Pulmonary Image Classification Based on Inception-v3 Transfer Learning Model

CHENG WANG [1], DELEI CHEN [1], LIN HAO[2], XUEBO LIU[2], YU ZENG[1], JIANWEI CHEN[3], AND GUOKAI ZHANG [4]

[1]College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China
[2]Department of Cardiology, Shanghai Tongji Hospital, Tongji University, Shanghai 200065, China
[3]Department of Mathematics and Statistics, San Diego State University, San Diego, CA 92182, USA
[4]School of Software Engineering, Tongji University, Shanghai 201804, China

Corresponding authors: Cheng Wang (wangcheng@hqu.edu.cn), Xuebo Liu (lxb70@hotmail.com), Jianwei Chen (jchen@mail.sdsu.edu), and Guokai Zhang (zhangguokai_01@163.com)

ABSTRACT Chest X-ray film is the most widely used and common method of clinical examination for pulmonary nodules. However, the number of radiologists obviously cannot keep up with this outburst due to the sharp increase in the number of pulmonary diseases, which increases the rate of missed diagnosis and misdiagnosis. The method based on deep learning is the most appropriate way to deal with such problems so far. The main research in this paper was using inception-v3 transfer learning model to classify pulmonary images, and finally to get a practical and feasible computer-aided diagnostic model. The computer-aided diagnostic model could improve the accuracy and rapidity of doctors in the diagnosis of thoracic diseases. In this experiment, we augmented the data of pulmonary images, then used the fine-tuned Inception-v3 model based on transfer learning to extract features automatically, and used different classifiers (Softmax, Logistic, SVM) to classify the pulmonary images. Finally, it was compared with various models based on the original Deep Convolution Neural Network (DCNN) model. The experiment proved that the experiment based on transfer learning was meaningful for pulmonary image classification. The highest sensitivity and specificity are 95.41% and 80.09% respectively in the experiment, and the better pulmonary image classification performance was obtained than other methods.

INDEX TERMS Pulmonary image, data augmentation, deep convolutional neural network, transfer learning, inception-v3.

I. INTRODUCTION

Lung disease is a kind of respiratory disease. With the rapid development of modern industry and transportation, air pollution is serious, and the incidence of lung disease has increased significantly. It is also known that there is nearly 600 million Chronic obstructive pulmonary disease (COPD) patients worldwide, and COPD is expected to become the third fatal disease in the world by 2020. However, It is reported that about 20% to 50% of pulmonary nodules are misdiagnosed or missed in chest X-ray examination, and even highly skilled

The associate editor coordinating the review of this manuscript and approving it for publication was Yongtao Hao.

radiologists inevitably make 3% to 6% clinical errors [1], [2]. Therefore, the computer-aided detection system appears.

The research of the computer-aided detection system has been concerned by many researchers. Although computer-aided diagnostic system (CAD) for the pulmonary image has been applied in the clinical stage, there are still many defects and problems to be solved urgently. M. Woźniak and Połap [3] proposed an automated pulmonary disease decision support system based on Bio-Inspired Method implemented for x-ray medical images. The method simulates the process of medical examinations of pulmonary diseases to model decision support that directs where diseased tissues are most likely to occur. The system has high accuracy,

