Multimed Tools Appl (2019) 78:1017–1033
https://doi.org/10.1007/s11042-018-6082-6

A deep CNN based transfer learning method for false positive reduction

Zhenghao Shi[1] · **Huan Hao**[1] · **Minghua Zhao**[1] · **Yaning Feng**[1] · **Lifeng He**[2] · **Yinghui Wang**[1] · **Kenji Suzuki**[3,4]

Received: 5 March 2018 / Revised: 26 April 2018 / Accepted: 30 April 2018 /
Published online: 8 May 2018

Abstract A low false positive (FP) rate is of great importance for the use of a Computer Aided Detection (CAD) system to detect pulmonary nodules in thoracic Computed Tomography (CT). However, due to the variations of nodules in appear and size, it is still a very challenging task to obtain a low FP rate. In this paper, we propose a deep Convolutional Neural Network (CNN) based transfer learning method for FP reduction in pulmonary nodule detection on CT slices. We utilized one of the state-of-the-art CNN models, VGG-16 [4], as a feature extractor to obtain nodule features, and used a support vector machine (SVM) for nodule classification. Firstly we transferred all the layers from a pre-trained VGG-16 model in ImageNet to our target networks. Then, we tuned the last fully connected layers to adjust the computer-vision-task-trained CNN model to pulmonary nodule classification task. The initial CNN filter weights were then optimized using the training data, i.e., the pulmonary nodule patch images and corresponding labels through back-propagation so that they better reflected the modalities in the pulmonary nodule image dataset. Finally, features learned in the fine-tuned CNN were used to train a SVM classifier. The output of the trained SVM was

✉ Zhenghao Shi
ylshi@xaut.edu.cn

✉ Lifeng He
helifeng@ist.aichi-pu.ac.jp

✉ Kenji Suzuki
ksuzuki@iit.edu

[1] School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China

[2] School of Information Science and Technology, Aichi Prefectural University, Nagakute, Aichi 4801198, Japan

[3] Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA

[4] World Research Hub Initiative, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, Kanagawa 226-8503, Japan

used for final classification. Experimental results show that the overall sensitivity of the proposed method was 87.2% with 0.39 FPs per scan, which is higher than 85.4% with 4 FPs per scan obtained by other state of art method.

Keywords False positive reduction · Nodule detection · Deep convolutional network · Support vector machine

1 Introduction

Lung cancer is the leading cause of cancer-related death, and the number of death due to lung cancer is greater than the combined number of prostate, colon, and breast cancer deaths [22]. Early detection of lung cancer is a helpful way to improve the survival rate of lung cancer patients. It is reported that the five-year relative survival rate of lung cancer diagnosed in advanced stages is only 16%, whereas that diagnosed in early stages increases to 70% [14]. Typically, lung cancer in its early stage manifests itself in the form of pulmonary nodules which usually refer to lung tissue abnormalities that are roughly spherical with round opacity on a CT (Computed Tomography) scan [3]. However, due to the large number of image slices at a CT scan, diagnostic reading on CT scans in a short time is a huge workload upon radiologists, and prone to errors. To assist radiologists with this challenge, various Computer-Aided Detection (CAD) systems have been developed in last recent years [3, 13, 30, 36, 38, 39].

The architecture of a conventional CAD system typically consists of two stages [2]: candidate nodules detection and false positives (FPs) reduction. In the first stage, a large number of candidate regions are generated using intensity-based imaging features upon prior assumptions about the expected morphology of the nodules. In the second stage, FPs resulting from the first stage are reduced using a classifier [26]. This stage is very important for the successful use of CAD in clinical practices. If there remains too many FPs, the performance of detecting true positive regions may potentially be degraded, and thereby may also be a subsequent increase in “unnecessary” follow-up examinations. In an attempt to reduce FPs as much as possible, hand-crafted features, such as Scale Invariant Feature Transform (SIFT) [4], Local Binary Patterns (LBP) [24] and Histogram of Oriented Gradients (HOG) [23], are extracted from the candidate regions and used to train a classifier for nodule classification. However, under such a condition, the performance of a CAD system relies heavily on the intermediate results of the image processing tasks for reliable features [3]. Whereas the selection, integration and optimization of hand-crafted features, are not an easy task, it is also another important issues in image classification and also has an important effect on the performance of a CAD system.

Recently, remarkable progresses have been obtained in the field of computer vision due to the recent revival of deep convolutional neural networks (CNN) [10, 12] and the availability of large-scale annotated datasets (i.e., ImageNet [7, 20]). Compared to the conventional machine learning methods, deep learning requires no hand-tuned feature extractor, and learns imaging representations directly from large volumes of data [25, 27, 28]. Accordingly, the issues of feature computing, selection, and integration can potentially be addressed by this new learning framework without a complicated pipeline of image processing and pattern recognition steps. Inspired by the success of deep CNNs in computer vision, increasing interesting of applying deep CNNs for medical image processing and analysis has been seeing in recent short time [1, 29, 33, 34].

However, before such successes can be transferred to benefit the medical imaging community, there are significant hurdles. First, to achieve state-of-the-art performance, a deep CNN network requires large volumes of labeled training data, whereas collecting and annotating large numbers of medical images of a specific modality still poses significant challenges. Second, there are significant variations in the appearance of medical images based on the individual diseases depicted.

