Automatic Detection of Lung Nodules Using 3D Deep Convolutional Neural Networks

*FU Ling*¹ (傅 玲), *MA Jingchen*¹ (马璟琛), *CHEN Yizhi*¹ (陈奕志), *LARSSON Rasmus*^{1,4}, *ZHAO Jun*^{1,2,3*} (赵 俊)

(1. School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

SJTU-UIH Institute for Medical Imaging Technology, Shanghai Jiao Tong University, Shanghai 200240, China;
 MED-X Research Institute, Shanghai Jiao Tong University, Shanghai 200030, China;

4. School of Technology and Health, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden)

© Shanghai Jiao Tong University and Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract: Lung cancer is the leading cause of cancer deaths worldwide. Accurate early diagnosis is critical in increasing the 5-year survival rate of lung cancer, so the efficient and accurate detection of lung nodules, the potential precursors to lung cancer, is paramount. In this paper, a computer-aided lung nodule detection system using 3D deep convolutional neural networks (CNNs) is developed. The first multi-scale 11-layer 3D fully convolutional neural network (FCN) is used for screening all lung nodule candidates. Considering relative small sizes of lung nodules and limited memory, the input of the FCN consists of 3D image patches rather than of whole images. The candidates are further classified in the second CNN to get the final result. The proposed method achieves high performance in the LUNA16 challenge and demonstrates the effectiveness of using 3D deep CNNs for lung nodule detection.

Key words: lung nodule detection, computer-aided detection (CAD), convolutional neural network (CNN), fully convolutional neural network (FCN)

CLC number: R 318 Document code: A

0 Introduction

Lung cancer is the leading cause of cancer deaths worldwide. The 5-year survival rate of it can be improved drastically from 17% to 57% when diagnosed at a localized stage rather than at the distant stage; however, more than one half of lung cancers are diagnosed at the distant stage^[1]. This raises a need for the early diagnosis of lung cancer. One of the potential precursors to lung cancer is the presence of lung nodule, the detection of which is closely related to the diagnosis of lung cancer. At present, spiral computed tomography (CT) is regarded as one of the most appropriate tools for the screening of lung nodules^[2]. However, it is a tedious work for the radiologists to detect the lung nodules manually in hundreds of CT slices among

Received date: 2017-11-09

numerous patients. In the recent decades, computer aided detection (CAD) systems for lung nodules have been developed to assist the radiologists by automatically locating the possible nodules, thus improving the accuracy and efficiency.

A lung nodule CAD system mainly consists of two stages: lung nodule detection stage and false positive reduction stage, and they are referred to as the first and second stages respectively. In the first stage, many lung nodule candidates are detected from hundreds of CT slices in one CT scan to achieve high sensitivity. Li et al.^[3] developed three selective enhancement filters based on Hessian matrices for specific shapes. Each of the filters can simultaneously enhance a specific object such as dot-like lung nodules and suppress other objects such as line-like blood vessels and plane-like airway walls. To estimate the nodule locations and exclude false positives, Tan et al.^[4] further developed the filter method by using the maxima of divergence of the normalized gradient of the image in 3D. Messay et al.^[5] used multiple gray level thresholds and corresponding specific morphological opening operation to obtain 15 intermediate candidate masks. These masks were further processed with an expert filter based on size and compactness followed by a

Foundation item: the National Natural Science Foundation of China (No. 81371624), the National Key Research and Development Program of China (No. 2016YFC0104608), the National Basic Research Program of China (No. 2010CB834302), and the Shanghai Jiao Tong University Medical Engineering Cross Research Funds (Nos. YG2013MS30 and YG2014ZD05) *E-mail: junzhao@sjtu.edu.cn