but it is less efficient in time complexity. Ke *et al.* [4] proposed a method based on neural network co-working with heuristic algorithms (Moth-Flame, Ant Lion) to detect degenerated lung tissues in the x-ray image. The proposed neuro-heuristic approach can detect small changes in lung tissue structure. Filho *et al.* [5] evaluates the performance of Optimal Path Forest (OPF) classifier in the task of diagnosing pulmonary diseases on CT images. Three feature extraction methods and seven distance measurement functions are compared in the evaluation. Moreover, the result in [5] shows that the OPF classifier is superior to other traditional supervised learning algorithms in computational efficiency and suitability. Woźniak *et al.* [6] proposed a new lung cancer classification method based on the probabilistic neural network. This method is relatively simple, but it also has good classification effect and can detect low contrast nodules. For extracting the features of lung image, Li *et al.* [7] used the stationary wavelet transform and convergent exponential filter to extract texture features of the Pulmonary, then used AdaBoost to generate white nodule diagram. Supanta *et al.* [8] proposed a method to extract and detect pulmonary nodules from digital radiography images to solve the problems of little visualization and highlighting of these features. Mukherjee *et al.* [9] proposed a method based on geometric features to detect pulmonary nodules automatically. This method classifies benign and malignant pulmonary nodules by shape parameters such as roundness, eccentricity, diameter and aspect ratio. Most of these methods need relevant professional knowledge or need to spend a lot of time and energy to design and implement methods and algorithms of feature extraction. With the gradual development of deep learning research, the technology that can be applied to the image has also made a qualitative leap [10], so the idea of medical image classification based on deep learning emerged. The research in medical image classification has achieved many research results, and it has been applied clinically [11]. Deep learning does not require any medical-related professional knowledge, also does not need to characteristics of engineering technology. It can go through the initial layer to learn essential characteristics such as color, edge, and then use the hidden layer to study more advanced on the specific characteristics of pulmonary imaging data sets. In terms of Pulmonary image classification methods research, Shin *et al.* [12] selected OpenI database, opened available chest X-ray radiology datasets and their diagnostic reports to train convolutional neural network/recurrent neural network (CNN/RNN). In this process, they use the feature of the graph extracted by CNN and the corresponding description of the image as input of RNN model. Then they use the state vector of the image at all T-Time in RNN model to mean pooling and get an average expression as image/text context. Finally, they retrain CNN/RNN. Their method adapts to medical image processing in model training strategy, but their experimental performance is not well. Bar *et al.* [13] explore the ability of CNN learned from a non-medical dataset to identify different types of pathologies in chest radiographs. In [13], he reduced the dimensionality of image data from the diagnostic imaging

department through PCA and standardized all features, and then he chooses the AlexNet model trained by ImageNet to train further and classify pulmonary nodules. Its main advantage is to solve the problem that large data sets are usually not available in the medical field, but the experiments were conducted only on private databases and could not be compared with the results of others.

Combined with the above research, this paper proposes a deep learning model based on transfer learning, The main contributions of this paper are as follows:

Combined with the above research, this paper proposes a deep learning model based on transfer learning, The main contributions of this paper are as follows:

(a). A method based on inception V3 migration learning is proposed for disease detection of lung images, thus provide an effective computer-aided diagnosis model for pulmonary X-ray film classification research in the absence of medical data.

(b). Compared with the original DCNN model, this method can effectively improve the accuracy of lung image classification.

II. METHODS
A. DATA AUGMENTATION

Data augmentation is an important part of pulmonary image classification. The appropriateness of data augmentation will directly affect the final effect of pulmonary image classification. Pulmonary image belongs to the category of medical image, so it is inevitable that pulmonary image also has the shortcomings of medical image. Labeling medical image, unlike voice data or text data, can not be outsourced to a professional company or team. Therefore, the labeling of these data can only be done by professional radiologists, and it needs complex experienced radiologists to spend half an hour or more time to observe image over and over again and hand out the location of the nodules. As a result, medical image data inevitably suffers from a lack of data due to the scarcity of professionals (far fewer than the number of professionals who can annotate other voice or text data). The significance of data augmentation lies in that, on the limited pulmonary image data set, the data amount of training can be expanded sufficiently and reasonably to improve the generalization ability of the model. Reasonable and appropriate noise data are added to improve the robustness of the model.

In order to eliminate dimensional effects between data, Z-score standardization method was used to normalize the data. Moreover, zero-phase Component Analysis(ZCA) was used to whiten the data to reduce the correlation of feature data and simplify the subsequent extraction process.

