Brain Tumor Segmentation using Cascaded Deep Convolutional Neural Network

Saddam Hussain, Syed Muhammad Anwar, Muhammad Majid

Abstract—**Gliomas are the most common and threatening brain tumors with little to no survival rate. Accurate detection of such tumors is crucial for survival of the subject. Naturally, tumors have irregular shape and can be spatially located anywhere in the brain, which makes it a challenging task to segment them accurately enough for clinical purposes. In this paper, an automated segmentation algorithm for brain tumor using deep convolutional neural networks (DCNN) is proposed. Deep networks tend to have a lot of parameters thus over-fitting is almost always an issue especially when data are sparse. Max-out and drop-out layers are used to reduce the chances of over-fitting since data are scant. Patch based training method is used for the model where two types of patches sized 37×37 and 19×19 with same center pixel are selected. The proposed algorithm includes preprocessing in which images are normalized and bias field corrected, and post processing where small false positives are removed using morphological operators. BRATS 2013 dataset is used for evaluation of the proposed method, where it outperforms state-of-the-art methods with similar settings in key performance indicators.**

Keywords—brain tumor, segmentation, deep learning, convolutional neural networks.

I. INTRODUCTION

Brain tumor is a disease that can lead to disability and in severe conditions can be fatal [1]. A study has evaluated that 25 people in 100,000 adults have tumor where almost 33% of them have been categorized as critical [2]. Mostly, brain tumors emerge in the cerebrum area, but there is another type of tumor known as metastatic brain tumor that often emerges in different body parts and then spreads to the brain [3]. Gliomas are the most dominant type of cerebral tumors in adults [3]. Glial cells degeneration is at the origin of gliomas tumors. World health organization (WHO) has divided tumors into four grades depending on their severity [4]. Grade 1 and grade 2 are mildly dangerous and are also known as low grade (LG) slowly advancing tumors, whereas, grade 3 and grade 4 are the high grade (HG) tumors. HG tumors are threatening and most likely lead to patient's death, having a maximum life expectancy of two years. On the other hand, LG tumors are associated to longer life expectancy that can span over many years depending upon their spreading ratio and grade.

Magnetic resonance imaging (MRI) is the most common method for cerebral tumor detection [1]. MRI is a widely used non-intrusive imaging modality, which gives sensitive tissue contrast. MRI offers ways to generally conform tissue contrast by normalization, which makes it an exceptionally adaptable device for imaging distinctive structures of interest in human brain like tumors. Due to the structure and uncertain locality of brain tumors, single modality of MRI is not adequate to detect irregular shaped tumors in all brain areas. Recently, distinctive MRI sequences have been utilized for finding and localizing tumor areas [5]. These sequences include T1-weighted MRI (T1), T1-weighted MRI with contrast improvement (T1c), T2-weighted MRI (T2) and T2-weighted MRI with fluid attenuated inversion recovery (T2flair) [6].

Due to challenging nature of tumor segmentation task and its importance in clinical enviroment, many algorithms have been proposed for automated, semi automated and manual segmentation. A survey on these algorithms can be found in [7, 8]. Most of these algorithms were tested on small datasets with relatively different metrics as finding a standard dataset of medical imaging for algorithm testing is a challenging task in itself. Thus making most of the prior techniques data dependent and less useful. Due to these issues, comparison between techniques and with manually segmented clinical data are not straightforward. MICCAI BRATS launched the brain tumor segmentation challenge in 2012 and 2013 providing standard MRI data that has been a benchmark for brain tumor segmentation [9].

Recently, deep learning based methods have provided good results in medical image analysis and retrieval [10, 11]. Pixel based predictions are a new trend in deep learning methods [12]. A deep convolutional neural network (DCNN) with a tree like structure has been used to boost output segmentation [13]. A convolutional neural network (CNN) has been proposed for both benign and malignant breast tumors classification, showing considerably high accuracy [14]. In [15], a CNN architecture is proposed that has used small kernels and ample pre-processing. A deep three dimensional CNN has been used for brain lesion segmentation employing 3D random field as fully connected layer, which is helpful in removing false positives [16]. In [17] another 3D CNN architecture has been proposed with considerably good accuracy for tumor segmentation.

