



3D-UNet Architecture Using Separable 2D Convolutions

Ashlin k Benny¹, Aksa Mariam George ², Akhila S³, Annu Saji⁴, Dr. Aswathy S U

¹Mangalam college of Engineering, India, ashlink98@gmail.com

²Mangalam college of Engineering, India, aksamariyam2020@gmail.com

³Mangalam college of Engineering, India, akhilasanthosh33@gmail.com

⁴Mangalam college of Engineering, India, annusaji83@gmail.com

⁵Mangalam college of Engineering, India, aswathy.su@gmail.com

ABSTRACT

In this decade the main challenge facing in the entire treatment sketch and the evaluation is how vast a brain tumor. one of the most dangerous reason for cancer. Accuracy in quantitative analysis and segmentation of brain are crucial for the treatment sketch. Even though many manual segmentations and magnetic resonance image has emerged they are highly time consuming and error prone. 2D and 3D convolutions using neural networks cannot satisfy the whole treating plans of brain tumors even though if possible they are highly expensive in cost of its computation and the demand in its memory. Here we propose 3D UNet architecture using separable 2D convolutions.

Key words: Neural networks, Brain tumor, UNet, BraTS

1. INTRODUCTION

SA leverages the use of various algorithms to make out the emotion (sentiment) from a subjective text. This is done by breaking down the textual data into components that can be fed into algorithms that weigh various aspects of specific constructs [1]. SA has found its way into many applications ranging from product review analysis to predicting the results of elections. There are many algorithms used for the same. Most of them widely used for these purposes only perform the basic binary analysis. This comes with a cost of reduced accuracy when dealing with sentences or words with multiple implied meanings [2]. This is particularly prevalent in opinions and comments where SA is usually performed.

SA can be performed on various levels. In each level, the actual meaning is finer grained and deals with varied complexities [2].

Surveys have been done covering all aspects of various methods and are referenced where stated.

We propose a S3D-UNet architecture for automatic brain tumor segmentation. In order to make full use of 3D volume information while reducing the amount of calculation, we

adopt separable 3D convolutions. For the characteristics of the isotropic resolution of brain tumor MR images, we design a new separable 3D convolution architecture by dividing each 3D convolution into three branches in a parallel fashion, each with a different orthogonal view, namely axial, sagittal and coronal. We also propose a separable 3D block that takes advantage of the state-of-the-art residual inception architecture. Finally, based on separable 3D convolutions, we propose the S3D-UNet architecture using the prevalent U-Net structure. This network has been evaluated on the BraTS 2018 Challenge testing dataset and achieved an average Dice scores of 0.68946, 0.83893 and 0.78347 for the segmentation of enhancing tumor, whole tumor and tumor core, respectively. Compared with the performance of the training and validation sets, the scores of testing set are lower. This may be due to the difficult cases in testing set because the median values are high. In the future, we will work to enhance the robustness of the network. For volumetric medical image segmentation, 3D contextual information is an important factor to obtain high-performance results. The straightforward way to capture such 3D context is to use 3D convolutions. However, the use of a large number of 3D convolutions will significantly increase the number of parameters, thus complicating the training processes. In the video understanding tasks, the separable 3D convolutions with higher computational efficiency have been adopted.

2. METHODS

2.1 Dataset

The BraTS'2018 challenge brain tumor data set is being used here. The training datasets contains around 285 subjects of brains on multimodal.

2.2 Brain MRI Data Acquisitions

Magnetic resonance imaging (MRI) is commonly used in medical image for analysis of brain tumors. MRI is a non-invasive system, which can be utilized alongside with other imaging modalities, such as computed tomography (CT), positron emission tomography (PET) to give accurate data for