The main purpose of Z-Score is to transform data of different magnitudes into the same magnitude to ensure the comparability of data. Compared with other normalization methods, Z-score can make the gradient descent converge faster, and reduce the influence of extreme values in eigenvalues to present data in a more explanatory way. The calculation

formula of Z-score standardization is shown in (1):

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (i = 1, 2, ..., n; \; j = 1, 2, ..., m) \quad (1)$$

where $\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$, $s_j = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2}$, n is the number of sample, m is the number of feature.

The z-score standardized method represents the length of the original data from the mean value, and the unit is a standard deviation. Therefore, when the value of the new feature data is less than zero, it means that the data is less than the mean. When the value is equal to zero, it means that the data is equal to the mean. When the value is greater than zero, it means that the data is greater than the mean.

Whitening is a step often used in data augmentation in the field of deep learning. After whitening, the correlation of feature data of pulmonary image will be significantly reduced, and all feature data have the same variance. This result can simplify the subsequent extraction process and make the model have better convergence. Whitening can be divided into Principal Component Analysis(PCA) whitening and ZCA whitening. PCA whitening refers to the singular value decomposition of the covariance matrix of the data in the determined data set, and the first k feature vectors are selected for dimensionality reduction of the data. Finally, the dimensionality reduction of the data is scaled and the scaling factor is divided by the square root of its corresponding feature value. The formula of PCA whitening is shown in (2):

$$\mathbf{X}_{PCA-whitening} = \mathbf{S}^{-\frac{1}{2}} \mathbf{U}^T \mathbf{X} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{\sqrt{\lambda_k}} \end{bmatrix} \mathbf{X}_{rotate} \quad (2)$$

where the dataset $\mathbf{X} \in \mathbb{R}^{n \times m}$, n is the number of feature, m is the number of sample, $\mathbf{U}$ is the eigenvector matrix of the covariance matrix $\sum$, $\mathbf{S}$ is the eigenvalue matrix of covariance matrix $\sum$, λ_i is the ith eigenvalue, $i = 1, 2, ..., k$, $\sum = \frac{1}{m} \mathbf{X}\mathbf{X}^T$, $\mathbf{X}_{rotate}$ is the projection of the original data on the principal axis.

ZCA whitening is to transform the data obtained by PCA whitening to the original space again. Compared with PCA, ZCA can ensure that the variance of data dimensions is the same so that the whitened data is close to the original input data. Therefore, ZCA was chosen for whitening. The formula of ZCA is shown in (3):

$$\mathbf{X}_{ZCA-whitening} = \mathbf{U}\mathbf{X}_{PCA-whitening} \quad (3)$$

In this paper, an iterative method is adopted to conduct eigenvalue standardization and ZCA whitening, and the processed images have a increase in contrast, the contrast of the images increased from 10123 to 12357 on average.

The images after ZCA and Z-Score are expanded by the operations such as rotation, flip and shift, which can provide

TABLE 1. Comparison of four famous neural network models.

Model	AlexNet[16]	VGG[17]	GoogleNet[18]	ResNet[19]	NASNet[20]
Time of presentation	2012	2014	2014	2015	2018
The number of layers	8	19	22	152	-
The error of Top-5	16.4%	7.3%	6.7%	3.57%	3.8%
Data enhancement	√	√	√	√	√
Inception (NIN)	×	×	√	×	√
The number of convolutions	5	16	21	151	-
The number of full connections	3	3	1	1	1
Random inactivation	√	√	√	√	√
Local response normalization	√	×	√	×	√
batch of standardized	×	×	×	√	√

enough training samples for deep learning model, greatly improve the training effect of the model and the result. The data were amplified 18 times by rotating at 20 degrees, amplified four times by the horizontal and vertical flip and amplified three times the translation in 20 pixels. The images were expanded from 247 chest radiographs to 47424 chest radiographs.