Aiming at issues mentioned above, we propose a deep CNN based transfer learning method for FP reduction for pulmonary nodule detection on CT slices in this work. We utilize VGG-16, one pre-trained state-of-the-art CNN model on ImageNet, as a feature extractor to obtain nodule features, and uses a support vector machine (SVM) for nodule classification. The use of pre-trained CNN is based on transfer learning [23]. It refers to the procedure employed to train a source model and then transfer the knowledge across different problems. Such a technique is expected to utilize the advantages of CNN model on feature learning and SVM mode on efficient non-linear classification. Performance of the proposed system is evaluated on the Lung Image Database Consortium database. It achieved 87.2% sensitivity with only 0.39 FPs per scan. Experimental results demonstrate the promising effectiveness of this method compared with other state-of-the-art approaches.

The main contributions of the paper are summarized below:

(1) We transfer an ImageNet-trained CNN network for learning the representation of pulmonary nodules in CT, which resolves the challenges associated with using CNN for image feature extraction problems with limited and unevenly distributed sample data.
(2) We propose a FP reduction method based on the combination of pre-trained CNN network and an SVM for nodule detection in CT slices. Such a model is expected to utilize the advantages of deep CNN on feature learning and the SVM on efficient non-linear classification. The use of pre-trained CNN based on transfer learning is expected to extract image features at different semantic levels, as well as to resolve the challenges associated with using CNNs on classification problems with limited and unevenly distributed sample data.

The remainder of the paper is organized as follows. Section 2 analyzes the related works. The methodology for pulmonary nodule detection is described in Section 3. The experimental results obtained are given in Section 4. Then a discussion of the work is followed. After that the complexity of the proposed method is discussed. We conclude this paper in Section 7.

2 Related work

Along with the tremendous success of deep learning in natural image processing and computer vision, deep learning techniques have also witnessed an increasing interesting for medical image processing with promising results [41]. Examples include automatic polyp detection in colonoscopy videos [6, 35], polyp detection in CT colonography [33], computer-aided detection of pulmonary embolism in CT datasets [5, 17, 32], automatic detection of mitotic cells in histopathology images [9], computer-aided detection of lymph nodes in CT images [18], and computer-aided anatomy detection in CT volumes [42].

Other recent studies show the potential for knowledge transfer from natural images to the medical imaging domain using pre-trained deep CNNs [40]. For instance, in [1], pre-trained CNNs were used as a feature generator for chest pathology identification. In [5], the authors applied a pre-trained CNN for classification of interstitial lung diseases. In [17], pre-trained CNNs were used to retrieve missing cardiac acquisition plane information from magnetic resonance imaging. However, due to lack of fine-tuning, these pre-trained CNNs are more reflective of the natural image dataset and may not necessarily reflect the subtle characteristics of medical images. Fine-tuning is the process of updating the pre-trained weights of CNN through the use of back propagation.

Different from existing methods which only use the pre-trained model, fine-tune pre-trained deep CNN is used in our work. In contrast to aforementioned pre-trained CNNs based methods, our work is more computationally effective in pulmonary nodule features extraction. By transfer learning, great benefits to medical imaging community could be achieved by the transfer of image representation ability learned by deep models to other generic vision tasks, for example, pre-trained state-of-the-art CNN model on ImageNet. In such a case, we only need to prepare a limited numbers of medical images of a specific modality.

3 Proposed method

3.1 Overview

Figure 1 shows the overview of the proposed method. It begins from copying all the layers from VGG-16 to our target networks. Then we modified the last fully connected layers for adapting the CNN model to pulmonary nodule classification task by replacing the last fully connected layer (intended for 1000 classes) with a new fully connected layer for the 2 classes in our dataset. Next, the initial CNN filter weights derived from the natural images were then fine-tuned (optimized) using the training data, i.e., the pulmonary nodule patch images and corresponding labels through back-propagation so that they better reflected the modalities in the pulmonary nodule image dataset. For the training of SVM, features extracted from the fine-tuned CNN are used. The SVM classifier is trained using features extracted from one CNN model. In the stage of nodule detection, our approach first extracts CT patches from CT slices. The obtained patches are then fed into the fine-tuned ensemble of deep CNN networks to compute discriminative features. Next, the computed discriminative features from the fine-tuned deep CNN network are fed into SVM classifier to label the input candidates as nodules or non-nodules. Finally, Outputs of the SVM are the detection result. Details of the proposed CAD system are described in the following sections.

3.2 Pre-trained CNN model—VGG-16

A CNN is a function to map input data (in our case, an input data is an image.) to an output, and it generally consists of convolutional layers, max or sum pooling layers, activation layers (e.g., Rectified Linear Unit (ReLU) or sigmoid activation layers), and a softmax layer which generates a well-formed probability distribution for classification on the outputs. If input data is images, typically, the first convolutional layer extracts edges, and subsequent convolutional layers act as higher-level features exactors. A pooling layer