logically OR operation to obtain the final candidate 1 mask. Setio et al.^[6] combined three algorithms specifically designed for solid nodules^[7], subsolid nodules^[8] and large nodules^[9]. All candidates from different algorithms were combined and then candidates located closer than 5 mm were merged. The number of lung nodule candidates detected in the first stage is much larger than the number of true nodules. These candidates are classified as true lung nodules or non-nodules in the second stage. Some handcrafted features are used in the second stage; the design of these features is mainly based on intensity, texture, shape and contextual information of lung nodules. Ciompi et al.^[10] developed the bag-of-frequencies descriptor, which has scale-invariant and rotation-invariant properties, to describe nodule morphology. Tan et al.^[4] used 45 features including gauge derivative invariant features, classical geometric or shape descriptors and regional or greyvalue descriptors. Recently, convolutional neural network (CNN) methods increasingly pervade the area of false positive reduction for lung nodule detection. Setio

et al.^[6] developed multi-view CNNs whose input was a set of 2D patches from 9 different oriented planes to reveal 3D spatial information to some degree. Dou et al.^[11] used 3D CNNs to take full advantage of the 3D spatial information. The networks were fed with different sizes of patches to incorporate different levels of contextual information. Both methods achieved leading results in lung nodule detection^[6,11].

Lung nodules can appear with substantial variations in intensity, shape, size and context. It is therefore hard to develop pre-defined rules or features for their detection. Some methods, developed for specific types of nodules, can be combined to improve the nodule detection sensitivity in the first $stage^{[6]}$. However, the number of false positives increases rapidly with the number of detected nodules, thus demanding more efforts to exclude all non-nodule candidates in the second stage. Though much work has been done in obtaining the lung nodule candidates, there remains a need for an efficient method that can detect various types of nodules with a relatively small number of false positives. Some researchers in recent years have focused on applying CNNs in false positive reduction in lung nodule detection, but not many of them have explored using of CNNs in generating lung nodule candidates. The aim of this work is to develop a framework based solely on CNNs for lung nodule detection. The lung nodule candidates are obtained using a 3D multi-scale fully convolutional neural network (FCN). The 3D CNN is used to distinguish the true nodules from other non-nodule candidates. To the best of our knowledge, this is a pioneer work that exploits 3D FCN for detecting lung nodule candidates; it also provides a general procedure for exclusively using CNNs for medical image detection tasks, besides lung nodule detection.

J. Shanghai Jiao Tong Univ. (Sci.), 2019, 24(4): 517-523

1 Method

The whole computer-aided lung nodule detection system is mainly divided into two stages. In the first stage, a 3D FCN is used to screen all possible lung nodules in the CT images. Excessive candidates are obtained in this stage to achieve a high sensitivity of lung nodules. The candidates are further classified in the second stage to get the final result.

1.1 Preprocessing and Lung Segmentation

The lung CT images used here have various slice thicknesses ranging from 0.6 to 2.5 mm and spatial resolutions ranging from 0.46 to 0.98 mm. The images are resampled to 1 mm isotropic voxel spacing. The pixel intensity is rescaled from $(-1\,000, 400\text{HU})$ to $(0, 1)^{[6]}$. The rough lung region is segmented on the basis of intensity and morphology. The region is then dilated 3 times by a disk kernel with a radius of 5 to ensure that all lung nodules especially juxta-pleural nodules are included in this region.

1.2 Lung Nodule Candidate Detection

The lung nodule candidates are obtained using a multi-scale 3D FCN. The inputs of the networks are the 3D CT image patches and corresponding lung nodule masks. A score map with the sizes equal to the patches can be obtained in the testing stage. The pixel value in the score map represents the possibility of being a lung nodule. The nodule candidates are obtained from the score maps.

1.2.1 Architecture of 3D FCN

A multi-scale 11-layer deep network originally developed by Kamnitsas et al.^[12] is adopted for brain lesion segmentation. There are 8 convolution layers, each followed by batch normalization^[13] and PRELU^[14]. The structure of the network is shown in Fig. 1, where the convolution kernels are described in the form of " $n_k \times s_k^3$ ", n_k is the number of kernels and s_k is the kernel size. In order to incorporate local and larger contextual information of the lung nodules, the network is designed to have two pathways for different image scales. The feature maps of the two pathways are concatenated after the second fully connected layer. More details about the analysis of the network can be found in Ref. [12].

