Stochastic Gradient Descent Optimizer for Segmentation of Brain Tumor using Mask R-CNN

E Murali[1], Dr. K Meena[2]
[1]Research Scholar, [2]Associate Professor, Dept. of CSE
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai.

Abstract:

Automatic segmentation of microscopy images is an important task in medical image processing and analysis. Tumor detection is an important task. Mask-RCNN is a recently proposed state-of-the-art algorithm for object detection, object localization, and object instance segmentation of MRI images. This paper explores segmenting brain tumor using two methods. The first method uses a cascade of a WNet and a UNet and the second method uses a Mask R-CNN framework to classify tumors in the brain. In this paper we demonstrate that Mask-RCNN can be used to perform highly effective and efficient automatic segmentations of a wide range of microscopy images.

Keywords — Brian Tumor, Mask R-CNN, WNet, UNet, Stochastic gradient descent.

I. INTRODUCTION

In the last few years, algorithms based on convolutional neural networks (CNNs) have led to dramatic advances in the state of the art for fundamental problems in computer vision, such as object detection, object localization, semantic segmentation, and object instance segmentation. [1], [2], [3], [4]. This has led to increased interest in the applicability of convolutional neural network-based methods for problems in medical image analysis. Recent work has shown promising results on tasks as diverse as automated diagnosis diabetic retinopathy, automatic diagnosis of melanoma, precise measurement of a patient's cardiovascular ejection fraction, segmentation of liver and tumor 3D volumes, segmentation of mammogram images, and 3D knee cartilage segmentation. This has led to increased interest in the applicability of convolutional neural network-based methods for problems in medical image analysis.

Magnetic resonance (MR) [1] image segmentation of a brain is a very important and exigent task that is needed for the purpose of diagnosing brain tumors and other neurological diseases. Brain tumors have different characteristics such as size, shape, location, and image intensities. They may deform neighboring structures and if there is edema with the tumor, intensity properties of the nearby region change. An automatic segmentation of the brain MRI image is necessary because manual segmentation requires more time and can be subjected to errors. A fast reliable technique is necessary to detect the brain tumor because treatment planning is the key method to improve the survival period of oncological patients. This paper presents a reliable detection method based on CNN that reduces operators and errors. The Convolutional Neural Network (CNN) is used in convolving a signal or an image with kernels to obtain feature maps. The image processing techniques such as image conversion, feature extraction and histogram equalization have been developed for extraction of the tumor in the MRI images of the cancer affected patients. A suitable Fuzzy Classifier is developed to recognize healthier tissue from cancer tissue. The whole system is divided into two phases: firstly learning/Training Phase and secondly Recognition/Testing Phase. The detection of tumor takes place in main three main stages: (1) pre-processing (2) classification by CNN and (3) post-processing. The aim of the project is to detect and extract the of tissue abnormalities by using the biochemical features. The specificity and the sensitivity of the method are evaluated and accuracy is determined.

This paper will explore two different methods of multimodal brain tumor segmentation on the Brain Tumor Image Segmentation Benchmark 2018 dataset. The first method used was Mark R-CNN, an instance segmentation framework, which identifies object at a pixel level [4]. The input to this framework included two datasets of images. One was the slices of the brain images, with segmentations of the whole tumor, tumor core, and the enhanced tumor core. The second was their corresponding masks identifying only the enhanced tumor core. The output was segmented image.

The second method was implementing UNet in conjunction with the first network (WNet) used by the model created by Wang et al., 2017 [5]. The input to this model was a MRI scan of a whole brain with a tumor in it. We then used the WNet which outputted the bounding box of the whole tumor. The image of the whole tumor was then fed into a UNet to output the segmented brain tumor.

II. RELATED WORK

We began research by reading two sources provided by the BraTs competition: The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) by Menze et.al., 2014 [6] and Advancing the Cancer Genome Atlas gliomas MRI collections with expert segmentation labels and radiomic features by Bakas et al.,2017 [7]. These provide background knowledge.