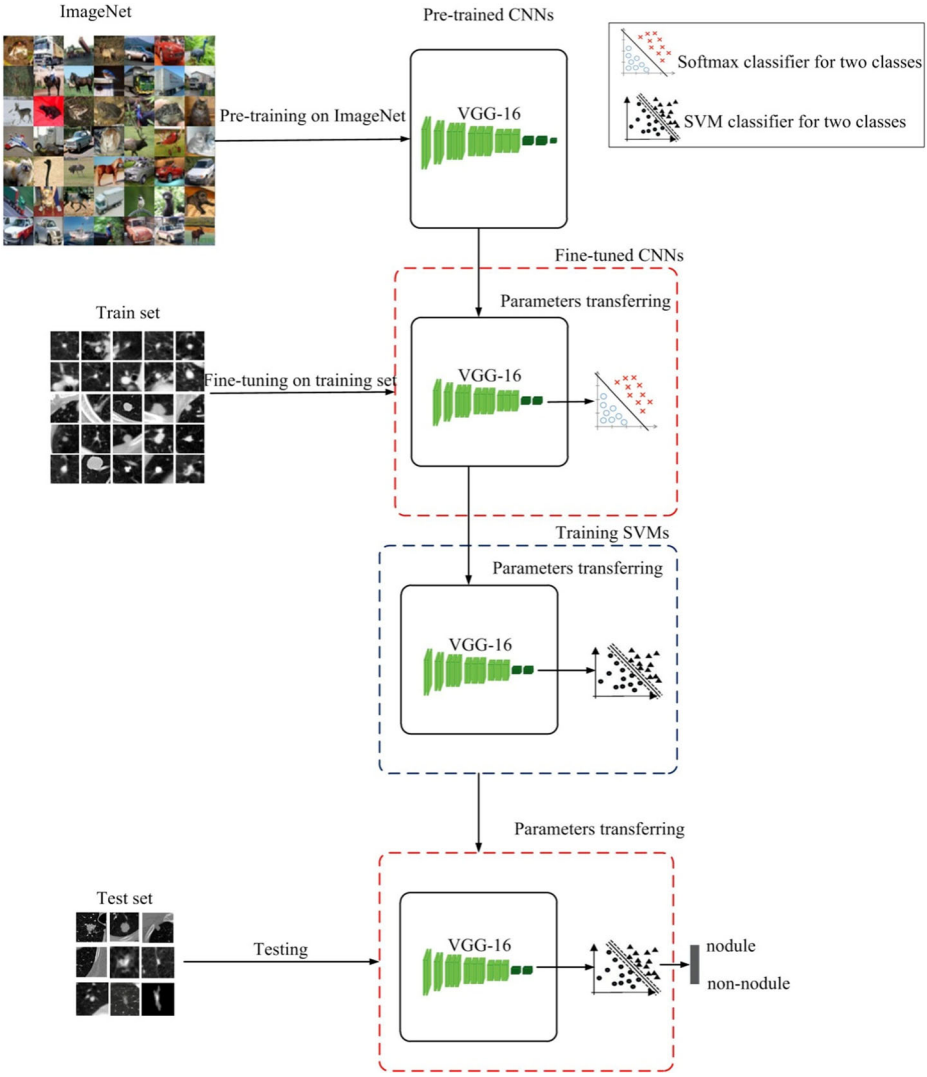


Fig. 1 An overview of the proposed method

is used to compute the maximum (e.g. the max pooling layer) or the average (e.g. the average pooling layer) of each patch of the feature map from the previous layer. The role of the activation layer, e.g. the Rectified Linear Units, is to address the issue of saturation. Mathematically, a CNN can be represented as a function 핏, which is the composition of a sequence of functions, i.e.,

$$\mathrm{F} = f_L f_{L-1} f_{L-2} f \ \ldots\ldots\ \ldots . f_2 f_1 \tag{1}$$

Each function 핏 represents a layer which takes the output of the previous layer.

After training, a CNN can be used as a classifier or as a feature extractor in the case of transfer learning. Transfer learning is a technique used to improve the performance of machine learning by harnessing the knowledge obtained by another task [1].

In this work, the pre-trained CNN model VGG-16 is referred. We choose the architecture because it is well established and has shown better performance in our image classification than other pre-trained models, such as AlexNet [10] and GoogLeNet [31].

VGG-16 has a similar architecture with AlexNet with more convolutional layers. It has 13 convolutional layers followed by rectification and pooling layers, and 3 fully connected layers. All convolutional layers use small 3×3 filters and the network performs only 2×2 pooling. VGG-16 has a receptive field of size 224 × 224. Although VGG-16 performs better than AlexNet and has a simpler architectural design, it has more parameters which needs more computation.

3.3 Parameters transferring and fine-tuning

As mentioned above, in this work, the state-of-the-art CNN network, VGG-16, pre-trained on ImageNet is transferred to learn lung nodule feature representations from the labeled training data. For the network architecture, firstly we copied all the layers from VGG-16 model to our target network except the last fully connected layer. Then we modified the last fully connected layer for adapting the CNN model to pulmonary nodule classification task by replacing the last fully connected layer (intended for 1000 classes) with a new fully connected layer for the 2 classes in our dataset. Next, the initial CNN filter weights derived from the natural images were then fine-tuned (optimized) using the training data, i.e., the pulmonary nodule patch images and corresponding labels through back-propagation so that they better reflected the modalities in the pulmonary nodule image dataset.

Given the i^{th} input image of nodules x_i and its corresponding label y, for m pairs of training data (x_i,y_i) for $i = 1,2,\ldots,m$, The classification loss function can be defined as

$$1(f(x_i\mathrm{w}), y_i) = \begin{cases} 1 & argmaxf(x_i\mathrm{w}) \neq y_i \\ 0 & others \end{cases} \tag{2}$$

To minimize the loss function, we use the stochastic gradient descent (SGD) method which calculates a random subset of training examples to estimate the mean gradient for all the training examples.

In general, the early layers of a CNN learn low level image features, which are applicable to most vision tasks, but the late layers learn high-level features, which are specific to the application at hand. Therefore, fine-tuning the last few layers is usually sufficient for transfer learning. However, because distance between images in imageNet and pulmonary nodules is significant, therefor in our work, we have fine-tuned the early layers as well. We start from the last layer and then incrementally include more layers in the updating process until the desired performance is reached.