brain tumor structure. However, using these systems alongside MRI is expensive and, in some of the case, can be invasive (e.g. PET). Therefore, different MR that are non-invasive and image both structure and functions are mostly used for brain imaging. MRI machines themselves come with different configurations and produce images with varying intensities. This makes tumor detection a difficult task when different MRI configurations (such as magnetic field strength 1.5, 3 or 7 Tesla) and the acquisition protocol (field of view value, voxel resolution, gradient strength, b0 value, etc.) are used. These configurations have different intensity values across voxels, which result in the same tumorous cells may have different grayscale values when pictured in different hospitals. MRI can show different tissue contrasts through different pulse sequences, making it an adaptable and widely used imaging technique for visualizing regions of interest in the human brain. MRI modalities are combined to produce multi-modal images giving more information about irregular shaped tumors, which are difficult to localize with a single modality. These modalities include T1-weighted MRI (T1), T1-weighted MRI with contrast enhancement (T1c), T2-weighted MRI (T2) and T2-weighted MRI with fluid attenuated inversion recovery (T2-Flair). This multi-modal data contains information that can be used for tumor segmentation with significant improvement in performance. T1-weighting is the most commonly used sequence for the structural analysis; it also allows for an easy annotation of the healthy tissues. In T1-weighted contrast-enhanced images (gadolinium-DTPA), the tumor borders appear brighter because the contrast agent accumulates there due to the disruption of the blood–brain barrier in the proliferative tumor region. In this sequence, the necrotic and the active tumor region can be distinguished easily. In T2-weighted MRI, the edema region, which surrounds the tumor, appears bright. T2FLAIR (FLAIR) is a special sequence, which helps in separating the edema region from the cerebrospinal fluid (CSF) because the free water signal is suppressed. The radiological definition of the tumor margins in the clinical context are often manually determined by the radiologist on the T2 and post-gadolinium T1 images by thresholding boundaries between T2 hyperintense/T1 contrast-enhanced lesions and the surrounding healthy tissue to define the outer margins of a tumor

2.3 Data pre-processing

MRI images was already skull stripped and resampled to Imm3 resolution in dataset. No non-uniform intensity normalization algorithm or non-parametric algorithm has been used to correct for intensity non-uniformities caused by the inhomogeneity of scanner’s magnetic field, because it will be obliterated the T2-Flair signal. We use SimpleITK. to read the NIFTI format data. and covert to numpy array format. The data size of each subject is 240x240x155, we only pick the 0-155th axial slices as training data due to the rest part of brain is very unlikely to have any tumor. The slices are then zero-mean normalized using the mean and standard deviation

so that it can effectively improve calculations and also retain the original information at its maximum.

2.4 Network Architecture

Our brain tumor segmentation model consists two stages. Firstly, we use a 9 layers U-net like architecture to segment full tumor. Secondly, using segmentation result as input for two 7 layers U-net like architectures to segment tumor core and enhancing tumor. The architecture of the proposed network is shown in figure2.4.1. The proposed architecture can be considered as an auto-encoder, where a contracting network tries to learn the image features, and an expanding network tries to use these features to reconstruct a low dimensional data representation of image, such as the ground truth of data. The true ground truth may need to be estimated from a collection of manual segmentations. The contracting layer consists of stacked convolution/pooling layers, while the expanding network is made up of up-sampling/transposed convolution layers. In order to localize, high resolution features from the contracting path are concatenated with the up-sampled output of different stages. These are called “skip connections”. The advantages of using skip connections are shown in detail by Drozdal et al. in. They make a distinction between using long skip and short skip, spanning over a single residual block. The difference between original U-net and our U-net is that we connect a batch normalization after each convolution layer to keep the gradient levels controlled, speed up convergence and reduce the effect of internal covariate shift, so that the parameters of the network do not change rapidly during backpropagation. The filter size of convolution layers is 3 x 3 and we use the same padding to keep the output size of convolution layers unchanged, which is also different from original U-net. To increase the receptive filed at the lowest resolution, two layers of convolution/batch normalization are stacked, which is the idea from VGG16. After two convolution layers and batch normalization, a 2 x 2 max-pooling with stride 2 down-sample image size to 1/2. At each down-sampling step they double the number of feature channels. Every step in the expansive path consists of an up-sampling of the feature map followed by a 2x2 transposed convolution (“up-convolution”) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions. Since it is binary segmentation for each model, the final operation is a 1 x 1 convolution with 1 filter, producing binary prediction that 1 is target and 0 is null.