The methodology we pursed was based mainly of four papers. Two of them discuss the use of UNets for image segmentation: U-Net: Convolution Netwrok for Biomedical Image Segmentation by Ronneberger et. Al., 2015 [8] and Radio frequency interface mitigation using deep convolutional neural network by Akeret et al., 2017. [9] The former was very useful for gaining intuition about various components of the UNet architecture.

Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Network by Wang et al., 2017 [4] explains the CNN implementation that we cloned as a basefor our model. This paper included a discussion which was particularly illustrative for understanding some advantages of using a cascade of fully convolutional neural networks given the hierarchical nature of the tumor sub-regions that we would like to label for segmentation. In order to learn more about the possible advantages and disadvantages of using Mask R-CNN for image segmentation, we consulted Mask R-CNN by He et al. 2018.

III. METHODS

i) Mask R-CNN

Mask R-CNN is a leading object segmentation framework, which follow from Faster R-CNN, an object detection framework. In the first stage of Mask R-CNN, images are scanned and proposals, areas likely to contain an object, are generated. In the second stage, proposals are classified and bounding boxes and masks are generated. The backbone of this framework is a CNN with ResNet101, where early layers detect low level features and later layers detect higher level features. Mask R-CNN utilizes Feature Pyramid Network (FPN) to improve upon the standard feature extraction pyramid by introducing an additional pyramid which takes high level features and feeds them to the lower layers, allowing all levels to have access to both higher and lower level features.

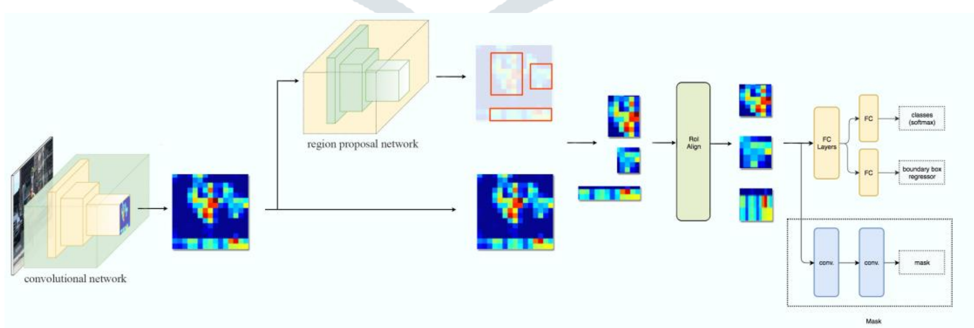

Fig1: Image Segmentation with Mask R-CNN

Mask R-CNN accomplishes object detection using Region Proposal Network (RPN), which partitions the image into anchors, upon which sliding windows traverse the image and find areas containing the object of interest. This particular framework uses a large amount of anchors which makes training slower but more thorough [15]. The output from the RPN are the bounding boxes. Next, the regions of interest (ROIs) are classified and the bounding boxes are further defined. Finally, Mask R-CNN uses the regions from the classified

ROIs and generates masks for them. The masks are low resolution and are represented by floating numbers so then can hold more details than binary masks.

Mask R-CNN uses a multi-task loss function given by $L = L_{cls} + L_{box} + L_{mask}$ where L_{cls} and L_{box} are same as in Faster R-CNN and L_{mask} is given by

$$\mathcal{L}_{mask} = -\frac{1}{m^2} \sum_{1 \leq i,j \leq m} [y_{ij} log \hat{y}^k_{ij} + (1 - y_{ij}) log(1 - \hat{y}^k_{ij})]$$

ii) WNet and UNet

UNet is a standard architecture for classifying to segment areas of an image by class. UNet is especially gaining popularity when dealing with medical images. As seen by the model below, we decided to combine the first network, WNet, of the model created by Wang et al., 2017 [4] and UNet for our second brain tumor classification method [16]. The output of the WNet was a bounding box of the tumor. This was then fed in as the input of the UNet. We choose to do this instead of feeding in the whole brain scan as the input as we hoped this would increase the accuracy of the UNet predictions.