3.4 Training SVM classifiers

For FP reduction, SVM classifier is employed in our work. The SVM classifier is trained using features extracted from the fine-tuned VGG-16. From VGG-16, we extracted 4096-dimensional output of the first fully connected layer after the rectified linear units (ReLU). Using the 4096-dimensional feature vector from VGG-16, we train the linear SVM classifier.

Let x_i be a training instance, and the training set is denoted as $X = (x_1, \ldots, x_N) \in R^{P \times N}$, where N and P are the number of trining instances and dimensionality of X, respectively. Let $y = (y_1, \ldots, y_N)^T$ be the label vector, with $y_i \in \{+1, -1\}$ being the class label of the i^{th} training

instance (in this paper, where+1 corresponds to a nodule and − 1to a non-nodule), the primal objective function of SVM can be defined as

$$\min_{\omega} \quad \frac{1}{2}\omega + c\sum_{i=1}^{N}\xi_i \qquad s.t\ y^n\left(\omega^T X_i + b\right) + \xi_i \geq 1 \xi_i \geq 0, \forall i \tag{3}$$

where ω is the normal of separating hyperplane, b is the bias term, $\xi = [\xi_1, \cdots, \xi_N]$ is the vector of slack variables, and C is the regularization parameter. The optimal C can be obtained by cross validation.

After optimization, SVM learned weighting $\alpha_i \in R^N$ on training samples in kernel feature space $\phi(x_i)$, where $\phi : \mathcal{X} \rightarrow \mathcal{H}$ defines a feature mapping from the original input space to a Hilbert space $\mathcal{H}$. Nodules are then separated from non-nodules by the maximal margin hyper plane in the feature space. The decision function can then be represented as

$$f(x) = \mathrm{sign}\left(\sum_{i=1}^{N} y_i a_i^* K(x, x_i) + b^*\right) \tag{4}$$

where $K(x, x') = \phi(x) \cdot \phi(x')$ is a kernel function used to assess the similarity of input samples, parameter a' and b' are obtained by solving the dual of the optimization in (1).

For final nodule classification, we used an ensemble of fine-tuned CNN and SVM. The ensemble consists of the fine-tuned VGG-16 and the SVM classifier trained with features extracted from the fine-tuned CNN. Outputs of the SVM are used as the finally output.

4 Experiments

4.1 Dataset

The dataset we used was obtained from the LIDC-IDRI database [30] which includes 1018 cases from seven institutions, and each case consists of images from a thoracic CT scan, as well as the annotations provided by four radiologists. We selected nodules greater than 3 mm as our experimental data which resulted in a set of 233 scans including 700 nodules. The ROI of nodules were extracted by an adaptive threshold segmentation method, moreover, we choice an appropriate number of vessels and other tissues as non-nodule images. Then we cropped each image into a 224×224 image around the center of ROI to obtain the final database. Our database consists of 1400 images, including 700 nodule images and 700 non-nodule images. Each slice has an image matrix of 224×224 (16 bits depth) pixels..

4.2 Experiment setting

In our experiments, for the fining-tune of VGG-16, the mini-batch stochastic gradient descent with a momentum factor of 0.9 is employed. The base learning rate is set as 0.001. A batch size of 20 is used due to large memory requirements for VGG-16. The test scheme is designed as two different strategies. One is 10-fold cross-validation and the other is that the dataset is divided into the training data (80%) and testing data (20%). Since all the previous works are based on the manually designed features while the proposed approach in this paper is based on feature learning and nodule recognition by deep convolutional neural networks, it is not possible to directly compare our method with them on the same LIDC dataset. All experiments are conducted on a desktop computer with Intel Core 2

CPU of 2.80GHz,8GB memory, and Windows 7.The algorithm is implemented using the MatConvNet [37] on Matlab2016.

To evaluate the performance of our proposed method, we present several experiments by compare it with some state-of-the-art methods, including detection of nodules by artificial features [15, 16, 21] and deep features [9, 18, 32] in terms of accuracy, sensitivity, and FP.

4.3 Results

In order to show the performance of the proposed method, we evaluated it by 10-fold cross-validation. Table 1 shows this experiment result. From Table 1, it can be seen that the transferred VGG-16 obtain a promising performance on pulmonary nodule recognition on CT images, where the average classification accuracy, sensitivity, specificity and FP/s of proposed model are 87.2%, 87.2%, 87.2% and 0.39.

We also compared the proposed method with the state-of-the-art methods designed for lung nodule detection with 10-fold cross validation, including method proposed in [21] where 4 features such as mean, contrast, entropy and standard deviation are manual extracted from the images to train a SVM classifier, method proposed in [16] where 16 different types of features are selected as the input of regression tree, method proposed in [15] where 22 basic texture and shape features to train a SVM classifier, method proposed in [11] where a 200 dimensional deep feature vector obtained from the layer 4 of their auto encoder architectures, method proposed in [19] where a Multi-View CNNs is used to extract deep features, and method proposed in [8] where a Multi-Group Patch-Based method is used to crop nodules patches and a two layer CNN is employed to extracted deep features.

