# A Transfer Learning Method for CT Image Classification of Pulmonary Nodules

Ran Wang[1,2](✉), Huadong Sun[1,2], Jialin Zhang[1], and Zhijie Zhao[1,2]

[1] School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China
wr0905@sina.com

[2] Provincial Key Laboratory of Electronic Commerce and Information Processing, Harbin University of Commerce, Harbin 150028, China

**Abstract.** A pulmonary nodule classification method of Computer Tomography (CT) images based on transfer learning of deep convolutional neural network (CNN) is proposed. Lung CT images with labels are quite limited compared with the large scale image database such as ImageNet. It is easy to produce over-fitting problem when using the limited data to train the deep CNN for classification task. In this paper, in order to overcome this difficulty, the deep CNNs GoogleNet and ResNet are pre-trained on the large scale database ImageNet. The fully connected layers and the classifiers of the pre-trained networks are replaced to complete the classification of CT images of pulmonary nodules. A sub set of the Lung Image Database Consortium image collection (LIDC-IDRI) is used to fine-tune the network and validate the classification accuracy. This is the process of transfer learning. It solves the problem of the deficiency of lung CT images as labeled training data for CNNs. By the knowledge obtained from the pre-trained CNNs which have been trained on ImageNet, the network is easier to converge and the training time is greatly reduced. The classification accuracy of Pulmonary Nodules can be reached up to 71.88% by using the proposed method.

**Keywords:** Transfer learning · Convolutional neural network · Computer tomography image · Pulmonary nodule

## 1 Introduction

In recent years, with the increase of environmental pollution, the incidence of lung cancer is on the rise. According to the published 2017 Chinese Cancer Annual Report, morbidity and mortality of lung cancer are both in the first place, seriously harm the health of Chinese people. The survey report of American Cancer Society (ACS) shows, at present the 5-years survival rate of lung cancer patients is only 17%, but the survival rate can be increased to 55% as long as early treatment is obtained [1]. Therefore, the early diagnosis of lung cancer is of great significance to patients. Early lung cancer is mainly characterized by pulmonary nodules. Therefore, early detection of pulmonary nodules is important for increasing the survival rate of lung cancer patients.


M. Jia et al. (Eds.): WiSATS 2019, LNICST 281, pp. 159–166, 2019.
https://doi.org/10.1007/978-3-030-19156-6_15

For the detection of pulmonary nodules, the currently widely accepted method which is safe and effective is Computer Tomography (CT). The advantage of CT images is that its resolution is high, and organ slices can be seen at different locations. By analyzing the CT image, it is not only possible to check whether there is a pulmonary nodule and locate the nodule, but also to analyze the size, density, morphology, internal structure and edge characteristics of the nodule [2]. However, the amount of CT images is large, one patient's lung CT images can reach more than 100 sheets. For doctors, viewing CT images of patients requires a large amount of time.

The emergence of the computer-aided diagnosis based on image processing and convolutional neural network (CNN) has reduced the workload of doctors for reading CT images [3]. The CNN can be used to classify the CT images when being successfully trained. Image classification based on CNN overcomes some shortcomings of traditional image classification methods. There is no need to manually design features, the network can extract image features and classify the images when having trained on a large data set. The CNN is going deeper and larger to increase its expression ability. However, if the training data is limited, it is easy to produce over-fitting and network degradation problems [4, 6]. In order to solve these problems, the GoogleNet [5], ResNet [6] architecture and the concept of transfer learning have been proposed [7–9].

Using these methods, we start with the concept of transfer learning. Then the steps of transfer learning of CNN and the medical image database we use to train the networks are shown. Finally, the experimental results and analysis are given to prove the validity of the presented method.