B. DEEP CONVOLUTIONAL NEURAL NETWORK (DCNN)

The development of the convolutional neural network model is not plain sailing. In 1980, Fukushima [14] proposed a neurocognitive machine based on visual pattern recognition mechanism which is widely regarded as the first version of the convolutional neural network. It decomposes and combines the visual system layer by layer to model it so that it can correctly identify objects with displacement or slight deformation. In 1989, LeCun *et al.* [15] proposed a convolutional neural network with a five-layer structure, named LeNet, which solved the problem of handwritten numeral recognition. It was the beginning of the convolutional neural network from research to application. However, due to the lack of training data and the weak computing ability of computer hardware, the convolutional neural network was not developed at that time. On the contrary, manual feature extraction with SVM performed well on small data sets, leading to its mainstream research at that time. Then in 2012, AlexNet, a famous neural network model established by Krizhevsky *et al.* [16], won the first place in the ImageNet image classification contest with an accuracy rate of 10% over the second place. Since then, the convolutional neural network model has returned to the vision of researchers. VGG model [17], GoogleNet model [18], ResNet model [19], NASNet [20] model and other network models have sprung up. Their neural network structures are compared as shown in Table 1.

In general, many independent neurons can form a two-dimensional plane, and the deep convolutional neural network is composed of many layers in which many two-dimensional planes of feature mapping form. A deep convolutional neural network consists of four core parts [21].

(a) The first part is the local perception which all neurons in the neural network do not need to perceive the global

TABLE 2. The structure of DCNN models in this paper.

Type	Patch size/Stride	output size
Conv2d	5×5/1	100*100*32
Activation	/	100*100*32
Maxpooling	2×2/2	50*50*32
Conv2d	5×5/1	50*50*32
Activation	/	50*50*32
Maxpooling	2×2/2	25*25*32
Conv2d	5×5/1	25*25*32
Activation	/	25*25*32
Maxpooling	2×2/2	25*25*32
Flatten	/	20000
Dense	/	1024
Activation	/	1024
Dense	/	512
Activation	/	512
Dense	/	256
Activation	/	256
Dense	/	2
Activation	/	2

TABLE 3. Experimental result of filter with different sizes.

The number of filters in each CNN layer	Training accuracy	Test accuracy	The elapsed time(s)
32-32-32	96.40%	64.01%	335
32-32-64	61.26%	64.02%	512
64-64-64	98.65%	76.07%	651
64-64-128	97.75%	65.21%	1226

image, but only the local information, and global information is obtained by gathering local information.

(b) The second one is the convolution. The function of convolution is to extract image features, and the number of parameters can be greatly reduced by using the convolution kernel.

(c) The third is weight sharing. Weight sharing means that the parameters in the same convolution kernel are used for the whole picture, and the weight in the convolution kernel will not be changed due to the different positions in the image. Moreover, weight sharing of convolution operation can greatly reduce the number of parameters in the convolution kernel.

(d) The fourth part is pooling. The pooling layer is usually placed behind the convolution layer in the convolutional neural network, which can be used to reduce the characteristic dimension of the output of the convolution layer of the previous layer, but at the same time retain the effective key information of image data.

In this paper, DCNN is used for extracting features. The reason why DCNN model is used for extracting features is that it has two characteristics: local connectivity and weight sharing. These two features are very useful for feature extraction in medical images, which can avoid a large amount of human resources input of repeated markers, and it can also ensure that the learned convolution kernel has a strong response to local features. Moreover, since the contour of Pulmonary images is roughly the same, the identification of two types of normal and abnormal Pulmonary images mainly depends on the existence of local features of pulmonary nodules. Therefore, DCNN model is a very suitable deep learning model for extracting features from chest X-ray films.

Considering the application scene of pulmonary image classification and the learning and fitting ability of the model, the deep convolutional neural network model used in this paper includes three convolutional layers, seven activation layers, and three pooling layers. The detailed structure is shown in table 2.