In this study, a patch based approach along with convolutional neural network is utilized to classify the center of each patch. Input cascaded approach for CNN, utilizing both global and local features is used. This is important for

Saddam Hussain, is with the Department of Software Engineering, University of Engineering &Technology, Taxila, 47080 Pakistan (e-mail: saddam4440@gmail.com).

Syed Muhammad Anwar is with the Department of Software Engineering, University of Engineering & Technology, Taxila, 47080 Pakistan (e-mail: s.anwar@uettaxila.edu.pk).

Muhammad Majid is with the Department of Computer Engineering, University of Engineering & Technology, Taxila, 47080 Pakistan (e-mail: m.majid@uettaxila.edu.pk).

978-1-5090-2809-2/17/$31.00 ©2017 IEEE

problems like tumor segmentation where tumor classes are closely related and have very similar features. Different classes are diffused together without clear boundaries, making it hard to learn features of a class. Smaller kernels have been used to learn features on a smaller scale so that the network can differentiate between two tumor classes in close vicinity of each other. The proposed system achieves improved results in all major brain tumor classification parameters. This can benefit the clinical decision support systems in segmenting brain tumors. The following sections present the proposed methodology, experimental setup, results achieved and conclusion.

II. Proposed Methodology

The proposed methodology, shown in Figure 1, is applied on multimodal MRI sequences and exploits the inherent pattern recognition capability of CNN to classify tumor pixels. A patch based approach is used for pixel classification, where pre-processed images are passed through a CNN and post processed to obtain the segmented image highlighting the tumor area. These steps are discussed in detail in the following subsections.

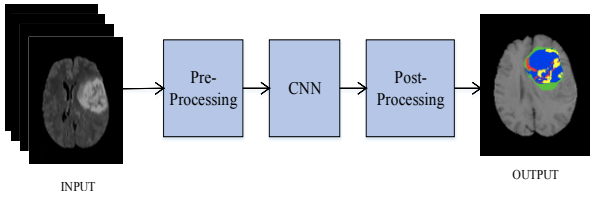

Figure 1. Block diagram of the proposed method.

A. Preprocessing

The MR images when extracted from volumetric data have artifacts due to different acquisition techniques and systems [18]. Especially in T1 and T1c modality, same type of tissues have different intensities across the dataset. N4ITK bias field correction is applied using 3D slicer toolkit [19, 20] to T1 and T1c modalities. Image normalization is performed to ensure zero mean and unit variance. Finally, patches are normalized with respect to mean and variance. Since concatenated architecture is used for the neural network, two types of patches are extracted: one having 37×37 pixels and the other having 19×19 pixels co-centric with the 37×37 patch.

B. Convolutional Neural Network (CNN)

CNN has an advantage over other classifiers as kernels used in convolution layers have same weights for all inputs, which detect same features making them translation invariant [21]. Usually, a non-linear activation function is used to convert features into class probabilities.

A novel cascaded architecture is implemented by concatenating output of first half with input of second half of the network as shown in Figure 2. The proposed architecture takes patches of multiple modalities as input and predicts the class of center pixel in respective patches. The BRATS dataset [9] lacks resolution in third dimension, therefore axial view is used to extract 2D patches. In the first convolution layer, input is the patches extracted from original MR images corresponding to different modaltities used. While, succeeding layers take feature maps produced by preceeding layer as input. A network containing six convolution layers is implemented that learns feature maps with different kernel sizes. For non-linearity, rectified linear units (RLUs) activation is used since it produces best results verified through grid search.

Three max-pooling layers are used to reduce input dimensionality going into the next layers. Max pooling layer summarizes the data in a small rectangle by selecting the max value and discarding the rest. In this way a summarized important structure is passed to next convolution layer discarding the irrelevant information. Pooling layers also have beneficial effects like invariance to position and lightning conditions [22]. The complexity is further reduced by using maxout layer, which reduces the dimensions in third axis, thereby, reducing the number of feature maps. It is used after convolution layer and selects the maximum of two adjacent feature maps, thus reducing the number of maps produced by convolution layer to half. It is observed that this improves performance by a small margin.

The input for second half of the network is extracted from the center of the input of first half. The concatenated input is fed into the second part of the network and the output layer i.e. softmax activation predicts class probabilities, which are accounted for in the loss function.

