3D Densely Connected Convolution Neural Networks for Pulmonary Parenchyma Segmentation from CT Images

Liang Zhao
College of Mechanical and Electrical Engineering, Shanghai Jian Qiao University, Shanghai, China
Email: 17089@gench.edu.cn

Abstract. Lung cancer is one of the deadliest diseases in the world today. It kills many people every year. For the diagnosis and treatment of lung cancer, accurate segmentation of lung tissue from CT images is an important process. It is necessary to design a fast and accurate segmentation method to accomplish this task. In the traditional computer-aided diagnosis system, the segmentation of Lung parenchyma is very complex, and the segmentation result depends on the performance of the parameters set in the previous stage. In order to solve these problems, we propose a 3D densely connected convolution neural network which based on deep learning. It has three densely connected blocks and three deconvolution layers. The experimental data set was taken from the public LIDC-IDRI database. A total of 888 samples with slice thickness less than 2.5 mm were selected in the experiments. And the number of samples of training set, test set and validation set is 708, 90 and 90 respectively. In addition, the experimental results show that our method is more accurate than 3D-Unet, but it requires less training parameters.

1. Introduction
Lung cancer is one of the malignant tumours with high morbidity and mortality [1]. Early detection of lung cancer is very important. Low dose computed tomography (CT) is an effective method to detect early asymptomatic lung cancer, which can detect 90% of the lesions. Automatic segmentation of pulmonary parenchyma in CT images is the basis of qualitative and quantitative analysis of medical diagnosis. It can help doctors diagnose and plan treatment for patients. It is the main premise of image-guided surgery, tumour radiotherapy and clinical treatment evaluation, and is an indispensable part of interventional ablation and magnetic induction hyperthermia. However, it is very difficult to accurate automatic segmentation for pulmonary parenchyma in CT images because of the complexity. In recent years, the use of deep learning technology for medical image segmentation is becoming a new research direction. So called deep learning is a kind of artificial neural network system with multiple hidden layers. Its essence is to extract features of the objects from low layer to high layer. It can process data such as images, sounds, videos and text. Among them, convolutional neural network (CNN) has achieved significant performance gains in large-scale image classification and segmentation tasks. Image segmentation methods based on CNN can be divided into two kinds: two dimensional segmentation network and three-dimensional segmentation network. Two dimensional segmented networks include the full convolutional neural network (FCN) [2] and 2D-Unet networks [3]; three dimensional segmented networks include 3D-Unet [4] and V-net networks [5]. For CT image, the accuracy of 2D segmentation network is lower than that of 3D segmentation network. The reason is that CT image is a kind of 3D image, and the 3D information in the image is very rich, but
the 2D segmentation network cannot use the information. The problem with 3D segmentation networks is that they require a lot of memory to run. The current limitations of GPU memory are such that 3D segmentation networks require a large-scale reduction in the size of the input image. This leads to the loss of image features, which affects the segmentation accuracy. For the same reason, the depth of 3D segmentation network needs to be reduced compared with 2D segmentation network, which leads to the decrease of receptive field. The large-scale reduction and the small receptive field of the images result in the unsatisfactory accuracy of the 3D segmentation network.

Recently, a new network connection model called densely connected convolutional network is proposed [6], which creates a cross-layer connection to connect the front and back layers of the network. All layers are connected directly and it must be guaranteed that maximum information can be transferred between layers in the network. Its main advantages are: 1. fewer parameters are needed. 2. The information (in forward computation) or the gradient (in backward computation) is kept better in the whole network, and the deeper model can be trained. 3. Dense connection has the effect of regularization, reducing over-fitting on fewer training sets. Based on the classical densely connected convolution network, and for fast and accurate segmentation of pulmonary parenchyma, a 3D densely connected convolution network is proposed in this paper.

2. Network Architecture
Figure 1 shows our proposed network architecture. It is a 3D fully convolutional network. There are three densely connected blocks named as DenseBlock1, DenseBlock2 and DenseBlock3 in the down-sampling path. DenseBlock1 is comprised of 4 dense connected layers, and was prefixed with a layer comprising of 64 convolution filters of size 3×3×3 on the input images. The dense connectivity scheme of DenseBlock1 is given in figure 2. It can be expressed as:

\[ X_L = H_L \left( [X_0, X_1, \ldots, X_{L-1}] \right) \]

where \( X_L \) is the output of the \( L \)th layer. \( H_L ([X_0, X_1, \ldots, X_{L-1}]) \) means implementing concatenation operation from layer 0 to layer \( L-1 \). Concatenation operation is the merging of channels, not the addition of values. As figure 2 plotted, each layer takes all the features extracted from the previous layers as input. Thus, the dense connectivity makes features to be reused among all these connected layers.

