



# DETECTION OF BRAIN TUMOR USING CNN BASED SEMANTIC SEGMENTATION OF MRI IMAGES

Dr.R.Satish Kumar <sup>[1]</sup>, Anusuya, <sup>[2]</sup>.

PG Scholar: Anusuya <sup>[2]</sup>, Guided by Dr.R.Satish Kumar, Principal  
Department of Medical Electronics Engineering,  
Sengunthar College of Engineering, Tiruchengode, Tamil Nadu, INDIA  
[harsiga@gmail.com](mailto:harsiga@gmail.com), [anusundaram@gmail.com](mailto:anusundaram@gmail.com)

## Manuscript History

Number: IJIRAE/RS/Vol.07/Issue03/Special Issue/01.MRAESCE10116

Received: 15, February 2020

Final Correction: 27, February 2020

Final Accepted: 10, March 2020

Published: 14, March 2020

**Editor:** Dr.A.Arul Lawrence selvakumar, Chief Editor, IJIRAE, AM Publications, India

Copyright: ©2020 This is an open access article distributed under the terms of the Creative Commons Attribution License, Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

**Abstract:** Accurate segmentation of brain tumor is an indispensable component for cancer diagnosis and treatment. In this paper, we propose a novel brain tumor segmentation method based on multi cascaded convolutional neural network (MCCNN) and fully connected conditional random fields (CRFs). The segmentation process mainly includes the following two steps. First, we design multi-cascaded network architecture by combining the intermediate results of several connected components to take the local dependencies of labels into account and make use of multi-scale features for the coarse segmentation. Second, we apply CRFs to consider the spatial contextual information and eliminate some spurious outputs for the fine segmentation. In addition, we use image patches obtained from axial, coronal, and sagittal views to respectively train three segmentation models, and then combine them to obtain the final segmentation result. The validity of the proposed method is evaluated on three publicly available databases. The experimental results show that our method achieves competitive performance compared with the state-of-the-art approaches.

**Keywords:** Brain tumor segmentation, convolutional neural network, multi-cascaded convolutional neural network, conditional random field, multi-modality.

## INTRODUCTION

Brain tumors are one of the common diseases of the nervous system and have great harm to human health, and even lead to death. Glioma is one of the intracranial tumors with the highest mortality and morbidity [1]. It is usually divided into high-grade glioma (HGG) and low-grade glioma (LGG), and the average life expectancy of patients who have evolved into HGG is about two years. Many imaging techniques have been employed for investigating brain tumors such as magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). And if the person delineating the region of interest (ROI) is not a well-trained technologist, it will usually yield poor segmentation results [4], [5]. Therefore, automatic segmentation or semi-automatic segmentation attracts the attention of researchers. The existing automatic segmentation or semiautomatic segmentation methods can be roughly classified as either generative model based methods or discriminative models based methods [6]. The generative models rely on prior knowledge of the brain anatomy, and their learning process is relatively complex. Prastawa et al [7] proposed a typical generative model of MR brain images, which mainly based on probabilistic image atlas. Other methods that require prior knowledge can be found in [8] [10]. On the contrary, discriminative models often rely on low level image features. They usually employ a discriminative classifier to transform local features into class probability, which is more suitable for multi-category identification problems. Typical discriminative models include conditional random fields [11] [13], random forests [14] [16], support vector machines [17], decision forests [18], etc.

In recent years, the deep learning based automatic segmentation methods have shown great advantages in medical image analysis. CNN is one of the most popular deep learning models, which can extract features that are favorable for classification from the original data. Compared with the traditional machine learning methods, CNN has more powerful capability of feature learning and characterization and can reduce the complexity of model due to the nature of weight sharing. Based on the advantages of deep learning, it has attracted an increasing attention in brain tumor segmentation. Proposed a two-path CNN model in which one channel extracts local detail features and the other extracts global context features. Most of the above brain tumor segmentation methods built CNN upon image patches. These patch-based methods classify each image patch into different classes, ignoring the spatial consistency of the entire image. They usually assumed that the label for each voxel is independent and did not take the local dependencies of labels into account. To consider the local dependencies of labels, Havaei et al. developed a cascaded architecture (Input Cascade CNN), which can obtain a better segmentation performance. Furthermore, they adopted a two-stage training strategy to solve the imbalance of label distributions.

