3D Region Proposal U-Net with Dense and Residual Learning for Lung Nodule Detection

Zhongliu Xie
Imperial College London
zhongliu.xie@outlook.com

***Abstract*—Detection of pulmonary nodules in thoracic computed tomography (CT) scans plays a crucial role in early diagnosis of lung cancer, which paves the way for early treatment and could significantly enhance a patient's chance of survival. Manual detection is highly time-consuming, tedious and prone to mistakes, calling for the use of computer-aided detection (CAD), both to improve the efficiency and lower the rate of misdiagnosis. Over the past years, a large number of such CAD systems have been proposed, many of which follow a two-phase paradigm: 1) nodule candidate detection, and 2) false positive reduction. We propose a 3D region proposal U-Net with dense and residual learning for pulmonary nodule detection. Our approach abandons the conventional two-phase paradigm, and trains a single network for end-to-end detection without the additional false positive reduction process. We borrow the region proposal network part from the popular Faster R-CNN framework for object detection in natural images, and extend it to 3D with a patch-based scanning scheme. Furthermore, our model design draws insights from DenseNet on dense learning, ResNet for residual learning and U-Net for small object detection. Overall, we achieved a FROC score of 0.9226.**

I. INTRODUCTION

Past decades have witnessed an increasing attention towards automatic detection of pulmonary nodules in thoracic computed tomography (CT) scans. Pulmonary nodules are often considered as a crucial indicator of primary lung cancer, which is a leading cause of death in recent years, accounting for an approximate 27% of cancer-related deaths in the United States [1] and 1.3 million deaths every year worldwide [2]. Fast and accurate detection of such pulmonary nodules plays a crucial role in early diagnosis of lung cancer, which paves the way for early treatment and could significantly enhance a patient's chance of survival [3, 4, 5].

Under the CT imaging, pulmonary nodules are often visible as small opaque structures that are roughly of spherical shape and brighter than their surroundings. Once correctly identified, size (diameter) measurement and appearance characterization are then conducted by radiologists/surgeons to carry out cancer malignancy diagnosis [6], which is followed by surgical intervention if necessary. However, manual detection is highly time-consuming, tedious and prone to mistakes, calling for the use of computer-aided detection (CAD), both to improve the efficiency and lower the rate of misdiagnosis.

Over the past years, a large number of such CAD systems have been proposed, many of which follow a two-phase paradigm: 1) nodule candidate detection, and 2) false positive reduction. In the first phase, a relatively coarse scan is typically performed over the whole image volume to locate a considerable set of potential candidates (often with high sensitivity but also high false positive rate), which then go through a more rigorous screening process to filter out false positive detection in the second phase. Traditional candidate detection methods include the use of intensity thresholding, shape curvedness and mathematical morphology [5, 7, 8]. Traditional false positive reduction methods include combinatorial analytics over shape, position, intensity, gradient features, texture characteristics and contextual information [5, 9].

More recently, deep learning with convolutional neural networks (CNN) has gained wider popularity. Most notably, Ding et al. [10] proposed using a 2D Faster R-CNN [12] that incorporates an ImageNet pre-trained VGG-16 model [11] to perform candidate detection, followed by a 9-layer 3D CNN (6 convolutional layers and 3 fully connected layers) for false positive reduction. Their methodology yielded a FROC-score (average sensitivity over 1/8, 1/4, 1/2, 1, 2, 4, 8 false positives per scan) of 0.893 on the LUNA dataset (topping the LUNA16 leaderboard in May 2017).

However, it requires training two separate networks for the two phases, introducing considerable engineering work. Moreover, the 2D Faster R-CNN method can be further divided into two components: a region proposal network (RPN) and a region-of-interest (ROI) classifier, both produce binary classification (nodule vs. non-nodule) and bounding box regression (if nodule) outputs, making the design a bit redundant. Other relevant works include the use of 2D multi-view CNNs proposed by Setio et al. [13] to learn representative features for pulmonary nodule detection. This method can incorporate relatively wide volumetric spatial information for detection by extracting many 2D patches from differently oriented planes, and achieved 82.66% In terms of false positive reduction alone, Dou et al. [14] proposed using a multi-level contextual 3D CNNs to perform classification over three differently sized patches centering the nodule candidate, followed by weighted label fusion. They achieved a FROC score of 0.827 on the false positive reduction track of LUNA16 challenge.

II. METHODS

We propose a 3D region proposal U-Net with dense and residual learning for pulmonary nodule detection. Our approach abandons the conventional two-phase paradigm, and trains a single network for end-to-end detection without the additional false positive reduction process. We borrow the region proposal network part from the popular Faster R-CNN framework [12] for object detection in natural images, and extend it to 3D with a patch-based scanning scheme. Furthermore, our model design draws insights from DenseNet [15] on dense learning, ResNet [16] for residual learning and U-Net [17] for small object detection. Overall we achieved a FROC score of 0.9226.

