2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Whole Tumor Segmentation from Brain MR images using Multi-view 2D Convolutional Neural Network

Ritu Lahoti, Sunil Kumar Vengalil, Punith B Venkategowda, Neelam Sinha, Vinod Veera Reddy

Abstract— **In this paper, a study is reported on the popular BraTS dataset for segmentation of brain tumor. The BraTS 2019 dataset is used that comprises four MR modalities along with the ground-truth for 259 high grade glioma (HGG) and 76 low grade glioma (LGG) patient data. We have employed U-Net architecture based 2D convolutional neural network (CNN) for each of the orthogonal planes (sagittal, coronal and axial) and fused their predictions. The objective function is aimed to minimize Dice loss between the binary prediction and its actual labels. Samples having tumor information are considered for each patient data to avoid training on non-informative data. The models are trained on 222 HGG data and tested on 37 HGG data using performance metrics such as sensitivity, specificity, accuracy and Dice score. Test-time augmentation is also performed to improve the segmentation performance. 7-fold cross validation is conducted to analyze the performance on different sets of training and testing data.**

Index Terms*— *Brain tumor segmentation, MRI, Multi-view Convolutional Neural Network

I. Introduction

Early detection and diagnosis are the keys for appropriate treatment of terminal diseases such as cancer. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans of the pathologic organ are the most commonly used diagnostic tools. Often structural neuroimaging such as 3D volumetric brain MRI are performed to perceive the shape, size and location of the tumor in the brain. These medical images are then analyzed by doctors and surgeons to examine the tumor condition by visualizing the 3D data slice by slice and then initiate and monitor the therapy. However, manually tracing and delineating the tumor regions from the data could be tedious for them and also prone to error. Computer aided diagnosis can be a helping hand to this manual system that can also reduce observational oversights.

Biomedical image segmentation is one of the critical fields of study where precision and accuracy matters a lot. Over the last few decades, there has been an immense shift in image segmentation techniques; from traditional image processing techniques such as watershed segmentation, k-means clustering to semi-automated machine learning approach involving feature extraction from ROI for training classifiers to fully automated data-driven deep learning methods such as the popular CNN. Brain tumor segmentation from MR images has gained attention of deep learning researchers to build an automated system to perform the segmentation task efficiently. Among various neural networks studied, CNN has shown reliable segmentation results. 3D CNN exploits volumetric characteristic of brain MR data but at the cost of large memory space and longer training time [1]. To overcome this yet reinstating the 3D information, multi-view 2D CNNs having similar network structure as 3D CNN have been found to be useful [2]. It is computationally efficient and considers information from three orthogonal planes (sagittal, coronal and axial). Predictions from these planes can be fused by averaging, max-pooling or majority voting. Furthermore, a deep neural network works well when trained on large dataset and data augmentation is one way to achieve it. Various augmentation techniques such as data flipping, rotation, translation, shearing, and deformation are used. Also, test-time augmentation is used to obtain better segmentation results.

In this study, we have performed brain tumor segmentation on high quality annotated multi-modal BraTS 2019 dataset of brain glioma [3-7]. U-Net architecture based multi-view 2D CNNs are employed. Segmentation operation is performed along the sagittal, coronal and axial axes, thus giving three trained models. These models predict the segmentation probability maps of the test data which are then fused by averaging. This paper is organized as follows; Methodology (model architecture, multi-view 2D CNN integration unit), experiments and results (dataset used, implementation details and segmentation results), discussion and conclusion.

II. Methodology

In this study, multi-view 2D CNN models are considered to combine results from three orthogonal views. The schematic diagram is shown in Fig.1.

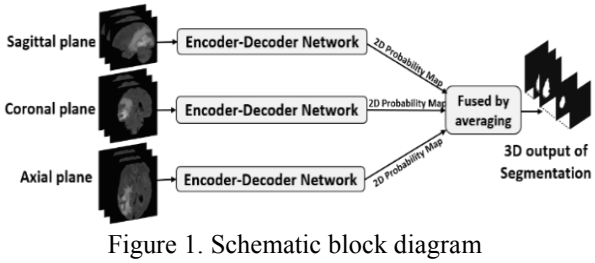

Figure 1. Schematic block diagram

Ritu Lahoti is with International Institute of Information Technology Bangalore, India (email: ritu.lahoti@iiitb.ac.in)

Sunil Kumar Vengalil is with International Institute of Information Technology Bangalore, India (email: sunilkumar.vengalil@iiitb.ac.in)

Punith B. Venkategowda is with International Institute of Information Technology Bangalore, India (email: Punith.Bv@iiitb.ac.in) and Siemens Healthcare Pvt. Ltd, India