From Table 2, we can see that the method in [21] achieved an overall accuracy of 83.6% with a sensitivity of 82.1%, the method in [16] achieved an overall accuracy of 84.8% with a sensitivity of 85.2% and a false positive of 3.13 per scan, the method in [15] achieved an overall accuracy of 84.1% with a sensitivity of 81.39% and a false positive of 4.5 per scan, the method in [11] achieved an overall accuracy of 75.01% with a sensitivity of 83.25% and a false positive of 0.39 per scan, the method in [19] achieved an overall accuracy of 80.1% with a sensitivity of 85.4% and a false positive of 4 per scan, and the method proposed in [8] achieved an overall accuracy of 80.06% with a sensitivity of 84% and a false positive of 4.7 per scan. While our proposed method achieved an overall accuracy of

Table 1 Evaluation of the fine-tuned VGG-16 CNN model and SVM classification results from 10-fold cross validation experiments

	Accuracy (%)	Sensitivity (%)	Specificity (%)	FP/scan
1	82.0	88.0	76.0	0.72
2	89.0	90.0	88.0	0.36
3	89.0	92.0	86.0	0.42
4	88.0	88.0	88.0	0.36
5	88.0	88.0	88.0	0.36
6	82.0	80.0	84.0	0.48
7	92.0	90.0	94.0	0.18
8	92.0	86.0	98.0	0.06
9	80.0	80.0	80.0	0.63
10	90.0	90.0	90.0	0.3
Avg	87.2	88.2	87.2	0.39

87.8% with a sensitivity of 88.2% and a false positive of 0.23 per scan, which has a better sensitivity and a lower FP/s. Our proposed method is an automatic way for feature extraction, compared with the artificial way of feature extraction [15, 16, 21], our method realizes the automation of feature extraction, and achieves higher accuracy, sensitivity, meanwhile, compared with the artificial way of feature extraction [8, 11, 19], our proposed CAD system has the highest accuracy.

5 Discussion

In this study, a novel pulmonary nodule detection CAD system using deep VGG-16 CNN model based on transfer learning is proposed. Compared to published CAD systems that are evaluated on the publicly available LIDC-IDRI dataset, our proposed CAD system achieves comparable or better performance, indicating the potential of using pre-trained deep CNN on ImageNet instead of using engineered features and classification as the FP reduction stage. It suggests that the possibility of learning features from large natural image training dataset allows the network to learn classifying objects with a high degree of variation, which is suitable for the problem of pulmonary nodule detection.

For deep CNN based image classification, the common used classifier is the Softmax function for its easy implementation. However, in this work, we used an SVM as a classifier for FPs reduction, Why? In fact, the Softmax is just a generalization of the logistic function, and by itself it is not a classifier. It sort of squashes values in a k-dimensional vector to be in the range (0,1)(0,1) and sum total of up to 1.0, which gives a probabilistic interpretation. Therefore, the Softmax is useful as a cost function for multi-class category detection. Whereas, an SVM on the other hand is a classifier with a hinge-loss cost function that results in a maximum margin hyperplane. This can be extended to non-linearly separable problems using kernel approaches by mapping the data into higher dimensional spaces. Based on what mentioned above, we select SVM as a classifier in our work.

We evaluated the performance of three different classifiers, including SVM, the Softmax, and KNN, using testing data. The results from the classifiers evaluation is shown in Table 3. As it can be seen that with SVM, an overall accuracy of 91.5% with a sensitivity of 90.1% and a FP rate of 0.23 was achieved, whereas with the Softmax, the overall accuracy is 80.5% with a sensitivity of 91.0% and a FP rate of 0.91, and with KNN,

Table 2 Comparison of CAD systems for nodule detection

Methods	Acc	Sen	FP/s [a]
Sivakumar, S et al. [21]	83.6%	Sen:82.1%	–
Lu L et al. [16]	84.8%	85.2%	3.13
Liu J-K et all. [15]	84.1%	81.3%	4.5
Kumar D et all. [11]	75.0%	83.35%	0.39
A. Setio et all. [19]	80.1%	85.4%	4
Jiang H et all. [8]	80.06%	84%	4.7
Proposed Method	87.8%	88.2%	0.32

[a] Acc is the accurary. Sen is the sensitivity. FP/s is the false positives per scan.[a] Acc is the accurary. Sen is the sensitivity. FP/s is the false positives per scan

Table 3 Comparison of the performance of classifiers with features extracted by the fine-tuned VGG-16

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	FP rate
SVM	91.5	90.5	92.5	0.23(15/66)
SoftMax	80.5	91.0	70.0	0.91(60/66)
KNN	83.5	68.0	99.0	0.03(2/66)

the overall accuracy is 83.5% with a sensitivity of 68.0% and a FP rate of 0.03. Noteworthy, the performance of SVM is higher than that the other two classifiers in accuracy.

We also directly applied the VGG-16 for lung nodule features extraction without fine-tuning, and a compare of the experiment results with that of the fine-tuned VGG-16 with 10-fold cross validation is shown in Table 4. From Table 4, it can be seen that in the case of directly applying the VGG-16 for lung nodule features extraction without fine-tuning only achieved an accuracy of 79.5% with a sensitivity of 87.57% and a FP rate of 0.39. Whereas with the fine-tuned VGG-16, an accuracy of 87.8% with a sensitivity of 88.2% and a FP rate of 0.32.

This preliminary work examines the effectiveness of using pre-trained deep CNN networks to distinguish nodules and non-nodules. This study can be extended to various classification tasks such as parenchyma mass and polyps and can be used hierarchically with models trained to localize nodule candidates and quantify. Some reasons for the success of the classification include the automatic extraction of features rather than an artificial calculation way [15, 16, 21]. In practice, it is impossible to extract effective features without knowing the real meaning of nodule features by CNN models. For that reason, we used a CNN model in our work for feature extraction. Also, the state-of-the-art deep CNNs performs well in image classification tasks, however, limited by the size of image library, it is hard to train a very deep CNN model for nodule detection. For that reason, we used a fine-tuned VGG model in our work for feature extraction which applied the state-of-the-art deep CNNs to nodule detection tasks. It can be concluded that Deep Learning using Convolutional Neural Networks would be a good option for focal lesion classification and the use of pre-training CNNs would be one of the best choices to achieve one of the best results.