The input of the DCNN model is a 100×100 gray-scale Pulmonary image. As far as filters are concerned, the more

the number of filters is, the more feature mappings can be extracted, and the better the performance of the model is. Therefore, in order to select the most suitable number of filters under the condition of considering both computing resources and DCNN network performance, the comparative experiments of 32-32-32, 32-32-64, 64-64, 64-64-128 filters are carried out while keeping the other hierarchical structures and other influencing factors unchanged. The performance of JSRT data sets is shown in Table 3.

Therefore, considering the performance considerations, 64-64-64 is selected as the convolution layer filter, and the size of each receptive field is 5*5.

There are many choices of activation functions on the activation layer, such as tanh, sigmoid, relu [21]. Its function is to add non-linear factors to enhance the expression of the models, so it must be non-linear. Therefore, in this experiment, Relu function, which has excellent performance in non-linear, is selected as the activation function. Its expression is shown in (4).

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \tag{4}$$

where x is the input of activation functions, $f(x)$ is the activation functions.

The method of Max pooling is selected in the pooling layer in this paper. The downsampling factor is 2*2, and the pooling step is set to 2. The function of the pooling layer is to reduce computation and parameters in the network while maintaining the main features of the input.

In order to avoid the trouble of changing the learning rate manually, the adaptive optimization algorithm Adadelta is chosen as the optimization algorithm in the DCNN. The algorithm strategy of Adadelta [22] is shown in (5):

$$\begin{cases} m_t = \beta_1 m_{t-1} + (1 - \beta_1)g_t \\ v_t = \beta_2 v_{t-1} + (1 - \beta_2)g_t^2 \\ \hat{m}_t = \dfrac{m_t}{1 - \beta_1^t}, \hat{v}_t = \dfrac{v_t}{1 - \beta_2^t} \\ W_{t+1} = W_t - \dfrac{\eta}{\sqrt{v_t} + \varepsilon}\hat{m}_t \end{cases} \tag{5}$$

where m_t and v_t are first-order momentum terms and second-order momentum terms respectively. β_1, β_2 are the dynamic values, which are usually 0.9 and 0.999 respectively; $\hat{m}_t$ and $\hat{v}_t$ are the revised values respectively. w_t represents the parameters of the t iteration model at time t, $g_t = \Delta J(w_t)$ denotes the gradient of t-iteration's cost function; ε is a very small number (usually 1e-8) to avoid denominator being zero.

For the loss function, the cross-entropy loss function is selected in this experiment, which is used to evaluate the current probability distribution and the real distribution in the training process. The Formula of cross-entropy loss is shown in (6):

$$L = -\left[y \log \hat{y} + (1 - y) \log(1 - \hat{y})\right] \qquad (6)$$

where y is the actual value of label, $\hat{y}$ is the predicted value of label.

C. TRANSFER LEARNING

With the increase of the depth of the neural network, the performance of the neural network will indeed be improved, but this improvement is at the expense of time and computing resources. Therefore, in order to reduce the cost of training, deep learning based on transfer learning emerges. In a word, transfer learning is to transfer the trained model parameters to the new model to help the new model training. Considering that some data or tasks are relevant, the learning efficiency of the new model can be accelerated and optimized by sharing the parameters of the trained model to the new model through transfer learning, instead of starting from scratch as most networks do. In recent years, due to the popularity of deep learning, the demand for data also increases, so transfer learning has been further developed. In order to train a high-quality classifier from the data of different distributions, Dai *et al.* [23] proposed an instance-based TrAdaBoost transfer learning algorithm, which uses the training data to select the useful parts from the different distributions data, so as to reduce the classification error rate. In order to solve the problem of domain Adaptation in transfer learning, Long and Wang [24] proposes Deep Adaptation Network, which extends Deep convolutional neural Network to domain Adaptation; Cao *et al.* [25] proposed Partial Transfer Learning based on selective adversarial networks (SAN). It only transferred the partial samples related to the source domain and the target domain, and then dealt with partial transfer problems through SAN. Because the adversarial network can learn domain invariant features well, it can play a great role in transfer learning. Busto and Gall [26] proposed Open Set Domain Adaptation. They use the relationship between the source domain and the target domain to label the samples of the target domain, and convert the source domain to the same space as the target domain, and finally classify the samples of the target domain.

