Residual Volumetric Network for Segmentation of Prostate from MR Image

Lequan Yu, Xin Yang, and Pheng-Ann Heng

Department of Computer Science and Engineering,
The Chinese University of Hong Kong

1 Introduction

In this manuscript, we will describe our automated method for segmentation of the prostate from MR images in PROMISE12 Challenge. As a paper is preparing, we will roughly describe our method and a detailed paper will be uploaded after publication. We propose a 3D convolutional neural networks (3D CNNs) based method for automated segmentation of the prostate from MR images. The proposed 3D network takes advantage of fully convolutional architecture to perform efficient, end-to-end, volume-to-volume prediction. More importantly, we introduce the recent proposed residual learning technique into our network, which can alleviate vanishing gradients problem and improve the performance of our network.

2 Method

Fig. 1 demonstrates the architecture of our proposed prostate residual volumetric network. It employs 3D fully convolutional architecture and is organized in a residual learning scheme. The building blocks of our network, such as the convolutional, max-pooling and deconvolutional layers, are all implemented with a 3D manner, thus the network can highly preserve and deeply exploit the 3D spatial information of the input volumetric data. Note that our network stems from the fully convolutional architecture [3], and hence can take arbitrary-sized volumetric data as input and output corresponding sized predictions within a single forward process, which is very efficient in handling large MRI dataset.

Previous studies [5] have shown that smaller convolutional kernels are more efficient in 2D network design. Because the effective receptive field size of stacked small kernels is equivalent to that of one large kernel (the effective receptive field of three $3\times3\times3$ kernels is same as one $7\times7\times7$ kernel), while giving lower computation cost. Therefore, we adopt small convolution kernels with size of $3\times3\times3$ in convolutional layers. Each convolutional layer is followed by a rectified linear unit (ReLU). Note that we also employ batch normalization layer (BN) [1] before each ReLU layer. The BN layer can accelerate the training process of our network. At the end of the network, we add a $1\times1\times1$ convolutional layer as a classifier to generate the segmentation results and further get the segmentation probability volumes after passing the softmax layer.

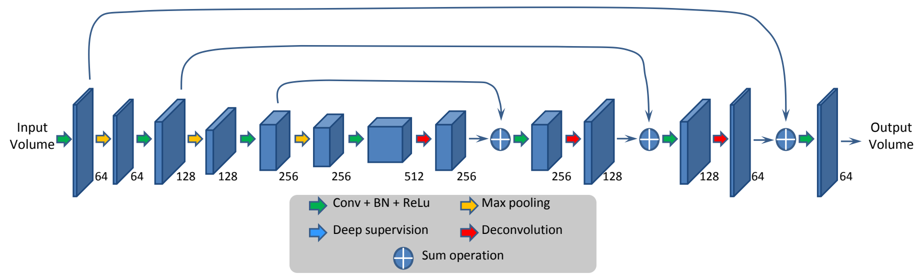

Fig. 1: Illustration of our proposed residual volumetric network. Numbers represent the number of feature volumes in each layer.

Note that our network might appear similar to U-Net [4], but there are critical differences: We use summation units instead of concatenation units when combining different paths, and thus we can reformulate our network as residual learning scheme; additionally, we adopt recently developed batch normalization technique to improve our performance.

3 Experiments and Results

The proposed method was implemented with C++ and Matlab under the open source deep learning library Caffe [2]. To evaluate our method, we used 10-fold cross validation on the training set available from the PROMISE12 dataset[1]. Before training, we first resize the samples to a common fixed resolution of $0.625\times0.625\times1.5$ mm and normalize them as zero mean and unit variance. Table 1 illustrates the quantitative cross validation results of our method and other automated challenge methods which also reported cross validation results.

Table 1: Quantitative comparison between the proposed approach with other methods

Methods	mean DSC	median DSC
CUMED(our)	0.8749	0.8928
ScrAutoProstate	0.8600	0.8900

References

1. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)

[1] http://promise12.grand-challenge.org

2. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv:1408.5093 (2014)
3. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3431–3440 (2015)
4. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)
5. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR abs/1409.1556 (2014)