1.2.2 Patch Extraction and Data Augmentation

The lung CT images are normalized to have zeromean and unit variance^[15]. The creation of nodule mask is based on the precise lung nodule boundary information provided in the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) database^[16-17]. In the nodule mask, the value of the lung nodule pixels annotated by at least 3 radiologists is set to be true while the value of other pixels is set to be false. As shown in Fig. 2, 25 pixel × 25 pixel × 25 pixel and 57 pixel × 57 pixel × 57 pixel patches are extracted from the same position of the



Fig. 1 Structure of 3D multi-scale FCN



Fig. 2 Structure of CNN

resampled images with 1 mm isotropic resolution. Here, the 25 pixel \times 25 pixel \times 25 pixel patches are used as the input of the first pathway while the 57 pixel \times 57 pixel patches are down-sampled by a factor of 3 and then used as the input of the second path way. The corresponding 9 pixel \times 9 pixel \times 9 pixel center part of the 25 pixel \times 25 pixel \times 25 pixel area is segmented from the nodule mask and used as the labels of the network. The positive samples and negative samples are randomly selected from the lung nodule regions and the lung regions respectively; equal numbers of positive samples and negative samples are used to prevent class-imbalance^[12].

1.2.3 Post Processing

In the test stage, the 3D FCN directly accepts 3D patches and outputs a 3D score map corresponding to the center part of the input patch. Sliding window strategy is used to get the whole score map. The value in the score map represents the possibility of being a nodule. An empirical set threshold of 0.25 is used in this map. Afterwards, regions smaller than 8 pixels are removed. The centers of the remaining connected components are calculated and regarded as the candidate positions of lung nodules.

1.3 False Positive Reduction

We extract and obtain 3D patches which are centered on the candidate positions obtained in the first stage. If the distance from a patch center to any true nodule center is less than the nodule radius, the label of the patch is set to 1; otherwise, this label is set to 0. The patches and labels are used to train a CNN. In the test stage, the CNN is used to classify all the lung nodule candidates.

1.3.1 Architecture of 3D CNN

The structure of the CNN is shown in Fig. 2. Rectified linear units (ReLUs) are used after every convolutional layer and the first fully connected layer^[18]. Dropout layers are adopted after the second and fourth convolution layers as well as after the first fully connected layer to prevent overfitting^[19].

1.3.2 Patch Extraction and Data Augmentation

Three geometric transformations (translation, rotation and flipping) are used for data augmentation. The center of the candidate can be translated by 1 mm in each axis or by 2 mm in one of three axes, which results in a total of 30 kinds of translations. Meanwhile, the samples can be rotated by 90°, 180° , 270° or flipped around one of the three axes. The total number of rotation and flipping combinations is 9. Under the combination of these three transformations, a total of 270 different combinations can be generated.

The number of the candidates is about 63 times of the number of true nodule samples. In order to prevent the prediction bias, the number of positive samples and the number of negative samples have to be balanced. In the training stage, transformations of pre-calculated types are randomly selected to ensure that the number of positive samples and the number of negative samples are equal, while in the testing stage all samples are augmented by 9 times.

2 Experiments

2.1 Dataset

We use the dataset released by the LUNA16 challenge. The data of the challenge comes from the LIDC-IDRI database which contains chest CT scans of 1010 patients. All the CT scans in the LIDC-IDRI database are read in a two-phase reading procedure by four experienced radiologists. The lung nodule annotations are independently determined by each radiologists in the first reading phase called blinded read. Then in the second reading phase, every radiologist can review the results of the other radiologists in the first reading phase and determines the final annotations of the nodules. The CT scans are acquired with different kinds of machines and protocols, and thus they have various slice thicknesses and spatial resolutions. In the challenge, 888 CT scans are selected by excluding scans with slice thickness larger than 2.5 mm or with inconsistent slice spacing^[6]. The nodules annotated by at least 3 radiologists and with diameters larger than 3 mm are regarded as true nodules while other nodules are considered as irrelevant findings.