Neelam Sinha is with International Institute of Information Technology Bangalore, India (email: Neelam.Sinha@iiitb.ac.in)

Vinod Veera Reddy is with International Institute of Information Technology Bangalore, India (email: vinod.reddy@iiitb.ac.in)

A. Model Architecture

A deep encoder-decoder architecture for CNN model have been employed to perform tumor segmentation, as shown in Fig.2. [8]. Encoder is the contracting path that down-samples the input slices into small representation with high-level features whereas the decoder in the expanding path up-samples the compressed form back to give segmentation results in same resolution as input data. Skip connection combines low-level features with high-level features to minimize loss of fine information and then serve as input to the next layer. Batch normalization is performed during training (batch size of 10) to supervise weights imbalance during training that may cause network instability. This is followed by rectified linear unit (ReLU) activation to enhance the learning ability. The last convolution layer has filter size of 2, one for each class labels i.e., tumor and non-tumor followed by sigmoid activation layer. To control over-fitting, dropout function is used with a rate of 0.2 after the initial convolution layer in the training step.

B. Multi-view 2D CNN Integration unit

Since 3D CNN computation requires considerable memory and 2D CNN along only one direction will restrain details from other two directions, multi-view 2D CNN is used i.e. along the three orthogonal axes (sagittal, coronal and axial). Data along each orientation are fed to three separate 2D CNN based encoder-decoder network. Sigmoid activation function on the last layer of the networks is applied to acquire probability maps of 3D volume for each orientation which are then integrated by averaging to get the final probabilities [9]. Thereupon, thresholding is performed to get the binary labels for all the voxels of the 3D volumetric brain data.

III. Experiments and Results

A. Dataset used

We have used volumetric 3D brain MRI of 259 HGG cases along with their ground-truth tumor segmentation labels from BraTS 2019 training dataset. All the cases have four MRI modalities (T1, T1-contrast enhanced, T2, FLAIR), each with input size of 240x240x155. For axial view, whole slices cropped to 160x160 (160x144 for sagittal and coronal views) are fed as input to the network. They are rigidly aligned and resampled to 1x1x1 mm isotropic resolution and are skull-stripped. The four modalities are treated as input channels. Since the intensity distribution varies across the dataset, we have normalized the images for each channel by subtracting the mean and dividing by the standard deviations. Moreover, only those slices are elected that carry tumor pixels for training to save on computation time and restrain from training unlabeled data.

B. Implementation details

The proposed workflow was developed and implemented using TensorFlow and Keras in python. Three 2D CNN models are trained using dataset provided by BraTS 2019 organizers, one model for each of sagittal, coronal and axial plane. To update the model parameters, adam optimizer is employed with learning rate of 0.0001. The models are trained with batch size of 8 and epoch number is set to 1. Furthermore, the tumor region which is of prime concern often contributes inconsiderably to the overall 3D brain volume. As a result, while training, the model become biased towards the dominant class i.e. healthy tissue, thereby performing poorly. To decrease the loss contribution by the data points (x) already well-classified, Dice loss function has been incorporated as the objective function.

The probability maps from sigmoid layer for each plane have probability value for each voxel of being tumor. These maps are then averaged to get the final segmentation probability map. The voxels having value more than 0.5 are classified as tumor and others as background (non-tumor). The method has been evaluated on HGG data by implementing a 7-fold cross validation with 222 training data and 37 test data. The choice of 7-fold was to have more data for training phase as well as for performance evaluation purpose i.e. testing phase and in this way segmentation performance was analyzed on all the 259 HGG data.