Recently, the advantages of the multi-scale features from CNNs have been demonstrated in segmentation tasks. Generally, there are two ways to extract multiscale features. One way is to use the feature maps from different levels of a network to represent multi-scale features. The other is to pass different scaled versions of the input image using the same network. In [13], the multi-scale features were extracted by Kamnitsas et al. using two patches with different sizes. We propose a new network architecture from Input Cascade CNN to obtain more discriminative multi-scale features through multi-cascaded networks. It is undeniable that CNN has achieved remarkable achievement. Some researchers adopt 3D-CNN models [13] or 2.5D models to deal with 3D images. They also trained the segmentation models from the axial, sagittal, and coronal views respectively. During the testing stage, they averaged the softmax outputs in these three views to obtain the final results. Although these methods achieve excellent performance, they increase computational complexity and memory consumption. Hence, some mathematical models, e.g., Markov Random Fields (MRFs) or Conditional Random Fields (CRFs), are usually introduced to consider the spatial contextual information [12], [13]. Formulated CRFs as neural networks and trained them with fully convolutional neural networks (FCNNs), their training process was cumbersome and computationally expensive. As a trade-off, we employ fully connected CRFs as a segmentation step to take the appearance and spatial consistency of the segmentation results into account. The contributions of this study can be summarized as follows.

We extend Input Cascade CNN to MCCNN architecture to obtain multi-scale features through multi-cascaded networks on brain MR images. The discriminative features are efficiently extracted by passing multiple scaled versions of the input image through three parallel sub-networks. In each sub-network, the high-level and low-level information are utilized jointly by fusing the features of the front-layer. In addition, different cascaded forms are employed to take the dependencies of labels and the neighborhood information of the central pixel into account to achieve a smoother boundary. We propose an efficient coarse-to-fine segmentation framework for brain tumor segmentation. Our framework mainly consists of a coarse segmentation and a fine segmentation. The former uses the proposed MCCNN to obtain a probability map by considering multi-scale features on the multi-modal brain images. The latter employs fully connected CRFs to refine the coarse segmentation by smoothing tumor edges and eliminating false positives.

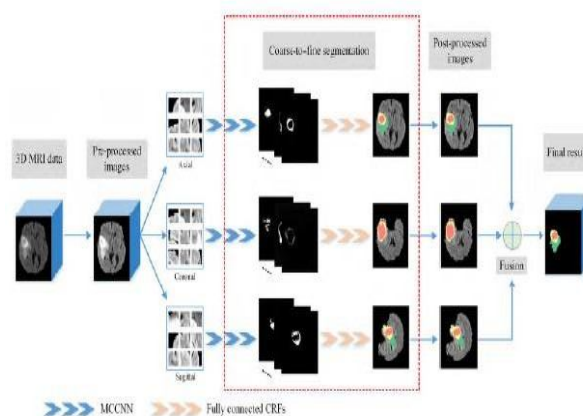


Fig. 1. Flowchart for the Proposed Method

Thus, in reference to the CRS, it is possible to say that patients undergoing symptoms of cardio vascular diseases (CVD) may also be suffering from CKD. Since it is already known that different CVDs leave characteristics traces in the patient’s ECG, it is also possible to detect the presence of CKD from the same ECG of the patient if the traces for CKD are known. From various important studies, it was found that ECG of any patient undergoing CKD, shows some significant changes which can be traced back to CKD [9, 10]. The QT duration was seen prolonged in CKD patients such that 460ms in female while 450ms in men and the QRS amplitude was also seen to be increased by approximately 0.18Mv. The traditional method of detecting kidney diseases are the invasive methods of blood tests (e.g. GFR test) or kidney biopsy, and the non-invasive methods are ultra sound imaging or urine tests to check for creatinine levels, and End stage CKD patients. Thus, this model opens up an option for patients to detect their kidney disease through a simple non-invasive way by means of their available bio signals i.e. ECG signal. ECG signal is widely used for CVD analysis as almost all types the heart abnormalities can be detected from this. Since, any patient with CVD are very likely to undergo an ECG test, through this model, the presence of CKD can also be detected from the same ECG, by checking for prolonged QT intervals and also checking RR intervals which can be used to measure the HRV. The following sections of this paper will discuss about the ECG data processing, model formation and validation stages.