2.2 Evaluation

The proposed method is evaluated by the detection sensitivity and false positives (FPs) per scan. A candidate is regarded as a true nodule, if the distance of its center to the nodule center is not larger than the approximate radius of the nodule. Free receiver operating characteristic (FROC) analysis is performed and a 95% confidence interval is computed using bootstrapping with 1 000 bootstraps^[20]. The competition performance metric (CPM) is also calculated^[21], which is defined as the average sensitivity at 7 predefined false positive rates: 1/8, 1/4, 1/2, 1, 2, 4, and 8 FPs per scan.

2.3 Experimental Setup

The FCN used in the first stage is implemented using Theano^[22-23]. In the training stage, the network is trained for 20 epochs, each of which consists of 20 subepochs. In each subepoch, 1 000 training samples of image patches are generated from 50 random CT scans. The configuration of the network is similar to that of the DeepMedic model^[12]. The learning rate is initialized as 0.001 and halved at 12, 15, 17 and 19 epochs. Besides, RMSProp is chosen as an optimizer and $L_1 = 10^{-6}$ and $L_2 = 10^{-4}$ are used as regularization constants^[24].

In the second stage, the CNN is implemented using Theano^[22-23] and Keras (https://keras.io). The learning rate is initialized as 0.01 and the weight decay is set to be 1×10^{-6} . The mini-batch gradient descent is used as an optimizer with the batch size set to 150. Training is stopped under the condition that the accuracy on the validation dataset does not improve for 4 epochs. The average prediction of the last 7 models is used as the final result.

2.4 Results

At the first stage, a sensitivity of 96.88% is achieved (1149/1186) with 84 FPs per scan. All the candidates are further classified in the second stage. The FROC of the whole stage is depicted in Fig. 3, where the dashed curves show the 95% confidence interval estimated by bootstrapping. The CPM score is 0.882. The sensitivity achieved at 2.2 FPs per scan is 93.93%. Figure 4 shows the distributions of the nodule sizes of detected and missed nodules both in the first and second stages. Figures 5 and 6 show some missed nodules in the first and second stages, respectively, where the nodule contours are annotated in red lines. Missed nodules in the second stage are defined as the nodules detected in the first stage and with a possibility of being a nodule less than 0.5 in the second stage. Missed nodules are mainly juxta-pleural nodules, juxta-vascular nodules, and nodules with irregular shape or non-solid nodules. Some false positives are shown in Fig. 7. Figure 8 shows some detected nodules, where the detected nodules are annotated in green lines while ground truth nodules are annotated in red lines.



Fig. 4 Distributions of nodule sizes of detected and missed nodules



Fig. 5 Some missing nodules in the first stage



Fig. 6 Some missing nodules in the second stage



Fig. 7 Samples of false positive detections annotated in green lines



Fig. 8 Samples of detected nodules and their corresponding ground truth nodules

3 Discussion

In the study, a lung nodule detection CAD system using 3D deep CNNs is developed. In the first of two stages, an FCN focuses on screening all possible lung nodules, and then a CNN in the second stage aims at reducing false positives. The FCN is used instead of a ruled-based screen method to detect different types of lung nodules. A detailed comparison with other methods for lung nodule candidate detection is shown in Table 1. Compared with other methods, the proposed method can achieve relatively high sensitivity with low FPs per scan in lung nodule detection stage. Meanwhile, it can also provide the contours of nodules. The successful application demonstrates the effectiveness of using FCN for lung nodule candidate detection. Table 2 lists the performances of some CAD schemes for lung nodule detection on the LIDC database. The performance of the proposed scheme is comparable with that of other methods in the literature.