## 2 The Transfer Learning of Convolutional Neural Network

Transfer learning is a machine learning method that solves different but related domain problems by using existing knowledge [7]. The basic idea of transfer learning is to learn a representation that is shared across related tasks. In transfer learning the higher the correlation between the source and target fields, the better the learning results will be [8]. Different levels of features have different transfer learning abilities; higher-level features have better transfer learning ability than lower-level features [9]. In transfer learning, the common features can be learned by solving an optimization problem, given as follows:

$$J(\Theta, U) = \arg\min_{\Theta, U}\left[\sum_{t\in\{T,S\}}\sum_{i=1}^{n_t} L\left(y_{t,t}, \left\langle\theta_t, U^T x_{t,t}\right\rangle\right) + \gamma\|\Theta\|_{2,1}^2\right] \tag{1}$$

In Eq. (1), $S$ and $T$ denote the tasks in the source domain and target domain, $x$ and $y$ denote the feature vectors and the labels, respectively. $\Theta = [\theta_s, \theta_t] \in R^{d\times 2}$ is a matrix of parameters. $\gamma$ is the regularization parameter. $U$ is a $d \times d$ orthogonal matrix (mapping function) for mapping the original high-dimensional data to low-dimensional representations. The $(r, p)$-norm of $\Theta$ is defined as:

$$\|\Theta\|_{r,p} = \left(\sum\nolimits_{i=1}^{d} \|\theta^{i}\|_{r}^{p}\right)^{1/p} \tag{2}$$

The optimization problem (1) estimates both of the low-dimensional representations $U^T X_T, U^T X_S$ and the parameters $\Theta$ of the model at the same time.

The transfer learning flow chart of CNN is shown in Fig. 1:

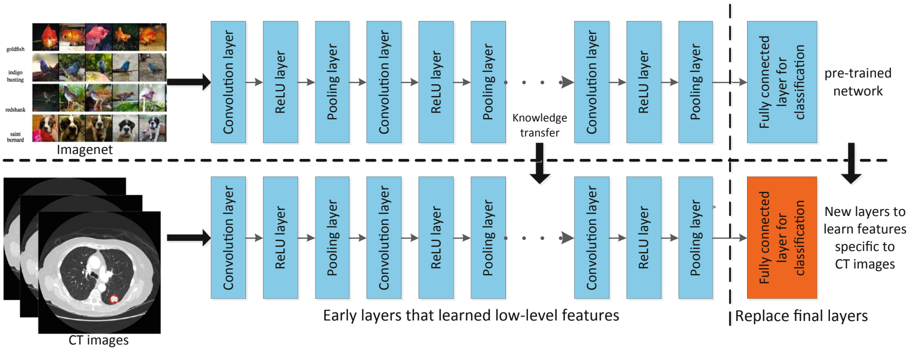



**Fig. 1.** The transfer learning flow chart of CNN.

First, a pre-trained CNN is loaded, it may have been trained on a large scale data set such as ImageNet; then, the last few layers of the loaded network are replaced by new layers to learn features specific to the dataset of new field; finally, the new dataset and training parameters are used to fine-tuning the network.

### 2.1 The Pre-trained CNNs for Transfer Learning

Different CNNs have different characteristics. The most important characteristics are network accuracy, speed, and size. Choosing a network is generally a tradeoff between these characteristics. A good network has high accuracy and is fast. Figure 2 shows the top1 accuracy versus operations, size/parameters of different CNNs [11].

Top-1 one-crop accuracy versus amount of operations required for a single forward pass is shown in Fig. 2. The size of the blobs is proportional to the number of network parameters. A legend is reported in the bottom right corner, spanning from $5 \times 10^6$ to $155 \times 10^6$ parameters. In general, GoogleNet and its derivative inception series have the characteristics of lower resource consumption and medium accuracy. ResNet series have the characteristics of medium resource consumption and higher accuracy. These two types of CNN architectures have the best comprehensive performance.