### PROPOSED METHOD

As shown in Fig. 1, the proposed method includes following steps, i.e., pre processing, coarse segmentation based on MCCNN, fine segmentation based on fully connected CRFs, post-processing, and fusion of the segmentation results of three different views. The details of each step are presented as below.

#### PRE-PROCESSING

Since MRI scans are affected by the bias field distortion [6], which makes the intensity of the same tissues to vary across the image. So we adopt the N4ITK method proposed by Tustison et al. [21] to correct the bias field of MRI data. In addition, it will also vary even if the images of the same patient are obtained in the same scanner at different time points, or in the case of pathology [22]. Therefore, in order to reduce the contrast and intensity ranges between patients and acquisitions, we also employ the method proposed by Nyúl et al. [48] to perform the intensity normalization on each modality. The pre-processing steps are as follows.

- Step 1. Remove the 1% highest and lowest intensities.
- Step 2. Apply an N4ITK bias correction to T1c modality.
- Step 3. Perform the intensity normalization on each modality.
- Step 4. Normalize the data within each input channel by subtracting the mean of all pixels from each pixel in the image of each channel and dividing by the standard deviation.

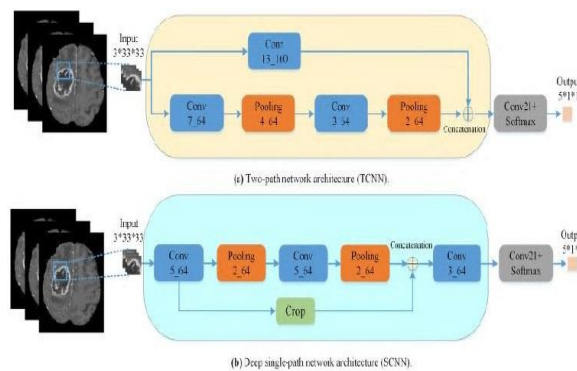


Fig. 2. Baseline Network architectures

#### COARSE SEGMENTATION USING OUR MCCNN ARCHITECTURE

CNN is a very powerful method in the field of image segmentation [2]. The basic architecture of CNN includes input layer, hidden layer, and output layer. The value of output can be obtained from the forward propagation of CNN, and the bias and weight can be adjusted by back-propagation. The classic CNN consists of three parts: convolution layers, pooling layers, and fully connected layers. There are usually a multitude of parameters that need to be trained in a deep learning model of CNN. In order to ensure that there are enough training samples and consider the complexity of training, our MCCNN is trained using image patches, which are extracted randomly from slice of different perspectives i.e., axial, coronal, and sagittal views in 3D brain MRI data. Overall, the glioma images are roughly segmented in this process, that is, the entire glioma contour is obtained as completely and accurately as possible.

For this purpose, we conduct a large number of experiments to design and select the network. Firstly, we propose two basic CNN architectures without cascade as baseline connected components. Then, we explore the performance of simple cascade architectures and the multi-cascaded architectures. Details of these models are as follows.

### METHODOLOGY

We train a suitable model by minimizing the energy function of CRFs1, which can be modeled as, computational complexity.

#### FINE SEGMENTATION USING FULLY CONNECTED CRFS

Although the form of cascading has been used to consider the local dependencies of  $\forall i, j, i \neq j$  (1) labels, it is not enough for medical images. This is mainly because anatomical structures have complex shapes that are difficult to model. Further, the temporal or while the pairwise potential measures the influence between any pair of pixels which is defined as, spatial correlation of MRI data also plays important role in the segmentation, which needs to be considered in the method. Therefore, it is necessary to further refine

$$\psi_P(y_i, y_j) = \sum^M \omega^{(m)} k^{(m)}(f_i, f_j) \quad (2)$$

the probability map obtained by CNN. In recent years, CRFs have been widely used in many medical image applications, since they show good performance for modeling the complex dependencies in spatial data. For the segmentation of brain tumors, CRFs can be used not only to model the relationship between image pixel features and its label,  $k(1)$  represents the appearance kernel, which tries to assign the same class labels to nearby pixels with similar intensity.  $k(2)$  represents the smoothness kernel, aiming to remove small isolated regions. They are respectively defined by,

$$k^{(1)}(f_i, f_j) = \exp\left(-\frac{|s_i - s_j|}{\theta_\alpha} - \frac{|e_i - e_j|}{\theta_\beta}\right) \quad (3)$$