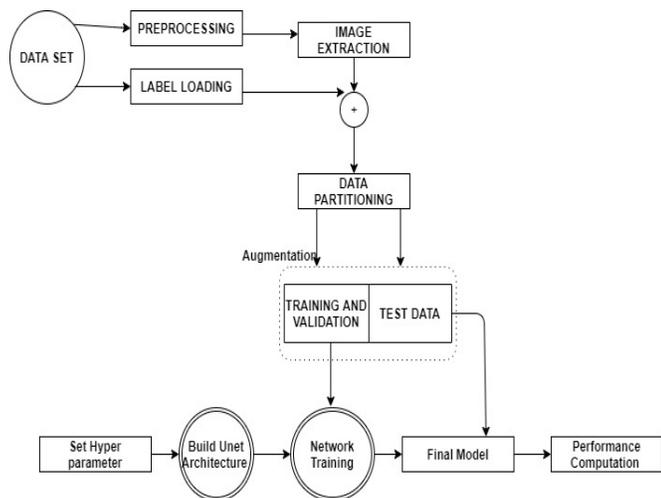


Figure.1: System architecture

2.5. UNet based deep convolutional networks

UNets are algorithm used for image segmentation in medical field. UNets are mostly based on Convolution neural network (CNN). CNN use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage.

3. Experimental Discussion

At first, we use only one U-net model to do three different segmentation tasks (Full Tumor, Tumor Core, Enhancing Tumor),

the result of full tumor is promising but the rest of two is not good. We find out that the problem is because the tumor core and enhancing tumor are too small compare to the whole brain, which means the pixel number of these two labels are too little (average 0.75% for tumor core and 0.45% for enhancing tumor per slice). Therefore, the data imbalanced problem cause that model tend to predict “no tumor” but still can get good result. we knew that tumor core must be inside of edema, and enhancing tumor is part of tumor core. To solve the problem, we use the advantage of fulltumor prediction and calculating the center point of full tumor, then use the center point to crop out the training data for tumor core and enhancing tumor. The number of cropping depend on the size of full tumor and weeven crop the overlap part to do data-augmentation, cropping size is 64x64 fixed. After cropping, the ratio of tumor core and enhancing tumor have achieved average 12.7% and 7.6% per training patch. Then we put the 64x64 training data into another U-net model to train and predict. The result of tumor core predictionand enhancing tumor prediction will be paste back to original full tumor prediction according the center point we got. Instead of using all MRI modalities as, we only use T2 and Flair. Data for full tumor segmentation and only T1ce for tumor core/enhancing tumor segmentation to accelerate training. According, to those radiologists, the edema was segmented primarily from T2 images and FLAIR was used to cross-check the extensionof the edema and discriminate it against ventricles and other

fluid-filled structures. Tumor core and enhancing tumor was both segmented by evaluating hyper-intensities in T1ce images. Figure 3.1 show how segmentation takes place as per training data.

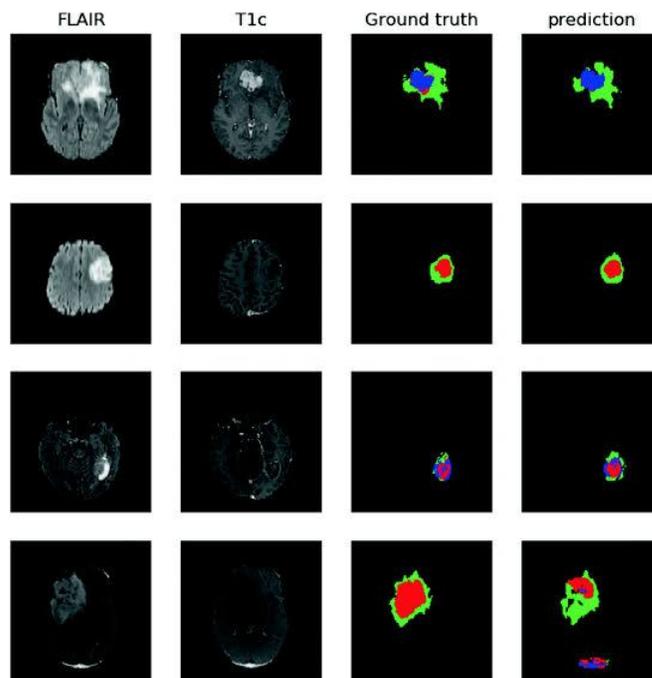


Figure 2: Example of segmentation from training data (BraTS 2018)

4 RESULT

We give numpy file to 9-layer UNet, numpy file is very easy to train. This 9-layer UNet predict tumor core but it cannot predict edema and narcosis so we resize the image and give this to 7-layer UNet we have two 7-layer UNet one predict tumor core and other predict enhancement tumor .we crop the image because it is easy to find the center point. In next step we paste the tumor core and enhancement tumor into full tumor as a result we can visible the edema and narcosis correctly we can give treatment to the particular section.