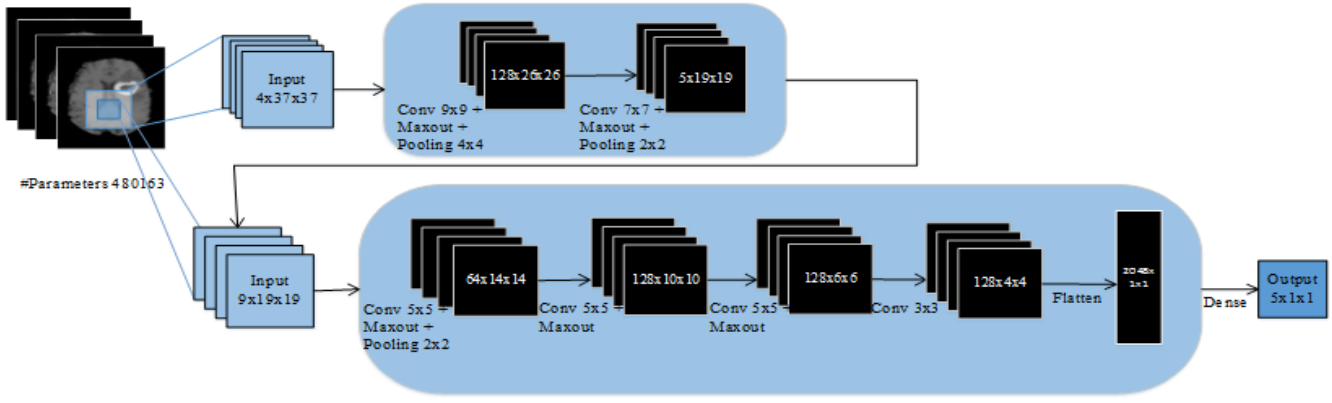

Figure 2. The proposed CNN architecture with input concatenation

C. Post Processing

In post processing, morphological techniques are used to remove small false positives around the corners of MRI scan, which emerge due to high intensity around skull portion.

III. Experimental Setup

A. Dataset

The proposed methodology is validated on BRATS 2013 dataset [9, 23]. It contains four modalities i.e., T1, T1c, T2 and T2flair. The dataset comprises of 30 training images (20 HG and 10 LG). The data are rather sparse and preprocessing steps like skull stripping have been performed to improve the data representation. In 2013 dataset, two more data subsets are provided i.e. leaderboard and challenge data. These two subsets comprise of 65 MR images. Manual segmentation is available for training data only. Five output labels are provided namely necrosis, edema, non-enhancing tumor, enhancing tumor and normal tissues. For evaluation purposes three classes are used namely, enhancing tumor, core tumor (necrosis, non-enhancing tumor and enhancing tumor) and complete tumor (all tumor classes).

B. Implementation Details

The algorithm is implemented in Keras library in python. It is high-level library used for implementing neural networks and can run over either Theano or TensorFlow framework. It supports both GPU and CPU processing. Hyper-parameters are tuned using grid search and the parameters on which model performed best on validation data are selected. Parameters such as learning rate and momentum are varied during training. Momentum is initially set to 0.5 and is gradually increased to 0.9. Learning rate on the other hand is initially set to 0.003 and then gradually decreased to 0.3×10^{-5}. The model is trained on one hundred thousand patches. A high number of network parameters and low number of patches increased the chances of over-fitting. Hence, a dropout value of 0.5 is used in the network to avoid over-fitting.

Segmenting brain tumor is an unbalanced classification problem where, most of the pixels are of healthy tissues. Training a network with true distribution of labels, causes it to be overwhelmed by healthy patches. To overcome this problem, two phase training for the network is used. In the first phase, the network is trained on one hundred thousand patches for 6 epochs such that all classes have equal representation in training data. During the second phase, model is trained on twenty thousand patches for 2 epochs where patches are extracted in accordance with their actual class distribution. All layers except the output layer are fixed and only the output layer is tuned during the second phase of training.

