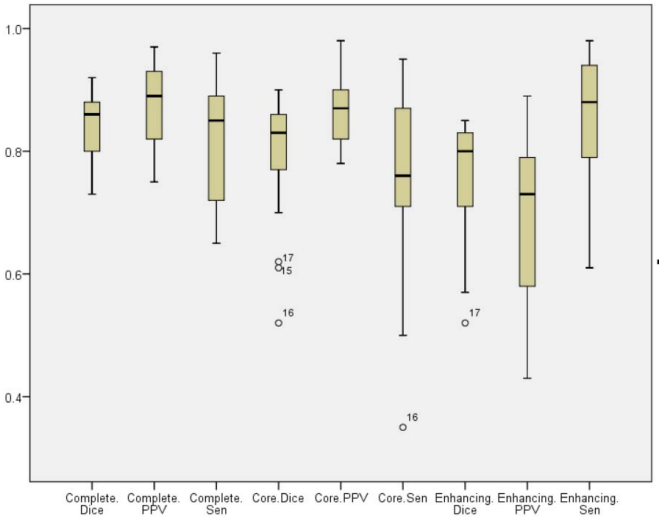

Figure 3. Boxplot of the evaluation metrics obtained in the BRATS 2015 databases.

TABLE II. THE RESULTS OF SEGMENTATION

	Complete			Core			Enhancing		
	Dice	PPV	Sensitivity	Dice	PPV	Sensitivity	Dice	PPV	Sensitivity
Proposed	0.84	0.88	0.82	0.79	0.86	0.75	0.75	0.70	0.86
References[13]	0.88	0.91	0.86	0.76	0.90	0.74	0.73	0.72	0.81

IV. CONCLUSIONS

In this paper, we put forward a novative segmentation method based on the 3D Convolutional Neural Networks. First, we apply the N4ITK method to correct bias field. Then we choose the activation function and initialization method that perform well by comparative experiment. During training, we add dropout to reduce overfitting, and adapt the batch normalization technique to speed up training. Finally, we use online evaluation platform to test the results, and comparing the results with References [13], some matrics such as dice in core regions, dice in enhancing regions and sensitivity have improved. Because of the GPU acceleration and improvement of the algorithm, our training time has been reduced greatly.

ACKNOWLEDGMENT

This work was supported by National Science Foundation of China (NSFC) under Grant 51479159 and The Excellent Dissertation Cultivation Funds

of Wuhan University of Technology under Grand 2016-YS-063.

REFERENCES

[1] Menze B H, Jakab A, Bauer S, et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)[J]. IEEE Transactions on Medical Imaging, 2015, 34(10):1993.

[2] Dolecek T A, Propp J M, Stroup N E, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009.[J]. Neuro-oncology, 2012, 14 Suppl 5(sup 2):v1.

[3] Gooya A, Pohl K M, Bilello M, et al. GLISTR: glioma image segmentation and registration.[J]. IEEE Trans Med Imaging, 2012, 31(10):1941-1954.

[4] Menze B H, Van L K, Lashkari D, et al. A generative model for brain tumor segmentation in multi-modal images.[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer-Verlag, 2010:151-9.

[5] Kamnitsas K, Ledig C, Newcombe V F, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation[J]. Medical Image Analysis, 2017, 36:61.

[6] Gillies R J, Kinahan P E, Hricak H. Radiomics: Images Are More than Pictures, They Are Data[J]. Radiology, 2016, 278(2):563.

[7] Zikic D, Ioannou Y, Brown M, et al. Segmentation of Brain Tumor Tissues with Convolutional Neural Networks[C]// Miccai Workshop on Multimodal Brain Tumor Segmentation Challenge. 2014.

[8] Urban G, Bendszus M, Hamprecht F A, et al. Multi-modal brain tumor segmentation using deep convolutional neural networks[C]// MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings, winning contribution. 2014.

[9] Havaei M, Davy A, Wardefarley D, et al. Brain tumor segmentation with Deep Neural Networks.[J]. Medical Image Analysis, 2017, 35:18.

[10] V. Rao, M. Sharifi, and A. Jaiswal. Brain tumor segmentation with deep learning[C]// MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 56–59, 2015.

[11] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.

[12] Dvořák P, Menze B. Local Structure Prediction with Convolutional Neural Networks for Multimodal BrainTumor Segmentation[J]. 2015.

[13] Pereira S, Pinto A, Alves V, et al. Brain Tumor Segmentation using Convolutional Neural Networks in MRI
Images. [J]. IEEE Transactions on Medical Imaging, 2016, 35(5):1-1.

[14] Havaei M, Dutil F, Pal C, et al. A Convolutional Neural Network Approach to Brain Tumor Segmentation [J]. 2015.

[15] Tustison N J, Avants B B, Cook P A, et al. N4ITK: improved N3 bias correction.[J]. IEEE Transactions on Medical Imaging, 2010, 29(6):1310.