A. Lung Segmentation

To narrow the scope of detection, lung segmentation is performed as a preliminary step. The lungs are the primary organs of the respiratory system in the human body, and consist of a left and a right lung, situated within the thoracic cavity of the chest. Under the CT imaging, the lung region mainly contains air, which has a value of around -1000 in Hounsfield scale. The lungs and their immediate surroundings may expand to over 10 liters of volume during breathing in, and contracting to around 1 liter during breathing out. For that reason, intensity thresholding can be applied to lung-specific level, followed by morphological methods including 3D connected component analysis, region erosion, region expansion, and volume thresholding to segment the lungs.

B. Nodule Detection

Since the image volume is too large to fit into a standard GPU for CNN processing directly even after lung segmentation. A window-based detection approach is employed, where a 3D window slides over the lung region and a cubic patch is extracted each time, fed to the 3D RPN for nodule detection. Figure 1 illustrates the detection framework, which is a dual-path network consisting of two U-shaped sub-networks: one follows the DenseNet paradigm while the other follows the ResNet paradigm. A label fusion is applied using non-maximal suppression (NMS) method [12]. The neural network implementation was based on the popular PyTorch framework [18]. Further details may not be published at this point due to paper preparation, and will be released later.

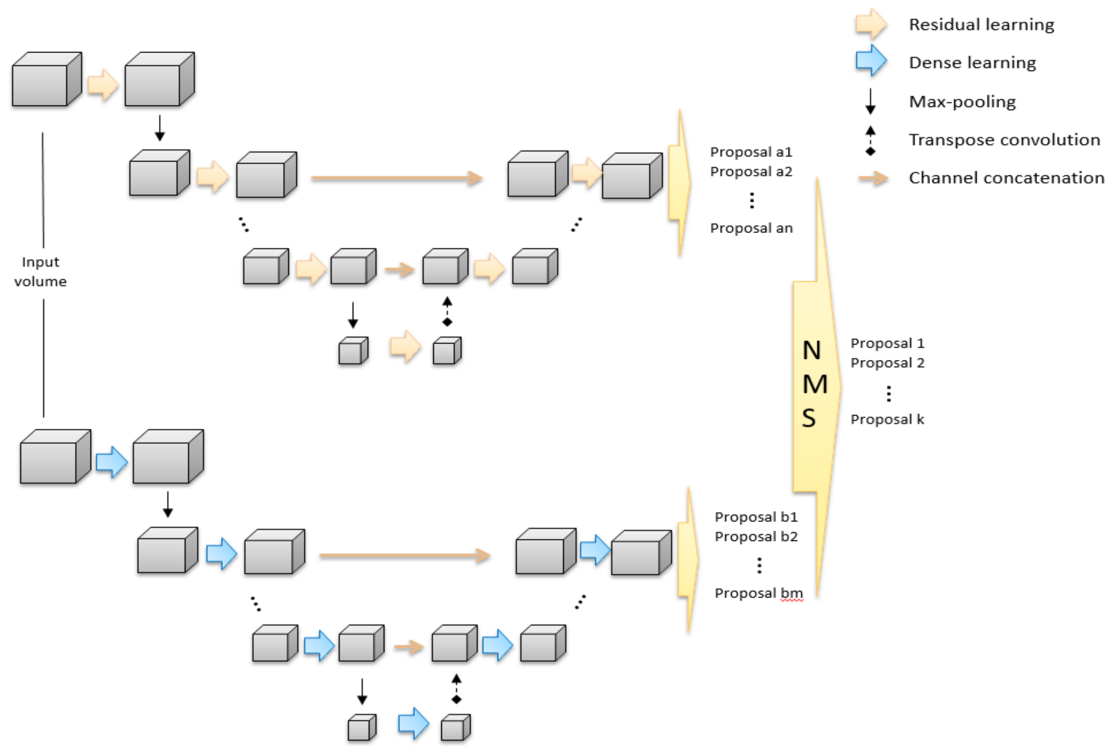

Figure 1. 3D region proposal U-Net with dense and residual learning for pulmonary nodule detection

III. Results

An experimentation has been carried out on the LUNA16 dataset to evaluate the proposed methodology using 5-fold cross validation. According to the LUNA16 standard, nodule detection performance is evaluated using the Free-Response Receiver Operating Characteristic (FROC) analysis, which calculates the average sensitivity over 7 false positive rates: 1/8, 1/4, 1/2, 1, 2, 4, 8 FPs per scan. Our best performance achieves detection sensitivity respectively at [0.8015, 0.8606, 0.9199, 0.9526, 0.9682, 0.9776, 0.9776], obtaining an average FROC score of 0.9226, as shown in Figure 2. We trained the model for 500 epochs, and the average FROC scores over different epochs are shown in Figure 3.