There remains some future works to improve the system. Firstly, more experiments are needed to study how to determine the best input size, output size, the depth of the network and the ratio of positive and negative samples used in the FCN. Secondly, the complete lung nodule detection can be divided into two parts:

Source	Year	Number of scans	Number of nodules	Agreement level	${\rm Sensitivity}/\%$	FPs per scan
This paper	_	888	1 186	3	96.9	84
Ref. [25]	2017	502	698	3	94.0	102
Ref. [6]	2016	888	1186	3	93.3	269
Ref. [26]	2015	865	1147	3	78.0	43
Ref. [27]	2014	84	148	1	97.9	271
Ref. [4]	2011	125	126	3	96.8	458

 Table 1
 Comparison among different methods for lung nodule candidate detection on the LIDC datasets

 Table 2
 Comparison among different CAD schemes for lung nodule detection on the LIDC database

Source	Year	Number of scans	Number of nodules	Agreement level	${\rm Sensitivity}/\%$	FPs per scan
This paper	_	888	1 186	3	$93.6(95.5^*)$	4
Ref. [11]	2017	888	1186	3	$-(90.7^{*})$	4
Ref. [25]	2017	502	698	3	88.9	4
Ref. [6]	2016	888	1186	3	$87.9(90.1^*)$	4
Ref. [28]	2015	98	631	2	85.2	3.1
Ref. [27]	2014	84	148	1	97.5	6.8
Ref. [29]	2014	108	68	3	75.0	2
Ref. [30]	2013	84	103	1	80.0	4.2
Ref. [31]	2012	84	148	1	97.0	6.1
Ref. [4]	2011	125	80	4	87.5	4.0

* The nodules annotated by fewer than 3 radiologists are not regarded as false positives.

lung nodule candidate detection and false positive reduction. These two parts are not independent. We try to achieve an end-to-end lung nodule detection system. Thirdly, the system can only output the location of nodules. But in the real lung nodule screening, the follow-up requirements of pulmonary nodules with different types and sizes are different. In the future, the type, size and characteristic of the nodule can be analyzed to provide recommendations for follow-up.

4 Conclusion

A CAD system for lung nodule detection using 3D deep CNNs is proposed. An FCN is selected to screen lung nodule candidates, and then a CNN is developed to classify the candidates. Experimental results in the LUNA16 challenge demonstrate the effectiveness of the proposed method for lung nodule detection.

References

- SIEGEL R L, MILLER K D, JEMAL A. Cancer statistics, 2016 [J]. CA: A Cancer Journal for Clinicians, 2016, 66(1): 7-30.
- [2] WANG B, TIAN X D, WANG Q, et al. Pulmonary nodule detection in CT images based on shape constraint CV model [J]. *Medical Physics*, 2015, 42(3): 1241-1254.
- [3] LI Q, SONE S, DOI K. Selective enhancement filters for nodules, vessels, and airway walls in two- and

three-dimensional CT scans [J]. *Medical Physics*, 2003, **30**(8): 2040-2051.

- [4] TAN M, DEKLERCK R, JANSEN B, et al. A novel computer-aided lung nodule detection system for CT images [J]. *Medical Physics*, 2011, **38**(10): 5630-5645.
- [5] MESSAY T, HARDIE R C, ROGERS S K. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery [J]. *Medical Image Anal*ysis, 2010, 14(3): 390-406.
- [6] SETIO A A A, CIOMPI F, LITJENS G, et al. Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks [J]. *IEEE Transactions on Medical Imaging*, 2016, **35**(5): 1160-1169.
- [7] MURPHY K, VAN GINNEKEN B, SCHILHAM A M R, et al. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification [J]. *Medical Image Analysis*, 2009, **13**(5): 757-770.
- [8] JACOBS C, VAN RIKXOORT E M, TWELLMANN T, et al. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images [J]. *Medical Image Analysis*, 2014, 18(2): 374-384.
- [9] SETIO A A A, JACOBS C, GELDERBLOM J, et al. Automatic detection of large pulmonary solid nodules in thoracic CT images [J]. *Medical Physics*, 2015, 42(10): 5642-5653.
- [10] CIOMPI F, JACOBS C, SCHOLTEN E T, et al. Bagof-frequencies: A descriptor of pulmonary nodules in computed tomography images [J]. *IEEE Transactions* on Medical Imaging, 2015, **34**(4): 962-973.