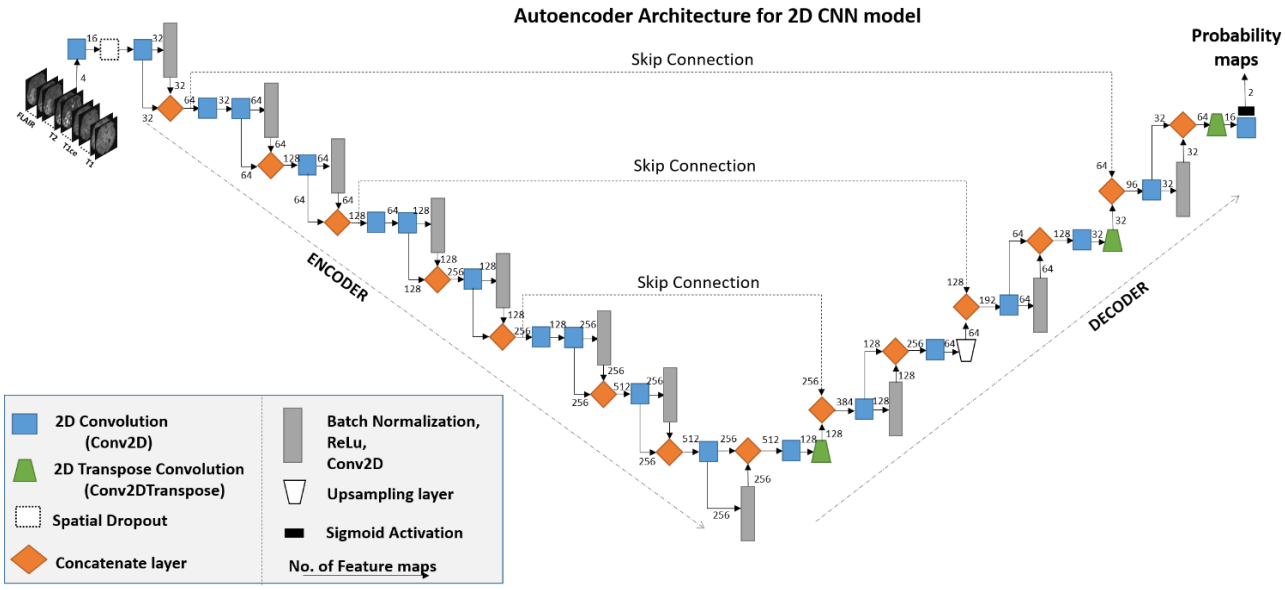

Figure 2. Diagrammatic representation of encoder-decoder network

C. Segmentation Results

We used the corresponding trained models to obtain the segmentation results and the associated evaluation metrics value for each fold trained on HGG dataset. During testing, we incorporated data augmentation by flipping method [9]. The results from the flipped version were averaged with the original results. The sensitivity, specificity, accuracy and Dice score values for 37 test data are illustrated in the boxplots shown in Fig.3. The horizontal line within the boxes shows the mean values of each folds for the test data.

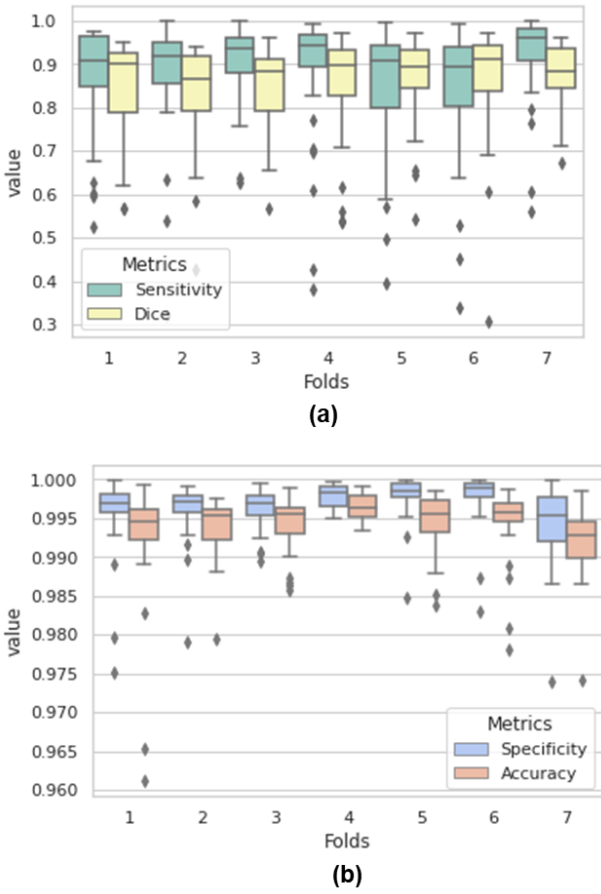

Figure 3. Segmentation evaluation metrics for 7-Fold cross validation experiments (a) Sensitivity, Dice score, (b) Specificity, Accuracy

In Fig.3 (a), the sensitivity boxplots shows outlier among the data with many actual tumor pixels that were classified as non-tumor pixels (false negatives) whereas the outliers in specificity boxplots, Fig.3 (b) shows those data point whose many non-tumor pixels were misclassified as tumor pixels (false positives). The overall tumor segmentation performance for such an imbalanced dataset is reflected in the Dice score boxplots whose outliers represent those data point that were not segmented well (either false positives and/or false negatives). The average performance across the different folds show sensitivity of 87.87%, specificity of 99.64% accuracy of 99.4% and Dice score of 0.8573.