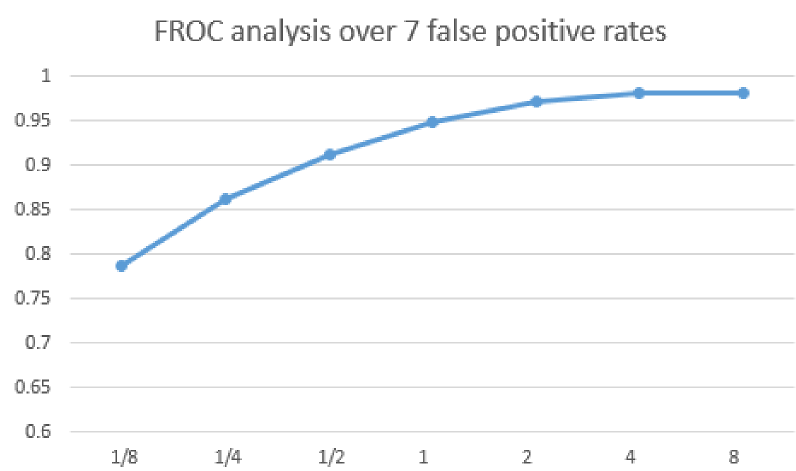

Figure 2. FROC analysis over 7 false positive rates

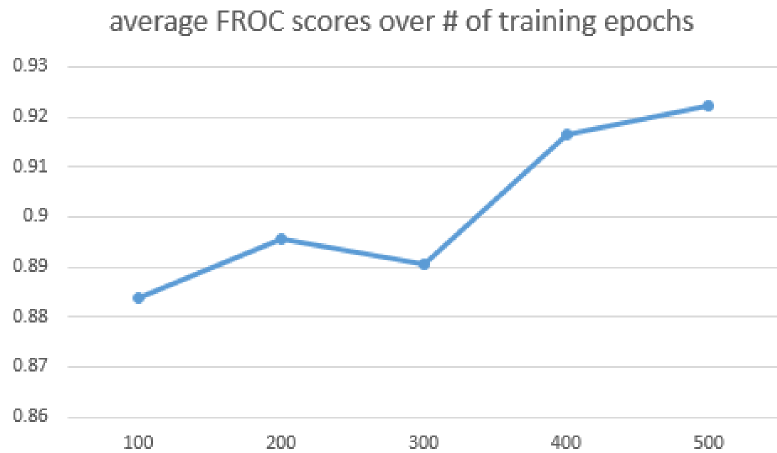

Figure 3. Average FROC scores over different numbers of training epochs

The experiments were performed on a standard machine with two NVIDIA K80 GPUs, where network training took around a week for each fold of cross validation. At test time, end-to-end workflow took less than 1 minute, among which, lung segmentation took around 20~30s and nodule detection by network inference took around 10~20s, depending on the size of CT image.

REFERENCES

[1] American Cancer Society, 2016. Cancer facts and figures 2016.

[2] Siegel, R.L., Miller, K.D., Jemal, A. (2015): Cancer statistics, 2015. CA: a cancer journal for clinicians 65(1), 5--29.

[3] I. Sluimer et al. (2006): "Computer analysis of computed tomography scans of the lung: A survey," IEEE Trans. Med. Imag., vol. 25, no. 4, pp. 385--405.

[4] B. van Ginneken et al. (2010), "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The anode09 study," Med. Image Anal., vol. 14, no. 6, pp. 707--722.

[5] C. Jacobs et al. (2014): "Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images," Med. Image Anal., vol. 18, no. 2,pp. 374--384.

[6] H. MacMahon et al. (2005): "Guidelines for management of small pulmonary nodules detected on CT scans: A statement from the fleischner society 1," Radiology, vol. 237, no. 2, pp. 395--400.

[7] A. A. Setio et al. (2015): "Automatic detection of large pulmonary solid nodules in thoracic CT images," Med. Phys., vol. 42, no. 10, pp. 5642--5653.

[8] K. Murphy et al. (2009): "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbor classification," Med. Image Anal., vol. 13, no. 5, pp. 757--770.

[9] T. Messay et al. (2010): "A new computationally efficient cad system for pulmonary nodule detection in ct imagery," Med. Image Anal., vol. 14, no. 3, pp. 390–406.

[10] J. Ding, A. Li, Z. Hu, and L. Wang (2017): "Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks," MICCAI 2017.

[11] K. Simonyan and A. Zisserman (2015): "Very Deep Convolutional Networks for Large-Scale Image Recognition," ICLR 2015.

[12] S. Ren, K. He, R. Girshick, J. Sun (2015): "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. on Pattern Analysis and Machine Intelligence.

[13] A. A. A. Setio et al. (2016): "Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks," IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1160--1169.

[14] Q. Dou, H. Chen, L. Yu, J. Qin, P.-A. Heng (2016): "Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection," IEEE Trans. on Biomedical Engineering, vol. 64, no. 7, pp.1558--1567.

[15] G. Huang, Z. Liu, L. van der Maaten, W. Q. Weinberger (2017): "Densely connected convolutional networks," CVPR 2017.

[16] K. He, X. Zhang, S. Ren and J. Sun (2016): "Deep Residual Learning for Image Recognition," CVPR 2016.

[17] O. Ronneberger, P. Fischer and T. Brox (2015): "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI 2015.

[18] PyTorch deep learning framework: http://pytorch.org/