6 Computation complexity

Now we analyze the computation complexity of the proposed method, which consists of the time complexity and the space complexity. The time complexity determines the training and test time of a method, and the space complexity determines the number of the parameters of a method. Firstly, we analyze the time complexity of the proposed method. The forward-passing through each layer mainly consists of two procedures:

Table 4 Comparing experiment with fine-tuning and without fine-tuning

Method	Accuracy (%)	Sensitivity (%)	FPs rate
VGG-16 with SVM	79.5	87.57	0.39
Fine-tuned VGG-16 with softmax	87.8	88.2	0.32

pre-computation of inner products, and approximate computation of layer response. For each convolution layer, the time complexity of the forward-passing is

$$\text{Time}\sim\text{O}\left(M^2\cdot K^2\cdot C_{in}\cdot C_{out}\right) \tag{5}$$

where

$$\text{M} = \frac{\text{X}-\text{K}+2*\text{Padding}}{\text{Stride}}+1 \tag{6}$$

M is the size of the output feature map of each convolution kernel, K is the size of each convolution kernel. For VGG-16, K = 3, C_{in} is the channel number of each convolution kernel, also the output channel number of a pre-layer. C_{out} is number of convolution kernel of the current layer. X is the size of input image. Here, we assume that all input image patches and convolution kernels are squares.

The proposed method is based on the VGG-16 model. The VGG-16 model has 16 weight layers, including 13 Convolution layers and 3 fully connected layers. Therefore for the proposed method, the time complexity of the forward-passing is

$$\text{Time}\sim\text{O}\left(\sum_{l=1}^{l=16} M_l^2\cdot K_l^2\cdot C_{l-1}\cdot C_l\right) \tag{7}$$

where d is the layer number of the VGG-16, also termed as depth. l is the l^{th} convolution layer, C_l is the number of output channel of the l^{th} convolution layer,

From what mentioned above, it can be seen that the time complexity of the proposed method largely depends on two parameters, M (the size of the output feature map of each convolution kernel) and K (the size of each convolution kernel). Large values of M and K will lead to more fine-grained quantization, but is less efficient in the computation and storage consumption. In practice, we usually vary these two parameters to balance the tradeoff between the test-phase efficiency and accuracy loss of the quantized CNN model.

About the space complexity of the proposed method, we defined it as

$$\text{Space}\sim\text{O}\left(\text{d} = \sum_{\text{l}=1}^{\text{d}=16}\text{K}_\text{l}^2\cdot\text{C}_{\text{l}-1}\cdot\text{C}_\text{l}\right) \tag{8}$$

7 Conclusion

Focusing on improving the performance of FP reduction, we investigated and presented a transfer learning method for lung nodule detection based on pre-trained deep CNN model and SVM. Experiments demonstrate the effectiveness of the proposed method in reducing FPs. The proposed methods could be applied for the detection of many other potential lesions, such as mass and polyp. In future, further investigations include evaluating it on more clinical data and promoting it in clinical practice with the aid of radiologists and surgeons.

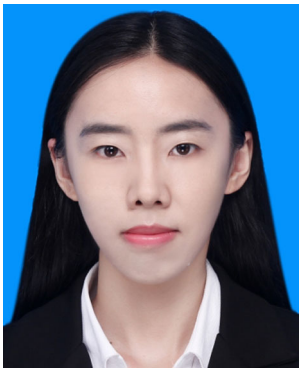
Acknowledgments This work was supported in part by a grant from the National Natural Science Foundation of China (No. 61202198, No.61401355), a grant from the China Scholarship Council (No.201608610048) and the Nature Science Foundation of Science Department of PeiLin count at Xi'an(GX1619), the Key Laboratory Foundation of Shaanxi Education Department, China (No.14JS072). The authors gratefully acknowledge the helpful comments and suggestions of the reviewers.