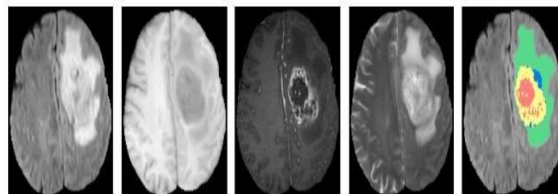


Fig. 3. Example of the Brain MRI Data

But also to model the dependencies between neighborhood pixel features and their labels [11]. Applied CRFs to image segmentation method is cumbersome and computationally expensive. By contrast, in this paper we employ the CRFs as a fine segmentation to refine the segmentation of MCCNN. It not only improves the segmentation accuracy of MCCNN, but also has lower where  $e_i$  and  $e_j$  are the intensities of pixel  $i$  and  $j$ ,  $s_i$  and  $s_j$  are the corresponding spatial coordinates.  $f_i$  and  $f_j$  represent the feature vectors of the pair of pixels, namely the intensity and position information.  $\theta_\alpha$ ,  $\theta_\beta$  and  $\theta_\gamma$  represent the parameters of the Gaussian kernels, respectively.

### POST-PROCESSING AND FUSION OF THE SEGMENTATION RESULTS

For the post-processing, we employ a simple method based on connected components to remove small spurious areas due to the high intensity value near the edges of the brain image. To utilize the spatial information of 3D MR brain image, we train three segmentation models using patches extracted from different perspectives, i.e., axial, coronal, and sagittal views. Then, the test images can be segmented slice by slice using these three models, respectively. Finally, we perform post-processing on these three results as described above, and then fuse them using a majority voting strategy.

### BRAIN MRI DATA

The proposed method is tested and evaluated on the BRATS 2013,2 BRATS 2015,3 and BRATS 20184 databases. The BRATS 2013 database contains training dataset of 30 glioma patients, comprising of 20 HGG cases and 10 LGG cases. The testing dataset contains 10 patient subjects with HGG. All the brain images were skull stripped and had the same orientation. For each patient

### EVALUATION METRICS

According to the protocol in the BRATS databases [6], the tumor structures are divided into three sub-regions for each patient, which are defined as follows:

- The complete tumor region (including all four sub tumoral classes, i.e, label 1, 2, 3, and 4).
- The core tumor region (including complete tumor region but excluded "edema" region, i.e., label 1, 3, and 4).
- The enhancing tumor region (only including "enhancing" region, i.e., label 4).

For each tumor region, we use Dice Similarity Coefficient (DSC) as well as Positive Predictive Value (PPV) and Sensitivity for the evaluation. DSC provides a measurement criterion for overlapping areas between manual delineated brain tumor regions and our segmentation results. subject, there are four MRI modalities, i.e., T1, T1c, T2, and FLAIR. All modalities. were co-registered to the T1c sequence in order to homogenize these data and then these images were re sampled to 1 mm isotropic resolution in a standardized axial orientation with a linear interpolator [6]. In addition, the manual annotations have been used as the ground truth in the training dataset. No ground truths are available for the testing dataset. The labels were divided into five classes, namely healthy tissues (label 0), necrosis (label 1), edema (label 2), non-enhancing (label 3), and enhancing where TP, FP, and FN denote the numbers of true positive, false positive, and false negative measurements, respectively. PPV is used to evaluate the number of TP and FP, which is defined as, In addition, Sensitivity is a useful measure of the number of TP and FN and it is given by, tumor (label 4). Examples of the brain MRI data along with the ground truth are shown in Fig 3.