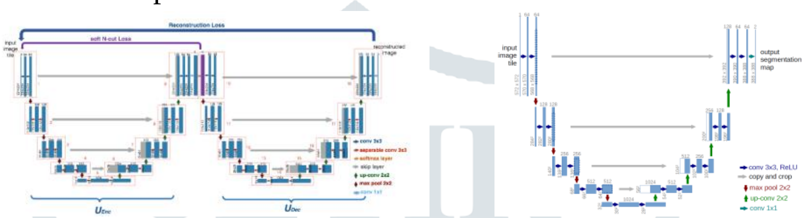

Fig2: WNet and UNet Model

The first network of this model acts like a normal CNN, taking in the image of the whole brain and outputing a bounding box for the whole tumor. As seen in the figure above, this netwrok uses 10 residual blocks with dilated convolution, residual connections and multi scale fusion. The dilated convolutions use batch normalization and PReLu. This reduces the impact of the earlier layers within the network and also helps with regularization. The residual connections help smooth propagation. The multi-scale fusion allows us to represent both low and high level features.

The second network of our model is the UNet. The whole network has 23 layers. The first half of the UNet is a contracting, downsampling path, meaning it acts like a regular CNN. However, the second half of the network is an expansive, upsampling path. Both halves use convolutions and ReLu. The first half also has max pooling and the second half has up-convolutions. Having a network such as this allows us to capture the context of the image while simultaneously allowing the localization to be more precise.

For both of these netwroks, a softmax with a cross entropy loss was used:

$$p_k(x) = \frac{e^{a_k(x)}}{\sum_{k'=1}^{K} e^{a_{k'}(x}} + \sum_{x \in \Omega} w(x) log(p_{l(x)}(x))$$

During training for both network, the Stochastic gradient descent optimizer was used.

$\theta = \theta - \eta \cdot \nabla J(\theta; x(i); y(i))$, where {x(i) ,y(i)} are the training examples

IV. RESULTS AND DISSCUSSION

In order to use the 3D images given the form of .nii files for the Mask R-CNN, they had to be converted to .png files. To work with the framework, each image frame had to have an accomopanying binary mask segmenting the enhanced tumor core in a COCO-style annotation. This was done based on the ground truth segmentation images the dataset provides.

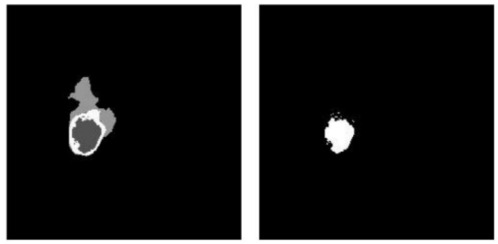

Fig3: Binary mask segmenting the enhanced tumor core

High Variance parameter updates for each training example cause the Loss function to fluctuate heavily due to which we might not get the minimum value of parameter which gives us least Loss value.

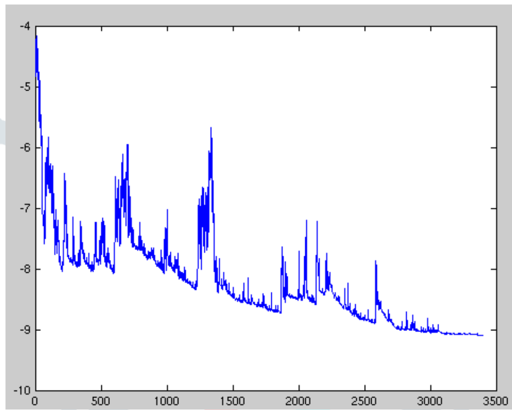

Fig4: Loss value on validation data shows overall decrease in loss

For the UNet, 1 was passed in as the number of channels because we were using gray scale and 2 was used as the number of classes since we were predicting the enhanced tumor core versus no enhanced tumor core [18] . The learning rate used was 0.2 and the batch size was 1. For this method we did not perform any hyperparameter tuning because the model we built off already had tuned hyperparameters. The metric that was used for this model was accuracy.

