Automatic 3D Prostate MR Segmentation with Densely-Connected Volumetric ConvNets

Qikui Zhu
School of Computer
Wuhan University
WuHan, China,430079
Email: QikuiZhu@163.com

1 Abstract

Automated prostate segmentation in 3D medical images play an important role in many clinical applications, such as diagnosis of prostatitis, prostate cancer and enlarged prostate. However, it is still a challenging task due to the complex background, lacking of clear boundary and various shape and texture between the slices. In this paper, we propose a novel 3D convolutional neural network with densely-connected layers to automatically segment the prostate from Magnetic Resonance(MR) images.

2 Introduction

Medical image segmentation is an essential part of medical image analysis. Accurate segmentation of medical image provides very useful information for computer aided diagnosis of many cancer as well as other diseases. Traditionally, those images are delineated by radiologists in a slice-by-slice manner, which is time-consuming and heavily depends on radiologists' experience and prone to inter- and intra-observer variations. Therefore, an automated medical image segmentation method is highly demanded in clinical practice.

However, automated medical image segmentation is very challenging for several reasons. First, due to many slices only have small part of segmented tissues specifically at the apex and base, which always led to those slices lack of clear boundary and make the automated segmentation harder. Second, due to image process and related diseases would bring imaging artifacts, and those imaging artifacts always influence the process of segmentation. Third, the tissues have a wide variation in size and shape among different slices. Fourth, the surrounding tissues always have similar appearance and intensity distribution.

3 Method

In order to fully leverage the 3D spatial contextual information of volumetric data to accurate segment prostate images, we extend the 2D dense block into 3D dense block to enable volume-to-volume segmentation. The proposed volumetric convolution network is extended from the 3D architecture reported in Yu et al[1]. Which consists of two parts: down-sampling path and up-sampling path. The down-sampling path consists of one convolutional layer, three dense blocks with residual connection and three average pooling layers. The pooling layers with stride 2, which gradually reduce the resolution of feature map and increases the receptive field of the convolutional layer. For the up-sampling path, which contains three deconvolutional layers and three dense blocks with residual connection. In addition, we also integrate a deep supervision mechanism in the up- sampling path. Three additional supervision layers be added after dense block. The additional supervision layer consists of one deconvolution layer which up-samples the feature map to its original size and one convolution layer which output segmentation result. Those additional supervision layers bring two advantages. First, which can accelerate the network convergence speed during training. Second, additional supervision can regularize the network and reduce the segmentation noise. To further improve the gradient information flow between down-sampling path and up-sampling path, inspired by U-Net, we employ long connections between the down-sampling path and up-sampling path. The long connections connect the block with same resolution in the down-sampling path and up-sampling path. Those connections bring many advantages. First, this can effectively propagate context and gradient information both forward and backward during training and overcome the vanishing-gradient problem. Second, which can alleviate the problem of information loss. When the feature map passes the average pooling layer, part of feature information is abandoned and many detailed information is lost. This always results in the segmentation results cannot obtain accurate boundaries. After adding the long connections, the later block can regain lost feature information from earlier block in down-sampling path and alleviate information loss.

4 Result

4.1Implementation Details

Our network is trained end-to-end with Stochastic Gradient Descent (SGD) method, and the framework is implemented under the open-source deep learning library Keras. In the training phase, the learning rate is initially set as 0.0001 and decreased by a weight decay. The momentum is 0.9, and due to the limitation of the memory, we choose 8 as the batch size. Experiments are carried out on GTX1080ti GPU with 11GB memory. During training the network, we employ a randomly cropping strategy. We randomly cropped 16×64×64 sub-volumes from the training data every iteration. In the test phase, we used overlapped sliding windows strategy to crop sub-volumes and used the average of the probability maps of these sub-volumes to get the whole volume prediction. The sub-volume size was also 16×64×64 and the stride was 8×32×32.

4.2 Results

We randomly split the 50 training cases into a training set with 40 cases and a validation set with 10 cases. The proposed model achieves a dice score of 0.897 over the entire volume on the validation set.

Part References

[1] Yu L, Yang X, Hao C, et al. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images[C]// AAAI. 2017.