## EXPERIMENTAL SCHEME AND IMPLEMENTATION DETAILS

In order to verify the effectiveness of the proposed method, a series of comparative experiments are conducted on the BRATS 2013 database. Firstly, we perform an experiment to evaluate the segmentation performance of the proposed two MCCNN architectures in comparison with two basic networks described previously and the simple concatenation forms between them. Secondly, to verify the validity of the proposed coarse-to-fine segmentation framework, we compare the segmentation performance of MCCNN with and without used CRFs. In addition, to investigate the effect of each modality on the proposed method, we carry out an experiment to compare the performance of our models based on input data of four imaging modalities (Flair, T1, T1c, and T2) and three imaging modalities (any three of these four imaging modalities).

It should be noted that all the experiments above are conducted on the images extracted from axial view. Moreover, in order to consider the information from different views, we evaluate the segmentation results of three different perspectives and demonstrate the effectiveness of fusing them. Finally, we compare the proposed method with some state-of-the-art approaches on the BRATS 2013, BRATS 2015, and BRATS 2018 databases. In the experiments, the training process consists of two stages. In the first stage, the coarse segmentation is carried out through the proposed MCCNN network, and the probability maps of three perspectives are obtained. In the second stage, refined segmentation results are obtained through the fully connected CRFs. The hyper- parameters of the different network architectures (kernel and pooling size for each layer and the number of feature maps) are found using the validation set, consisting of two HGG cases. Other hyper-parameters of the model are also tuned with the validation set. The final results of our experiments are obtained using ReLU as activation function in convolution layers. The learning rate is set by  $10^{-4}$ . Momentum is 0.9 and weight decay is  $10^{-3}$ .

## RESULTS ANALYSIS

### PERFORMANCE OF THE PROPOSED NETWORK ARCHITECTURES

It is generally accepted that the learning ability of the network will improve by increasing the complexity of the network. In this study, we provide a comparison based on different network architectures, i.e., basic CNN architectures (TCNN and SCNN), simple concatenation architectures (TTCNN, SSCNN and TSCNN), and multi-cascaded architectures (MCCNN1 and MCCNN2). It is need to note that the proposed multi- cascaded architectures consist of three sub- networks, because going deeper will not improve the performance. Experiments demonstrate that this architecture achieves comparable segmentation results while guaranteeing the time efficiency. Therefore, MCCNN2 is selected as the final recommended multi-cascaded network architecture. In the following part, we will use MCCNN to represent this network architecture.

### EVALUATION OF THE IMPACT OF EACH MODALITY

As can be seen from Fig. 3, different modalities in MRI data emphasize different information. In order to explore the effect of each modality on training, we establish segmentation models respectively using all four imaging modalities, i.e., Flair, T1, T1c, and T2, and the combination of any three imaging modalities. In general, different modalities have different features. Since the T1c sequence strengthens the features of tumor boundary, making the boundary clear and easy to distinguish [2]. The information reflected by this sequence is effective for the identification and classification of core tumor and enhancing tumor regions. Therefore, in the case of removing the T1c sequence, the model cannot learn robust features of the tumor edge, resulting in poor segmentation results.

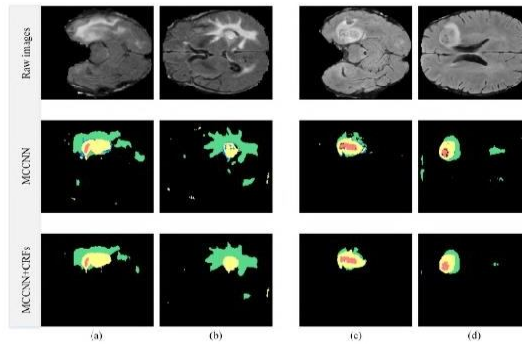


Fig 4. Visual effect on the Brats 2013 Dataset

### EVALUATION OF THE VALIDITY OF FUSING THE SEGMENTATION RESULTS OBTAINED BY THREE PERSPECTIVES

Considering that more robust information can be extracted from different perspectives, we separately train three segmentation models using patches obtained from three different views, i.e., axial, coronal, and sagittal views. These three models are used to segment the brain images to produce three segmentation results, which are then fused to obtain the final segmentation result. It shows that fusing the segmentation results usually yields better segmentation performance at any stage. It is worth noting that after the CRFs, the fusion results are greatly improved without the post- processing compared with the negligible improvement after applying the post- processing.