References

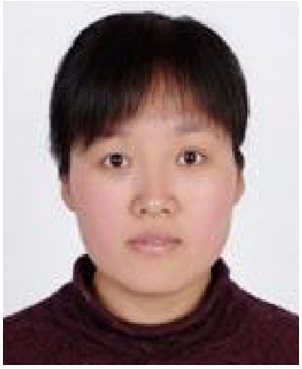
1. Bar Y, Diamant I, Wolf L, and Greenspan H (2015) Deep learning with non-medical training used for chest pathology identification, in Proc. SPIE Med Imag, 94 140V–94 140V
2. Camarlinghi N (2013) Automatic detection of lung nodules in computed tomography images: training and validation of algorithms using public research databases. Eur Phys JPlus 128(9):21
3. Dhara AK, Mukhopadhyay S, Khandelwal N (2012) Computer-aided detection and analysis of pulmonary nodule from CT images: a survey. IETE Techn Rev 29(4):265–275
4. Farag A, Ali A, Graham J, Farag A, Elshazly S, Falk R (2011) Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest. In: 2011 I.E. international symposium on biomedical imaging: from nano to macro. IEEE, 169–172
5. Gao M et al. (2015) Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. In: 1st Workshop Deep Learn. Med. Image Anal., Int. Conf. Med. Image Comput. Comput. Assist. Intervent., Munich, Germany, [Online]. Available: www.research.rutgers.edu/~minggao/files/MingchenGao_MICCAIworkshop2015.pdf
6. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 I.E. Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 580–587
7. Girshick R, Donahue J, Darrell T, Malik J (2016) Region-based convolutional networks for accurate object detection and semantic segmentation. IEEE Trans Pattern Anal Mach Intell 38(1):142–158
8. Jiang H, Ma H, Qian W, Gao M, Li Y (2017) An automatic detection system of lung nodule based on multi-group patch-based deep learning network. IEEE J Biomed Health Inform. https://doi.org/10.1109/jbhi.2017.2725903
9. Khan A, ElDaly H, Rajpoot N (2013) A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. J Pathol Inf 4(1)
10. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst:1097–1105
11. Kumar D, Wong A, Clausi DA (2015) Lung nodule classification using deep features in CT images[C]// conference on computer and robot vision. IEEE Comput Soc 133–138
12. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444
13. Li F, Arimura H, Suzuki K, Shiraishi J, Li Q, Abe H, Engelmann R, Sone S, MacMahon H, Doi K (2005) Computer-aided detection of peripheral lung cancers missed at CT: ROC analyses without and with localization. Radiology 237:684–690
14. Liang MZ, Tang W, Xu DM, Jirapatnakul AC, Reeves AP, Henschke CI, Yankelevitz D (2016) Low-dose CT screening for lung Cancer: computer-aided detection of missed lung cancers. Radiology 281(1):279–288
15. Liu J-K et al (2017) An Assisted Diagnosis System for Detection of Early Pulmonary Nodule in Computed Tomography Images. J Med Syst 41:30
16. Lu L, Tan Y, Schwartz LH et al (2015) Hybrid detection of lung nodules on CT scan images. Med Phys 42(9):5042–5054 20
17. Margeta J, Criminisi A, Lozoya RC, Lee DC, and Ayache N (2015) Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition. Comput Methods Biomech Biomed Eng, Imag Vis 1–11
18. Roth H et al. (2014) A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: proc. MICCAI, P. Goll, N. Hata, C. Barillot, J. Hornegger, and R. Howe, (eds.), vol. 8673, LNCS, pp. 520–527,
19. Setio A et al (2016) Pulmonary nodule detection in CT images using multi-view convolutional networks. IEEE Trans Med Imag 35(5):1160–1169
20. Simonyan K, Zisserman A (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition ArXiv, [Online]. Available:arXiv:14091556
21. Sivakumar S, Chandrasekar C (2013) Lung nodule detection using fuzzy clustering and support vector machines. Int J Eng Technol 5(11):179–185
22. Society, AC (2016) February 8, 2016 [cited 2016 February 28, 2016]; Available from: http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lungcancer-key-statistics
23. Song Y, Cai W, Zhou Y, Feng DD (2013) Feature-based image patch approximation for lung tissue classification. IEEE Trans Med Imaging 32(4):797–808

24. Sorensen L, Shaker SB, De Bruijne M (2010) Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans Med Imaging 29(2):559–569
25. Suzuki K (2012) Pixel-based machine-learning in medical imaging. Int J Biomed Imaging 2012: article ID 792079, 18 pages
26. Suzuki K (2013) Machine learning in computer-aided diagnosis of the thorax and Colon in CT: a survey. IEICE Trans Inf Syst E96-D:772–783
27. Suzuki K (2017) Overview of deep learning in medical imaging. Radiol Phys Technol 10(3):257–273
28. Suzuki K (2017) Machine learning in medical imaging before and after introduction of deep learning. J Med Imaging Inform Sci 34(2):14–24
29. Suzuki K (2017) Survey of deep learning applications to medical image analysis. Med Imaging Technol 35(4):212–226
30. Suzuki K, Armato SG III, Li F, Sone S, Doi K (2003) Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. Med Phys 30(7):1602–1617
31. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: IEEE, CVPR pp. 1
32. Tajbakhsh N, Lian J (2015) Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. In: Proc MICCAI
33. Tajbakhsh N, Suzuki K (2017) Comparing two classes of end-to-end learning Machines for Lung Nodule Detection and Classification: MTANNs vs. CNNs. Pattern Recogn 63:476–486
34. Tajbakhsh N, Gurudu SR, and Liang J (2015) A comprehensive computer-aided polyp detection system for colonoscopy videos. In: Information processing in medical imaging. New York: Springer, pp. 327–338
35. Tajbakhsh N, Gurudu SR, and Liang J (2015) Automatic polyp detection in colonoscopy videos using an ensemble of convolutional neural networks. In; Proc. IEEE 12th Int Symp O Biomed Imag, pp. 79–83
36. Valente IRS, Cortez PC, Neto EC, Soares JM, de Albuquerque VHC, Tavares JMR (2016) Automatic 3D pulmonary nodule detection in CT images: a survey. Comput Methods Prog Biomed 124:91–107
37. Vedaldi A and Lenc K (2014) Matconvnet – convolutional neural networks for matlab. CoRR, abs/14124564
38. Xu Z, Mei L, Liu Y, Hu C, Chen L (2016) Semantic enhanced cloud environment for surveillance data management using video structural description. Computing 98(1–2):35–54
39. Ye J, Ding Y (2018) Controllable keyword search scheme supporting multiple users. Future Generation Comp Syst 81:433–442
40. Zhang J, Xia Y, Xie Y, Fulham M, Feng DD (2017) Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features," IEEE journal of biomedical and health informatics, available online 20 November 2017. https://doi.org/10.1109/JBHI.2017.2775662
41. Zhang J, Xia Y, Cui H, Zhang Y (2018) Pulmonary nodule detection in medical images: a survey. Biomed Signal Process Control 43:138–147
42. Zheng Y, Liu D, Georgescu B, Nguyen H, and Comaniciu D (2015) 3D deep learning for efficient and robust landmark detection in volumetric data. In: Proc. MICCAI, pp. 565–572