The number of Pulmonary images in the JSRT database used in this paper is 247, which is seriously insufficient compared with the amount of data required by the neural network model, so this paper adopts the method of transfer learning. Considering the performance of the model and computing resources, we chose Inception-v3 model as transfer learning framework which is trained on ImageNet datasets [27] (including more than 1 million copies a total of 1000 categories of image data) and has a good performance on small data set. This paper refers to the structure of Inception-v3 of Google Net and makes some improvements because of its

structure. In order to make the model more suitable for our experiment, the structure of the model is fine-tuned that the last three layers of the model are cut out, and the results of bottleneck layer are taken as the feature results. The structural schematic diagram of the Inception-v3 model [28] (47 layers in total) in this paper is shown in figure 1.

D. THE PROCESS OF METHOD

The process of method in this paper is divided into four steps. The first step is data preprocessing. The second step is to extract the image features after preprocessing, and the third step is to put the extracted features into the classifier for training. Finally, the model after training is tested. The process is shown in figure 2.

For the process of feature extraction, deep convolution neural network is used for extracting features in this method.

For the process of classification, three classification models are selected to classify the extracted features in this paper. The selected classification models are SVM, Logistic and Softmax.

III. EXPERIMENT
A. DATA DESCRIPTION

In order to compare and analyze the results of other researchers on the classification of pulmonary images, the standard public digital image database JSRT [29], released by the Japanese Society of Radiological Technology, was selected as the experimental data set in this paper. All Pulmonary images (pulmonary nodules and non-pulmonary nodules) in the database were CT image and were judged by three chest radiologists as having pulmonary nodules. There were 154 images with pulmonary nodules and 93 images without pulmonary nodules in JSRT, and each pulmonary image had 2048 × 2048 matrix size and 4096-pixel gray level. Each pixel corresponded to 0.175 mm high resolution, and the diameter of pulmonary nodules ranged from 5 mm to 40 mm. All pulmonary images were labeled by thoracic radiologists according to the patient's age, gender, diagnosis results, nodule coordinates, nodule location. There were 154 images with pulmonary nodules and 93 images without pulmonary nodules in JSRT, and each pulmonary image had 2048 × 2048 matrix size and 4096-pixel gray level. Each pixel corresponded to 0.175 mm high resolution, and the diameter of pulmonary nodules ranged from 5 mm to 40 mm. All pulmonary images were labeled by thoracic radiologists according to the patient's age, gender, diagnosis results, nodule coordinates, nodule location. This paper mainly classifies the pulmonary nodules according to whether the pulmonary exist or not, that is, pulmonary images with pulmonary nodules are classified as the abnormal samples, while those without pulmonary nodules are classified as the normal samples.