The segmentation results for some of the test data are displayed in Fig.4 (b). The samples displayed shows how well

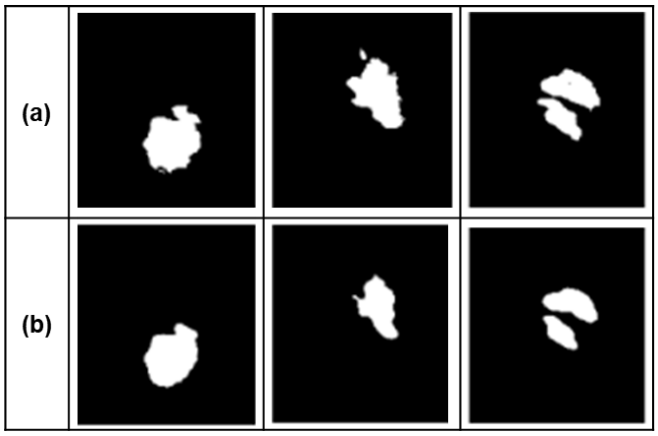

Figure 4. (a) Ground Truth, (b) Segmentation Results

the tumor is segmented when compared to ground-truth with few false negatives.

We have compared our method with selected existing models for whole tumor segmentation on 125 validation data of BraTS 2019 and their mean values are illustrated in Table. 1. It can be observed that 3D approaches outperforms 2D approaches. However, among the 2D approaches, our model shows decent performance.

TABLE 1. Dice score, sensitivity, specificity of some existing methods and our method are present in the table (FCNN: Fully Convolutional Neural Network, DC: Dense Channels, Res: Resolution, MSCN: Multi-Step Cascaded Network, N/w: Network, MV: Multi-View, NA: Not Available)

Method	**N/w Type**	**Dice Score**	**Sensitivity**	**Hausdorff_95**
		Range:0 -1 (greater the better)	Range:0-1 (greater the better)	Range:>=0 (smaller the better)
FCNN [10]	2D	0.73	0.67	12.8
DC U-Net [11]	2D	0.885	NA	19.74
Our Model	2D MV	0.817	0.751	11.23
Multi-Res [12]	3D	0.86	0.85	8.42
MSCN [13]	3D MV	0.886	0.921	6.23

IV. DISCUSSION AND CONCLUSION

In this study, we evaluated the multi-view tumor segmentation on BraTS dataset. For this, we performed a 7-fold experiment to segment high grade glioma (259 data) from the BraTS dataset. Due to tumor diversity, we cannot limit to single view analysis as it won't be generalizable and hence, we utilized the three orthogonal planes to exploit the 3D contextual features which improves the segmentation efficiency compared to single-view approach. We analyzed the sensitivity, specificity, accuracy and Dice score of segmentation results. With test-time augmentation, the results

further improved by 0.05-0.2%. In future study, we plan to further investigate and undertake more exhaustive evaluation of multi-view approach with variation in fusing strategy and also improve the performance on outlier data.

ACKNOWLEDGMENT

No funding was received for conducting this study. The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Casamitjana, A., et al., 3D Convolutional Neural Networks for Brain Tumor Segmentation: A Comparison of Multi-resolution Architectures. 2016. 150-161.

[2] Wang, G., et al., Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation. Frontiers in Computational Neuroscience, 2019. 13(56).

[3] B. H. Menze, A.J., S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans Med Imaging, 2015.

[4] Bakas, S., et al., Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. ArXiv, 2018. abs/1811.02629.

[5] Bakas S, A.H., Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C., Segmentation Labels for the Pre-operative Scans of the TCGA-GBM collection. The Cancer Imaging Archive, 2017.

[6] Bakas S, A.H., Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C., Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection. The Cancer Imaging Archive, 2017.

[7] S. Bakas, H.A., A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Nature Scientific Data, 2017.

[8] Myronenko, A., 3D MRI brain tumor segmentation using autoencoder regularization. BrainLes@MICCAI 2018. 2018.

[9] Nalepa, J., M. Marcinkiewicz, and M. Kawulok, Data Augmentation for Brain-Tumor Segmentation: A Review. Frontiers in computational neuroscience, 2019. 13: p. 83-83.

[10] Rupal Agravat MSR, Brain Tumor Segmentation and Survival Prediction. BrainLes 2019; 2019; China: Springer.

[11] Brain Tumor Segmentation Using Dense Channels 2D U-net and Multiple Feature Extraction Network. Cham: Springer, 19 May 2020.

[12] Mehdi Amian MS, Multi-Resolution 3D CNN for MRI Brain Tumor Segmentation and Survival Prediction. Lecture Notes in Computer Science (LNCS) BraTS proceedings; 2019 19 May 2020; Cham: Springer.

[13] Wang XLL, Multi-step Cascaded Networks for Brain Tumor Segmentation2019 19 May 2020: Springer