- [11] DOU Q, CHEN H, YU L Q, et al. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection [J]. *IEEE Transactions on Biomedical Engineering*, 2017, 64(7): 1558-1567
- [12] KAMNITSAS K, LEDIG C, NEWCOMBE V F J, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation [J]. *Medical Image Analysis*, 2017, **36**: 61-78.
- [13] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on Machine Learning. Lille, France: JMLR, 2015, 37: 448-456.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification [C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1026-1034.
- [15] JARRETT K, KAVUKCUOGLU K, RANZATO M, et al. What is the best multi-stage architecture for object recognition? [C]// 2009 IEEE 12th International Conference on Computer Vision (ICCV). Kyoto, Japan: IEEE, 2009: 2146-2153.
- [16] ARMATO III S G, MCLENNAN G, BIDAUT L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans [J]. *Medical Physics*, 2011, **38**(2): 915-931
- [17] CLARK K, VENDT B, SMITH K, et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository [J]. Journal of Digital Imaging, 2013, 26(6): 1045-1057.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1097-1105.
- [19] SRIVASTAVA N, HINTON G E, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [20] EFRON B, TIBSHIRANI R J. An introduction to the bootstrap [M]. Boca Raton, USA: CRC press, 1994.
- [21] NIEMEIJER M, LOOG M, ABRÀMOFF M D, et al. On combining computer-aided detection systems [J].

IEEE Transactions on Medical Imaging, 2011, **30**(2): 215-223.

- [22] BASTIEN F, LAMBLIN P, PASCANU R, et al. Theano: New features and speed improvements [C]//Proceedings of the Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop. Nevada, USA: NIPS, 2012: 1-9.
- [23] BERGSTRA J, BREULEUX O, BASTIEN F, et al. Theano: A CPU and GPU math compiler in Python [C]//Proceedings of the 9th Python in Science Conference. Austin, USA: [s.n.], 2010: 3-10.
- [24] TIELEMAN T, HINTON G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude [J]. COURSERA: Neural Networks for Machine Learning, 2012, 4(2): 26-31.
- [25] MA J C, ZHOU Z E, REN Y C, et al. Computerized detection of lung nodules through radiomics [J]. *Medical Physics*, 2017, 44(8): 4148-4158.
- [26] VAN GINNEKEN B, SETIO A A A, JACOBS C, et al. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans [C]//2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). New York, USA: IEEE, 2015: 286-289.
- [27] CHOI W J, CHOI T S. Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor [J]. Computer Methods and Programs in Biomedicine, 2014, 113(1): 37-54.
- [28] LU L, TAN Y Q, SCHWARTZ L H, et al. Hybrid detection of lung nodules on CT scan images [J]. Medical Physics, 2015, 42(9): 5042-5054.
- [29] BROWN M S, LO P, GOLDIN J G, et al. Toward clinically usable CAD for lung cancer screening with computed tomography [J]. *European Radiology*, 2014, 24(11): 2719-2728.
- [30] TERAMOTO A, FUJITA H. Fast lung nodule detection in chest CT images using cylindrical noduleenhancement filter [J]. International Journal of Computer Assisted Radiology and Surgery, 2013, 8(2): 193-205.
- [31] CASCIO D, MAGRO R, FAUCI F, et al. Automatic detection of lung nodules in CT datasets based on stable 3D mass-spring models [J]. Computers in Biology and Medicine, 2012, 42(11): 1098-1109.