C. Evaluation Parameters

The experimental results are evaluated based on three metrics, namely dice similarity coefficient (DSC), sensitivity and specificity. Dice score is calculated by overlapping predicted labels with actual labels and the intersection of two contributors determine the dice score. Dice score is calculated for three categories i.e. whole tumor, enhancing tumor, and core tumor and is given by,

$$DSC = 2 \times (|L \cap P|) / |L| + |P|, \quad (1)$$

where L and P stand for actual labels for tumor region and predicted tumor regions respectively. Sensitivity is the measure of tumor pixels that have been correctly classified and is given by,

$$Sensitivity = |P| \cap |L| / |L|, \quad (2)$$

Specificity is the measure of healthy pixels that have been classified correctly and is given as,

$$Specificity = |P_o| \cap |L_o| / |L_o|, \quad (3)$$

where P_o and L_o stand for predicted and actual non-tumor or healthy portion respectively.

IV. Results And Discussion

Experiments have been performed on BRATS 2013 dataset that has two types of tumors, HG and LG glioma, divided in to four tumor classes. There are 30 volumetric images in 2013 dataset containing slices varying in range of 150 to 220. The dataset is divided randomly into training and testing sets with 80:20 ratios. The dataset also contains synthetic data with low variance in intensity values of a similar class that are comparatively easy to classify. Therefore, only real patient data are used for evaluating the model. Evaluation metrics are determined for three tumor regions namely a) the complete tumor area (all four tumor labels), b) the core tumor area (labels 1, 3 and 4), and c) the enhancing tumor region (label 4). For evaluating the effectiveness of the proposed model, a comparative analysis is presented in Table 1. Evaluation results are compared to classification results presented in [24, 25]. A probabilistic model that combines sparse representation and Markov random field to classify tumor pixels has been proposed [24]. In [25], random decision forests are trained on image features to classify voxels. It is evident from Table 1 that proposed model outperforms state-of-the-art methods in terms of dice score, sensitivity, and specificity. Some of the segmentation results generated using the trained neural networks are shown in Figure 3. The proposed algorithm performs well in specifying tumor region as is evident from the lack of false positives in detections. It also detects enhancing tumor better than most state-of-the-art techniques and gives comparable results on other metrics.

TABLE I. Segmentation Results on MICCAI 2013 BRATS Training Data.

Method	Dice			Sensitivity			Specificity		
	Complete	*Core*	*Enhancing*	*Complete*	*Core*	*Enhancing*	*Complete*	*Core*	*Enhancing*
SR [24]	0.78	0.52	0.52	0.85	0.54	0.58	-	-	-
RDF [25]	0.79	0.62	0.57	0.80	0.57	0.57	0.82	0.80	0.60
Proposed	0.80	0.67	0.85	0.82	0.63	0.83	0.85	0.82	0.88

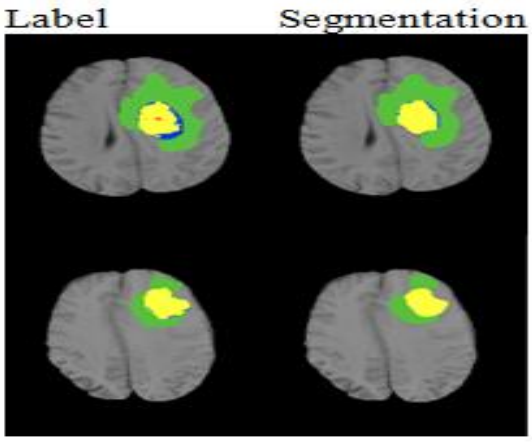

Figure 3. Segmentation results for images from BRATS 2013 with label (first column) and segmented region (second column). Red, Green, Blue and Yellow show Label 1, 2, 3, and 4 respectively.

It is observed that the trained model faces difficulty, predicting minority classes. But by increasing training data, this problem can be dealt with. Two variations of the proposed architecture are also tested where number of features in fully connected layer is varied. It has been observed that too many features in fully connected layers lead to over-fitting and if features are reduced enormously, the model does not learn significantly leading to under-fitting. It is also observed that fully connected layers are time consuming compared to convolutions and thus a trade-off between segmentation time and accuracy is achieved with 2048 features in fully connected layer.