5. CONCLUSION

In this paper, we demonstratethat the U-Net with separable 3D convolutions can achieve promising results in thefield of medical image segmentation. In the future work, we will continue to improve the structure of the network and usesome post-processing methods such as fully connected conditional random field tofurther improve the segmentation performance. Image segmentation, especially semantic segmentation, is a fundamental and classicproblem in computer vision. It refers to partitioning an image into several disjointsemantically meaningful parts and classifying each part into a pre-determined class. Inthe application of brain tumor segmentation, the task includes the division of severalsub-regions, such as GD-enhancing tumor, peritumoral edema, and the necrotic andnon-enhancing tumor core [1]. Accurate segmentation and quantitative analysis ofbrain tumor are critical for diagnosis and treatment planning.

Therefore, there is a strong need for a fully automatic method for brain tumor segmentation. However, brain tumor segmentation is a challenging task because MRI images are typically acquired using various protocols and magnet strengths, which results in the non-standard range of MR images. In addition, brain tumors can appear anywhere in the brain, and their shape and size vary greatly. Furthermore, the intensity profiles of tumor regions are largely overlapped with healthy parts. Due to the challenge of brain tumor segmentation and the broad medical prospect, many researchers have proposed various methods to solve the problem of brain tumor segmentation. Brain tumor segmentation methods can be divided into different categories according to different principles [2]. Broadly, these methods can be divided into two major categories: generative methods and discriminative methods. Generative methods typically rely on the prior information about the appearance of both healthy tissues and tumors. The proposed models often regard the task of segmentation as a problem of a posteriori distribution estimation. On the contrary, discriminative methods use very little prior information and typically rely on a large number of low-level image features to learn the distribution from the annotated training images.

REFERENCES

1. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* 34, 1993–2024 (2015)
2. Ms. K. Nirmala Devi, Ms. K. Mouthami, Dr. V. Murali Bhaskaran 'Sentiment Analysis and Classification Based on Textual Reviews', 2012.
3. Study of Random Tree and Random Forest Data Mining Algorithms for Microarray Data Analysis - Ajay Kumar Mishra, Bikram Kesari Ratha *ISSN(Online): 2349-9338*
4. Sona Thaheri I, Musa Mammadov *Int. J. Appl. Math. Comput. Sci.*, 2013, Vol. 23, No. 4, 787–795
5. LEARNING THE NAIVE BAYES CLASSIFIER WITH OPTIMIZATION
6. www.sciencedirect.com/topics/neuroscience/support-vector-machines
7. www.medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72
8. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text C.J. Hutto Eric Gilbert Georgia Institute of Technology, Atlanta, GA 30032
9. <https://www.kdnuggets.com/2018/08/emotion-sentiment-analysis-practitioners-guide-nlp-5.html>
10. Comparative study of classification algorithms used in a sentiment analysis. Amit Gupte, Soirabh Joshi, Pratik Gadgil, Akshay Kadam.
11. Cluster computing, "Brain Tumor Detection and Segmentation Using A Wrapper Based Genetic Algorithm For Optimized Feature Set", (2018) (SCIE indexed).
12. Cogent Engineering, "An Improved Segmentation Algorithm From T2 and FLAIR Multimodality MRI Brain Images by Support Vector Machine and Genetic Algorithm", 2018 (Scopus Indexed, ESCI).
11. International Journal of Applied Engineering Research, "A Survey on Segmentation of Brain Tumor From MRI Brain Images", 2014.
12. IEEE Digital Explore Library "A Survey on Detection of Brain Tumor from MRI Brain Images", 2014.
- Indian Journal Of Science and Technology, "Quick Detection of Brain Tumor Using A Combination of EM And Level Set Method", 2015.