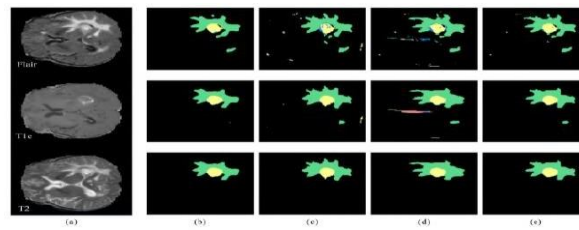


Fig 5 Visual effects from three views on case 0301 of the BRATS 2013

We can find that the segmentation results of the axial view are optimal in the coarse segmentation stage. In the fine and final segmentation stages with post-processing, the coronal view is the best. And in any stage, the segmentation results from sagittal view are the worst. Overall, the best performance is achieved by combining these three different perspectives.

### RESULTS ON THE BRATS 2015 DATABASE

To illustrate the robustness of our method, we also use the proposed method to segment the 3D MR images provided by the BRATS 2015.

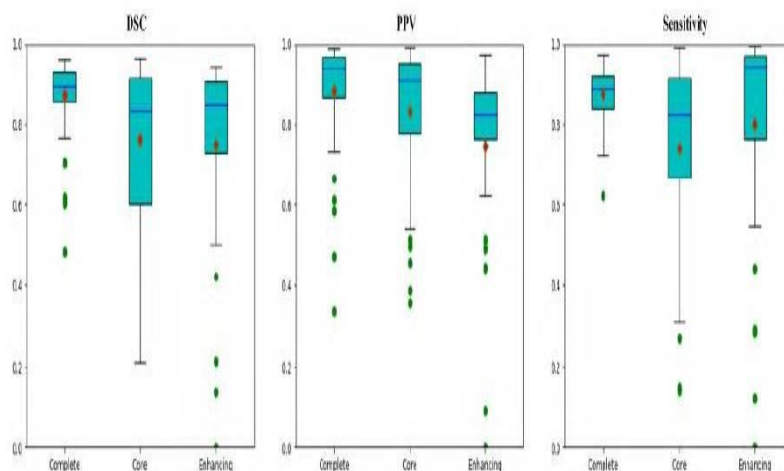


Fig 6 Boxpolts of the Results on the Brats 2015

## RESULTS ON THE BRATS 2018 DATABASE

To further evaluate the effectiveness of the proposed method, we conduct an experiment on the BRATS 2018 dataset. The data provided by BRATS 2018 differs from the data provided during the previous BRATS challenges.