V. Conclusion

Brain tumor segmentation has a very important role in diagnostic procedures. With accurate segmentation, clinical diagnostic not only becomes easy, but also the chances of subject's survival increase tremendously. In this paper, a CNN architecture for brain tumor segmentation is presented. This algorithm incorporates both global and local features since context is important when it comes to tumor segmentation task. The use of max-pooling, max-out and drop-out complement the learning process, improving training and testing speed by reducing features in fully connected layer as well as reducing number of parameters, which in turn reduce the chances of over-fitting. Evaluation results show that the proposed network architecture is promising and performs particularly well in detecting enhancing tumor as well as specifying tumor to actual tumor region only.

References

[1] L. M. DeAngelis, "Brain tumors," *New England Journal of Medicine,* vol. 344, pp. 114-123, 2001.

[2] B. A. Kohler, E. Ward, B. J. McCarthy, M. J. Schymura, L. A. Ries, C. Eheman*, et al.*, "Annual report to the nation on the status of cancer, 1975–2007, featuring tumors of the brain and other nervous system," *Journal of the National Cancer Institute,* 2011.

[3] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in medicine and biology,* vol. 58, p. R97, 2013.

[4] A. B. T. Assocation, *About Brain Tumors: A Primer for Patients and Caregivers*: American Brain Tumor Association, 2013.

[5] A. Drevelegas and N. Papanikolaou, "Imaging modalities in brain tumors," in *Imaging of Brain Tumors with Histological Correlations*, ed: Springer, 2011, pp. 13-33.

[6] S. Abbasi and F. T. Pour, "A hybrid approach for detection of brain tumor in MRI images," in *Biomedical Engineering (ICBME), 2014 21th Iranian Conference on*, 2014, pp. 269-274.

[7] M. M. Sharma, "Brain Tumor Segmentation Techniques: A Survey," *Brain,* vol. 4, 2016.

[8] A. Ravi and S. Sreejith, "A Review on Brain Tumour Detection Using Image Segmentation," *International Journal of Emerging Technology and Advanced Engineering,* vol. 5, pp. 60-64, 2015.

[9] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby*, et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE transactions on medical imaging,* vol. 34, pp. 1993-2024, 2015.

[10] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annual Review of Biomedical Engineering,* vol. 19, 2017.

[11] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical Image Retrieval using Deep Convolutional Neural Network," *arXiv preprint arXiv:1703.08472,* 2017.

[12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.

[13] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE transactions on pattern analysis and machine intelligence,* vol. 35, pp. 1915-1929, 2013.

[14] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation," *Expert Systems with Applications,* vol. 42, pp. 990-1002, 2015.

[15] 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, pp. 1240-1251, 2016.

[16] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon*, et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis,* vol. 36, pp. 61-78, 2017.

[17] D. Yi, M. Zhou, Z. Chen, and O. Gevaert, "3-D Convolutional Neural Networks for Glioblastoma Segmentation," *arXiv preprint arXiv:1611.04534,* 2016.

[18] G. Collewet, M. Strzelecki, and F. Mariette, "Influence of MRI acquisition protocols and image intensity normalization methods on texture classification," *Magnetic Resonance Imaging,* vol. 22, pp. 81-91, 2004.

[19] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich*, et al.*, "N4ITK: improved N3 bias correction," *IEEE transactions on medical imaging,* vol. 29, pp. 1310-1320, 2010.

[20] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol*, et al.*, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic resonance imaging,* vol. 30, pp. 1323-1341, 2012.

[21] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for LVCSR," in *Acoustics, speech and signal processing (ICASSP), 2013 IEEE international conference on*, 2013, pp. 8614-8618.

[22] N. Carlevaris-Bianco and R. M. Eustice, "Learning visual feature descriptors for dynamic lighting conditions," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, 2014, pp. 2769-2776.

[23] M. Kistler, S. Bonaretti, M. Pfahrer, R. Niklaus, and P. Büchler, "The virtual skeleton database: an open access repository for biomedical research and collaboration," *Journal of medical Internet research,* vol. 15, p. e245, 2013.

[24] Y. Li, F. Jia, and J. Qin, "Brain tumor segmentation from multimodal magnetic resonance images via sparse representation," *Artificial intelligence in medicine,* vol. 73, pp. 1-13, 2016.

[25] J. Festa, S. Pereira, J. A. Mariz, N. Sousa, and C. A. Silva, "Automatic brain tumor segmentation of multi-sequence MR images using random decision forests," *Proceedings of NCI-MICCAI BRATS,* vol. 1, pp. 23-26, 2013.