In this paper, GoogleNet and ResNet are chosen for transfer learning. The pre-trained GoogleNet and ResNet we used in this paper have been trained on more than one million images and can classify images into 1000 object categories. The training images are a subset of the ImageNet database [10], which is used in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). Using a pre-trained network with

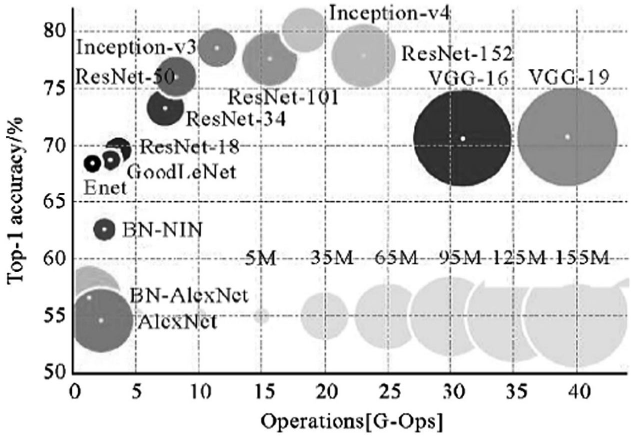



**Fig. 2.** Top1 accuracy versus operations, size/parameters of different networks

transfer learning is typically much faster and easier than training a network from scratch.

## 2.2 Fine-Tuning of the CNNs

For large scale CNNs as GoogleNet and ResNet, the lower level features show the color or edge feature of the image, well the higher level features show the texture feature, the more class-specific features and the entire objects with significant pose variation successively [8]. The effects of transfer learning of different levels of CNNs have been studied: high-level features of CNNs have better transfer learning ability than lower-level features [12]. In this paper, all the parameters and structure of the pre-trained networks are preserved besides the last three layers to obtain the higher level features before the fine-tuning of the networks.

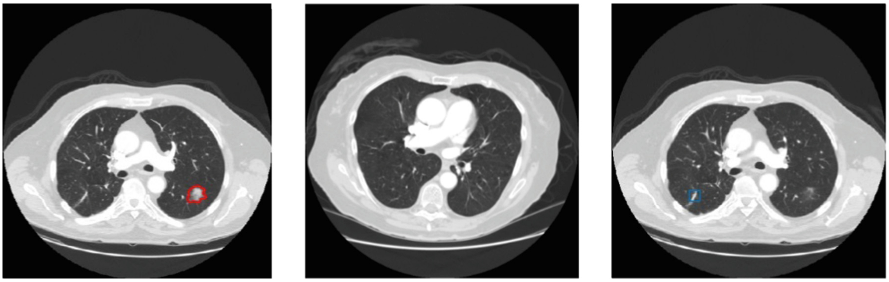
The last three layers of the pre-trained CNNs are replaced by three new layers to complete classification of CT images of pulmonary nodules. The new layers are one fully connect layer, one softmax layer and one classifier of three categories. A sub set of CT images of the Lung Image Database Consortium image collection (LIDC-IDRI) is divided into training set and test set. After the replacement of the last three layers, the training set is used to fine-tune the network. When the fine-tuning is finished, the test set is used to validate the classification accuracy of the network.

GoogleNet can be iterated quickly and try out different parameter settings such as data preprocessing steps and training options. From that we may find which parameters are fit for our problem. Then a more accurate network: ResNet will be tried to see if the predict accuracy is improved.

## 3 The Lung Image Database Consortium Image Collection

In this paper, a sub set of the Lung Image Database Consortium image collection (LIDC-IDRI) is used for classifying the pulmonary nodules. LIDC-IDRI consists of diagnostic and lung cancer screening thoracic computed tomography (CT) scans with marked-up annotated lesions [13]. It is a web-accessible international resource for development, training, and evaluation of computer-assisted diagnostic (CAD) methods for lung cancer detection and diagnosis. It initiated by the National Cancer Institute (NCI), further advanced by the Foundation for the National Institutes of Health (FNIH). The data set contains 1018 cases. Each subject includes images from a clinical thoracic CT scan and an associated XML file that records the results of a two-phase image annotation process performed by four experienced thoracic radiologists.