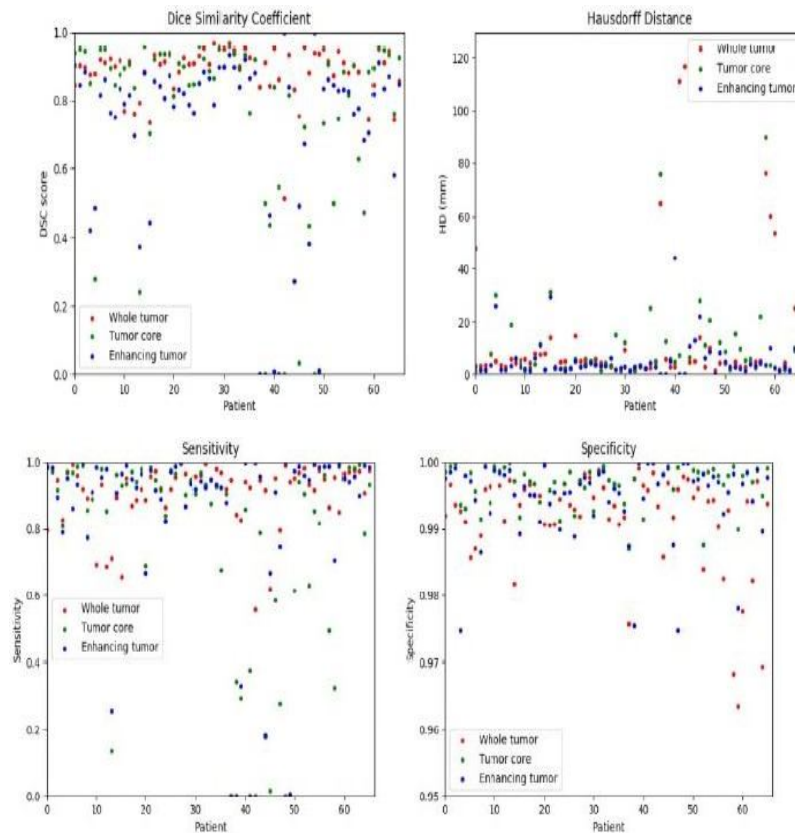


Fig 7 Scatter diagram of the results on the BRATS 2018 dataset

## CONCLUSION

In this paper, we firstly proposed a novel MCCNN architecture to extract more discriminative multi-scale features for brain tumor segmentation. Especially, we considered different network frameworks and studied their impact on the segmentation performance and found that high performance can be achieved by stacking three CNNs to model the dependencies of labels and exploiting the nature of fully convolution. Secondly, we presented a coarse to segmentation framework by combining MCCNN and fully connected CRFs. In this framework, we first performed a coarse segmentation using MCCNN in order to obtain the tumor contour. Then, we used CRFs to refine the segmentation results by minimizing energy function. In particular, we train three segmentation models using 2D patches obtained from different perspectives (axial, coronal, and sagittal views), respectively, and then to obtain the final segmentation results using a voting based fusion strategy. Finally, the experimental results have demonstrated the performance of the proposed method compared with some state-of-the-art approaches. Our experimental results also show that different modalities have different effects on the performance. In addition, we found that the Flair and T1c sequences play an essential role in the training of a segmentation model, and the segmentation model based on Flair, T1c, and T2 modalities obtained similar segmentation results and lower computation time compared to the segmentation model based on all four modalities. Overall, our method provided promising performance on the BRATS 2013, 2015, and 2018 datasets compared to the state-of-the-art approaches. It took about 1.5-3 minutes to segment an entire brain, which is faster than most patch-based methods. However, our model has a decrease in performance when the data differing significantly. Maybe 2D CNN still cannot fully utilize 3D information of MRI data [13]. Our experimental results have shown that the fusion of information from different views can improve the performance of tumor segmentation, but not each single view can achieve good segmentation results. Therefore, our future research is to consider the interaction information of images from more perspectives and effectively integrate them into a 3D CNN for brain tumor segmentation.

## REFERENCES

1. S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, 2013.
2. C. Direko, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Procedia Comput. Sci.*, 2016.
3. M. Goetz, C. Weber, F. Binczyk, J. Polanska, R. Tarnawski, B. Bobek- Billewicz, U. Koethe, J. Kleesiek, B. Stieltjes, and K. H. Maier-Hein, "DALSA: Domain adaptation for supervised learning from sparsely annotated MR images," *IEEE Trans. Med. Imag.*, Jan. 2016.
4. S. D. Olabariaga and A.W. Smeulders, "Interaction in the segmentation of medical images: A survey," *Med. Image Anal.*, 2001.
5. J. Yao, "Image processing in tumor imaging," in *New Techniques in Oncologic Imaging*. London, U.K.: Taylor & Francis, 2006.
6. B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, 2015.
7. M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Med. Image Anal.*, 2004.
8. M. B. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. G. Villemure, and J. P. Thiran, "Atlas-based segmentation of pathological MR brain images using a model of lesion growth," *IEEE Trans. Med. Imag.*, Oct. 2004.
9. B. H. Menze, K. van Leemput, D. Lashkari, M.-A. Weber, N. Ayache, and P. Golland, "A generative model for brain tumor segmentation in multimodal images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Beijing, China: Springer, 2010.
10. Gooya, K. M. Pohl, M. Bilello, L. Cirillo, G. Biros, E. R. Melhem, and C. Davatzikos, "GLISTR: Glioma image segmentation and registration," *IEEE Trans. Med. Imag.*, Oct. 2012.