![Figure 1](image1.png)  
**Figure 1.** The architecture of the 3D densely connected convolution neural networks.  

![Figure 2](image2.png)  
**Figure 2.** The dense connectivity scheme of DenseBlock1.
In each transformation layer of dense block, there are a batch normalization layer, a rectified linear unit, and a 3×3×3 convolution layer and the growth rate is 16. In figure 1s and 2, “BN”, “ReLU” and “Conv” denote these three layers respectively. DenseBlock2 and DenseBlock3 are similar to DenseBlock1, and are comprised of 8 and 16 transformation layers respectively. Between two adjacent dense blocks there are several transition layers. They can change the size of the feature map through convolution and pooling. In a transition layer, there are a batch normalization layer, a rectified linear unit, a 1×1×1 convolution layer and a 2×2×2 max pooling layer. In the down-sampling path, there are a batch normalization layer, a rectified linear units, two 1×1×1 convolution layer and three 2×2×2 deconvolutional layers. The deconvolutional layers denoted as “Deconv” in figure 1. They make the output images the same size as the input ones.

3. Experiments and Results

3.1. Dataset

The experimental data set was taken from the public LIDC-IDRI database [7], which contains 1018 CT image samples. These CT images were produced by seven different agencies. The thickness of the slices was 0.6-5.0 mm, and the median was 2.0 mm. A total of 888 samples with slice thickness less than 2.5 mm were selected in this study. In this experiment, the CT images of lung and its corresponding segmentation image were used as training set to train the network. In order to prevent the model from over fitting, some training data need to be used as verification set. In this paper, 10% of the training data are selected randomly as the verification set. The number of samples of training set, test set and validation set is 708, 90 and 90 respectively.

3.2. Network Training and Testing

The experimental steps are: first, process the data; second, use the training data to train the model; third, use the trained model to build the algorithm; finally use the test set to verify the performance of the algorithm. Data processing is mainly on the image resolution standardization and cutting. Image resizing and gray value standardization are carried out dynamically during training and testing. The network parameters are set as follows: batch size is set to 4, the loss function is set as Dice loss function, and the learning rate is set to 0.001. The hardware configuration of this experiment is: CPU, Intel i7-8700K; Memory, 64G; GPU, 8G. And software configuration is: Windows 10, Cuda9.0, cudnn, tensorflow and Keras.

3.3. Result Analysis

Figure 3 shows the prediction results for a test sample using the proposed method. Obviously, the segmentation results are very close to the ground truth. Dice coefficient is a function used to estimate the similarity between two samples. It is one kind of important segmentation evaluation coefficient, can be expressed as:

\[
\text{Dice} = 2 \frac{V_{seg} \cap V_{gt}}{V_{seg} + V_{gt}}
\]  

(2)

where \(V_{seg}\) and \(V_{gt}\) represent the segmentation results and the ground truth respectively. The higher Dice values, the closer segmentation results are to the ground truth. The proposed method is compared to 3D-Unet on the test set, and the results are shown in table 1. It can be observed that the proposed method can generally achieve better performance than 3D-Unet. However, the number of training parameters of the proposed method is only half that of 3D-Unet.
4. Conclusions
A 3D densely connected convolution neural networks is proposed to automatically segment pulmonary parenchyma from CT images. It is a 3D fully convolutional network. There are three densely connected blocks in the down-sampling path, while three deconvolutional layers in the down-sampling path. They make the output images the same size as the input ones. By experiment, we found that the proposed method is more accurate than 3D-Unet, however it requires less training parameters.

Acknowledgements
This work was funded by the Shanghai Jian Qiao University school level project, the project No. SJQ-17003.

References
[1] Sluimer I, Schilham A, Prokop M and van Ginneken B 2006 Computer analysis of computed tomography scans of the lung: A survey IEEE T. Med. Imaging 25 385-405.
[2] Shelhamer E, Long J and Darrell T 2017 Fully convolutional networks for semantic segmentation IEEE T. Pattern Anal. 39 640-651.
[3] Ronneberger O, Fischer P and Brox T 2015 Proc. of the 2015 Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Munich) vol 9351 (Cham: Springer) pp 234-241.
[4] Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O 2016 Proc. of the 2016 Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Athens) vol. 9901 (Cham: Springer) pp 424-432.
[5] Milletari F, Navab N, Ahmadi S 2016 Proc. of the 4th Int. Conf. on 3D Vision (California) vol 1 (Piscataway: IEEE) pp 565-571.
[6] Huang G, Liu Z, Weinberger K Q and van der Maaten L 2016 Densely connected convolutional networks Preprint gr-qc/0806993.
[7] Armato S G, Mclellan G, Bidaut L, et al. 2011 The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans Med. Phys. 38 915-931.