V. CONCLUSION

In this paper we demonstrate that the Mask-RCNN model, while primarily de-signed with object detection, object localization, and instance segmentation of MRI images in mind, can be used to produce high quality results for the challenging task of segmentation. There are several similar tasks in medical image analysis for which it is likely that a Mask-RCNN based model could easily be adapted to improve performance without extensive modification or customization. Examples of this include the task of segmentation of the left ventricle of the brain, where accurate segmentations can be used to estimate a cardiac patient's ejection fraction and improve their outcomes, or liver and tumor segmentation as described in [5]. Future work will explore the efficacy and performance of Mask-RCNN-based models for a range of such tasks.

VI. REFERENCE

[1] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classi_cation with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (USA, 2012), NIPS'12, Curran Associates Inc., pp. 1097-1105.

[2] Simonyan, K., and Zisserman, A. Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014).

[3] He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. CoRR abs/1512.03385 (2015).

[4] Zagoruyko, S., Lerer, A., Lin, T., Pinheiro, P. H. O., Gross, S., Chintala, S., and Doll_ar, P. A multipath network for object detection. CoRR abs/1604.02135 (2016).

[5] Wang, Guotai, Li, Wenqi, Ourselin, Sebastien & Vercauteren, Tom (2017) Automatic Brain Tumor Segmentation using Cascaded Anisotropic Convolutional Neural Networks, In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Pages 179-190, Springer, 2018.

[6] He, Kaiming & Gkioxari, Georgia & Dollar, Piotr & Girshick, Ross (2017) Mask R-CNN. International Conference on Computer Vision (ICCV) 2017

[7] Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, Freymann JB, Farahani K & Davatzikos C (2017), Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features Nature Scientific Data, 2017.

[8] C_ ic_ek, • O., Abdulkadir, A., S. Lienkamp, S., Brox, T., and Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. 10 2016. 2

[9] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. Dermatologist-level classi_cation of skin cancer with deep neural networks. Nature 542 (2017), 115 - 118.

[10].Sergio Pereira, Adriano Pinto, Victor Alves, and Carlos A. Silva " Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images "IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016

[11]. Rasel Ahmmed, Md. Foisal Hossain "Tumor Detection in Brain MRI Image Using Template based K-means and Fuzzy C-means Clustering Algorithm" 2016 International Conference on Computer Communication and Informatics (ICCCI -2016), Jan. 07 –09, 2016, Coimbatore, INDIA

[12]. S.C. Turaga, J.F. Murray, V. Jain, F. Roth, M. Helmstaedter, K. Briggman,W. Denk, and H.S. Seung. Convolutional networks can learn to generate affinity graphs for image segmentation. Neural Computation, 22(2):511–538, 2010.

[13] C.-H. Lee et al., "Segmenting brain tumors using pseudo-conditional random fields," in Medical Image Computing and Comput.-Assisted Intervention-MICCAI 2008. New York: Springer, 2008, pp. 359–366.

[14] R. Meier et al., "A hybrid model for multimodal brain tumor segmentation," in Proc. NCI-MICCAI BRATS, 2013, pp. 31–37.

[15] E. Murali, K Meena, "A Hybrid Approach to the Classification of Brain Tumors from MRI Images using Fast Bounding Box Algorithm" in IJSST, vol. 19, no.4, pp. 3.1-3.8, Aug, 2018.

[16] E. Murali, K Meena, "A Phenomenological Survey on Various Types of Brain Diseases Using Soft Computing Techniques" in IJCIET, vol. 8, issue 9, pp. 1209-1220, Sep, 2017.

[17] R. Meier et al., "Appearance-and context-sensitive features for brain tumor segmentation," in MICCAI Brain Tumor Segmentation Challenge (BraTS), 2014, pp. 20–26.

[18] D. Zikic et al., "Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR," in Medical Image Computing and Comput.-Assisted Intervention-MICCAI 2012. NewYork: Springer, 2012, pp. 369–376.

[19] S. Bauer et al., "Segmentation of brain tumor images based on integrated hierarchical classification and regularization," Proc. MICCAIBRATS, pp. 10–13, 2012.

[20] S. Reza and K. Iftekharuddin, "Multi-fractal texture features for brain tumor and edema segmentation," SPIE Med. Imag. Int. Soc. Opt. Photon., pp. 903503–903503, 2014.

[21] N. Tustison et al., "Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR," Neuroinformatics, vol. 13, no. 2, pp. 209–225, 2015.