Figure 3 shows Lung CT images of a patient in LIDC-IDRI. All the images are in the format of bitmap, they have three channels and the resolution of them is $512 \times 512$. Figure 3(a) shows CT images with pulmonary nodules, the edge of the nodule is marked with a red line. Figure 3(b) shows CT images without nodules. Figure 3(c) shows CT images with a suspected nodule which is marked with a green line.



(a) Pulmonary nodule (b) Without nodule (c) Suspected nodules

**Fig. 3.** Lung CT images of a patient in LIDC-IDRI (Color figure online)

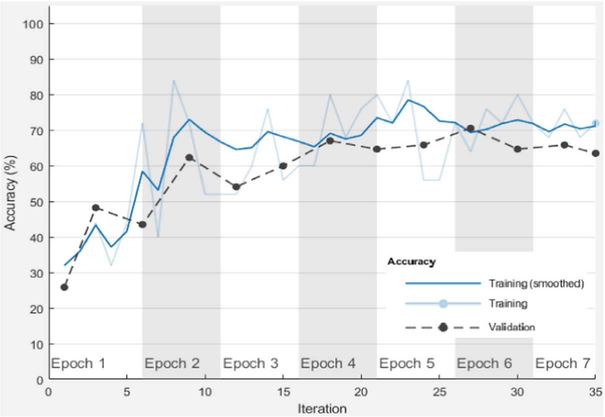
## 4 Experimental Results and Analysis

The experimental environment is Windows 7 operating system and Matlab R2017b with deep learning toolbox.

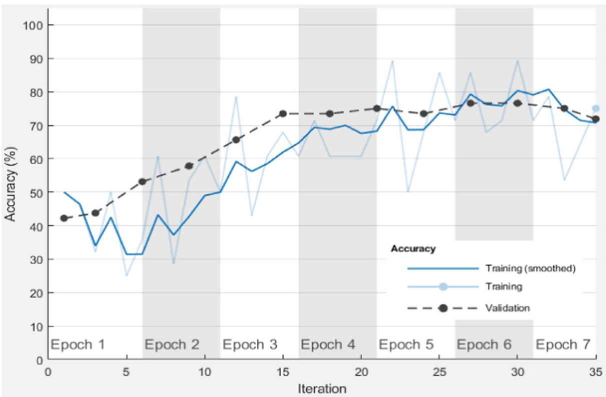
**Table 1.** Training options for fine-tuning

| Training options | GoogleNet | ResNet |
|---|---|---|
| Mini Batch Size | 28 | 28 |
| Epochs | 7 | 7 |
| Initial Learn Rate | $1.4 \times 10^{-4}$ | $10^{-4}$ |
| Verbose Frequency | 1 | 1 |
| Network Depth | 22 | 50 |

We choose 213 CT images of three patients in LIDC-IDRI which are divided into 3 categories: 73 CT images of lung tissue without nodules, 68 CT images of pulmonary nodules and 72 CT images of suspected nodules. 149 of 213 (70%) CT images are chosen stochastically from all 3 categories as training data and the rest 64 CT images are defined as validating data. All the images are resized to the resolution of 224 $\times$ 224 to fit the input size of the networks. Then the pre-trained GoogleNet or ResNet is loaded directly by googlenet and resnet50 function of Matlab. The last three layers: fully connected layer, softmax layer and classification layer are replaced by new ones which can be seen as a classifier of 3 classes. Then we use the CT images to fine-tune the network. The training options we use for fine-tuning GoogleNet and ResNet are shown in Table 1. For 149 training images and the Mini Batch Size of 28, there are 5 iterations per epoch and 35 iterations in total. Figure 4 shows the Classification accuracy of the two networks:




(a) Classification accuracy of GoogleNet




(b) Classification accuracy of ResNet

**Fig. 4.** Classification accuracy of the two networks

Figure 4(a) and (b) is the classification accuracy of GoogleNet and ResNet, respectively. In Fig. 4, training accuracy means the classification accuracy on each individual mini-batch. Smoothed training accuracy is obtained by applying a smoothing algorithm to the training accuracy. It is less noisy than the unsmoothed accuracy, making it easier to spot trends. Validation accuracy is the classification accuracy on the entire validation set. As shown in Fig. 4, the validation accuracy of the ResNet is higher than the GoogleNet. The validation accuracy of the GoogleNet has larger variation in the initial 3 epochs because of the larger Initial Learn Rate. It becomes smoother in the last 4 epochs. The final results of validation accuracy of the GoogleNet and the ResNet are 63.53% and 71.88%, respectively.

## 5 Conclusion

In this paper we presented a pulmonary nodule classification method in CT images based on transfer learning of CNN. The application of transfer learning solved the problem of the insufficiency of training data of CT images for training the network. By using transfer learning, we need only about two hundreds CT images in total to finish fine-tuning and validating the pre-trained GoogleNet or ResNet. Compared with training from scratch, the training time is greatly reduced when using transfer learning.

Two architectures of deep CNNs GoogleNet and ResNet have been pre-trained on a subset of ImageNet which contains millions of images in 1000 categories. By the replacement of last few layers and fine-tuning of the network, the knowledge which the networks learned from ImageNet by pre-training was transferred to pulmonary nodule CT images classification task.

It has been shown in our experiments that GoogleNet is iterated quickly, while ResNet is more accurate. The classification accuracy of the GoogleNet and the ResNet reaches up to 63.53% and 71.88% respectively, and the computing resource requirement is quite small.

**Acknowledgment.** This work is supported by the Basic Scientific Research Operating Expense Project of Provincial Institutions of Higher Education in Heilongjiang (17XN003); the Doctoral Research Startup Project of Harbin University of Commerce (2016BS28); and the Humanities and Social Sciences Research Projects of the Ministry of Education (18YJAZH128).

## References

1. American Cancer Society: Cancer Facts and Figures 2016. American Cancer Society, Atlanta (2016)
2. Diciotti, S., Picozzi, G., Falchini, M., Mascalchi, M., Villari, N.: 3-D segmentation algorithm of small lung nodules in spiral CT Images. IEEE Trans. Inf. Technol. Biomed. **12**(1), 7–19 (2008)
3. Wang, Y., Zhou, T., Lu, H., Wu, C., Yang, P.: Computer aided diagnosis model for lung tumor based on ensemble convolutional neural network. J. Biomed. Eng. **4**(34), 543–551 (2017)

4. Hinton, G.E., Srivastava, N., Krizhevsky, A.: Improving neural networks by preventing co-adaption of feature detectors, pp. 1–18. arXiv:1207.0580 [cs.NE] (2012)
5. Szegedy, C., Liu, W., Jia, Y.: Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE Computer Society, Washington (2015)
6. He, K., Zhang, X., Ren, S.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, United States, pp. 770–778 (2016)
7. Zhuang, F., Luo, P., He, Q., Shi, Z.: Survey on transfer learning research. J. Softw. **26**(1), 26–39 (2015)
8. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8689, pp. 818–833. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_53
9. Zhou, B., Lapedriza, A., Xiao, J.: Learning deep features for scene recognition using places database. In: Proceedings of Advances in Neural Information Processing Systems, pp. 487–495. MIT Press, Cambridge (2014)
10. ImageNet Homepage. http://www.image-net.org. Accessed 8 July 2018
11. Canziani, A., Paszke, A., Culurciello, E.: An analysis of deep neural network models for practical applications, pp. 1–7. arXiv:1605.07678 [cs.CV] (2016)
12. Razavian, S., Azizpor, H., Sullivan, J.: CNN features off-the-shelf: an astounding baseline for recognition, pp. 1–8. arXiv:1403.6382 [cs.CV] (2014)
13. LIDC-IDRI Collection. https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI. Accessed 4 Aug 2018