Zhenghao Shi received the BS degree in material science and engineering from Dalian Jiaotong University, Dalian, China, in 1995, the MS degree in computer application technology from Xi'an University of Technology, Xi'an, China, in 2000, and the Ph.D degree in computer architecture from Xi'an institute of microelectronics, Xi'an, China, in 2005. In 2000, he joined Xi'an University of Technology, Xi'an, China. From 2000 to 2005, he was an assistant professor in the department of computer science and engineering at Xi'an University of Technology. From 2006 to now, he is an associate professor of the department of computer science and engineering at the same University. Since December 2016, he joined IDBE laboratory, University of North Carolina at Chapel Hill as a visiting associate professor. During the period of 2006 to 2007, of 2008 to 2009, he was on leave with the department of computer science and engineering at Nagoya institute of technology, Nagoya, Japan, for image research as a postdoctral researcher, respectively. From 2007 to 2008, he was a research associate in the Kurt Rossmann Laboratories for Radiologic Image Research, the Department of Radiology, the Division of Biological Sciences, the University of Chicago. His research interests include neural networks for image processing and pattern recognition, computer-aided diagnosis, and image processing suggested by the human visual systems. He is a member of the IEEE, also a member of ACM.



Huan Hao is currently study for her master degree in computer science and technology in Xi'an University of Technology, Xi'an, China. Her research interests focus on medical image processing.



Minghua Zhao received her Ph.D degree in computer science from Sichuan University, Chengdu, China, in 2006. After that, she joined Xi'an University of Technology, Xi'an, China. Currently, she is an associate professor of the department of computer science and engineering at the same University. Her research interests include image processing and pattern recognition.

Yaning Feng received her Ph.D degree in material science from Nagoya Institute of Technology, Nagoya, Japan, in 2008. After that, she joined Xi'an University of Technology, Xi'an, China. Currently, she is an assistant professor of the same University. Her research interests include material image processing and pattern recognition.

Lifeng He received the B.E. degree from Northwest Institute of Light Industry, Shaanxi, China, in 1982, the second B.E. degree from Xian Jiaotong University, Shaanxi, in 1986, and the M.S. and the Ph.D. degrees in artificial intelligence and computer science from Nagoya Institute of Technology, Aichi, Japan, in 1994 and 1997, respectively. He is currently a Professor at Aichi Prefectural University, Aichi, and a Guest Professor at Shaanxi University of Science and Technology, Shaanxi. From September 2006 to May 2007, he was a Research Associate at The University of Chicago. His research interests include intelligent image processing, computer vision, medical image processing.

Yinghui Wang received his Ph.D. degrees in computer science from Northwest University,Xi'an, Chin, in 2001. He is currently a Professor at Xi'an University of Technology, Xi'an, China. His research interests include image processing, computer vision.

Kenji Suzuki joined Department of Electric and Computer Engineering and Medical Imaging Research Center, Illinois Institute of Technology, as Associate Professor in 2014. Before that, he worked in Research and Development Center, Hitachi Medical Corporation, Japan, from 1993 to 1997. From 1997 to 2001, he worked in Faculty of Information Science and Technology, Aichi Prefectural University, Japan, as a faculty member. In 2001, he joined Department of Radiology, the University of Chicago, as Research Associate. He was promoted to Assistant Professor of Radiology, Graduate Program in Medical Physics, and Cancer Research Center in 2006. His research interests include computational intelligence and machine learning in medical imaging, computer-aided detection and diagnosis, and medical image processing and analysis. He published more than 300 papers (including 110 peer-reviewed journal papers). His papers were cited more than 8000 times by other researchers, and he has an h-index of 39. He is one of top 10 most cited researchers in the world in the computer-aided diagnosis field by Google Scholar. He is inventor on 30 patents (including 12 granted patents), which were licensed to several companies and commercialized. He published 10 books and 22 book chapters, and edited 12 journal special issues. He was awarded more than 25 grants as the P.I. including NIH R01 and ACS grants. He served as a grant reviewer for funding agencies, including National Science Foundation (NSF), National Institutes of Health (NIH), US Army Medical Research and Materiel Command (USAMRMC), U.S. Department of State, and American Association for the Advancement of Science (AAAS). He served as the Editor-in-Chief/Associate Editor of 43 leading international journals, including Pattern Recognition (IF: 3.4), Medical Physics (IF: 2.5), Neurocomputing (IF: 2.4), BioMed Research International (IF: 2.1), and Academic Radiology (IF: 2.0). He served as a referee for 80 international journals, an organizer of 30 international conferences, and a program committee member of 150 international conferences. He gave 90 invited talks and keynote speeches at international conferences and universities. He supervised more than 70 graduate/undergraduate students and 7 postdocs/computer scientists, and accepted 16 visiting professors in his laboratory from foreign countries, supported by grants/scholarships from their governments/institutions. He received 25 awards for his research and teaching, including Paul Hodges Award, 3 Certificate of Merit Awards from RSNA, IEEE Outstanding Member Award from IEEE, Cancer Research Foundation Young Investigator Award, Kurt Rossmann Award for Excellence in Teaching from University of Chicago, IEICE Best Paper Award, and Most Cited Journal Paper Award 2016 from EANM. He has been a Senior Member of IEEE since 2004.