The neural network model based on transfer learning needs to take into account the size of the new data set and the original data set and the similarity between the two data sets.
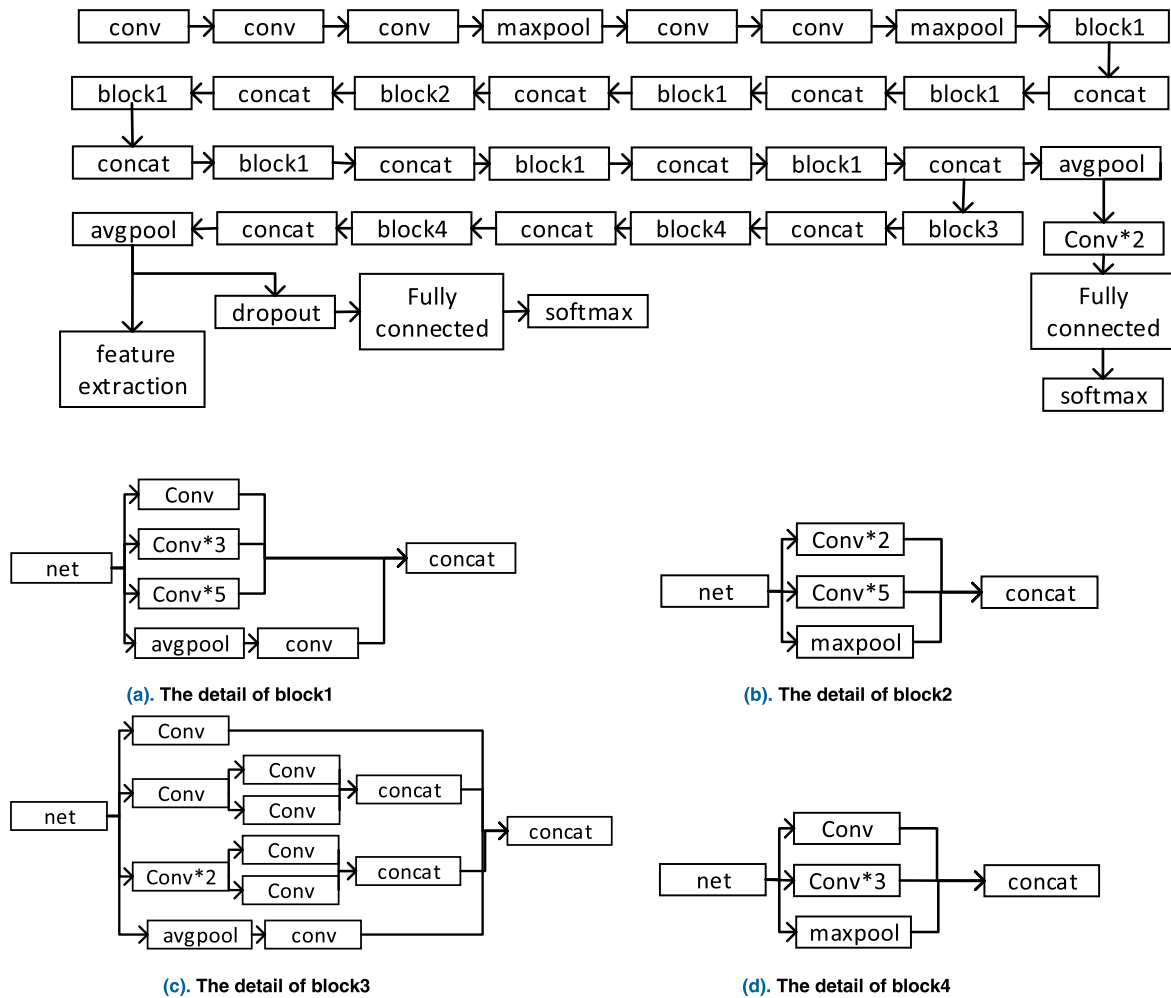
FIGURE 1. The structural schematic diagram of the Inception-v3 model. The detail of block 1~4 are shown in Figure 1(a), Figure 1(b), Figure 1(c) and Figure 1(d).

Considering 247 Pulmonary images in the JSRT database used in this paper, the amount of data needed for the neural network model is far from enough. Considering the relevance of training tasks, we choose the weight obtained by training Inception-v3 model on the glioma data set of TCIA as the initial weight.

B. THE METHOD OF EVALUATION

This paper uses the method of ten-fold cross-validation to verify the experimental results. Ten-fold cross-validation is to divide the data set into ten parts at the same size, take one part in sequence each time as the test set, the rest as the training set, and finally take the average of ten results as the estimate of the metrics of the model. The steps of ten-fold cross-validation:

(a). Dividing the dataset into ten parts on average. Each part contains an equal number of images.

(b). For each training, One part was selected in sequence as the test set, and the remaining dataset was used as model input to train the model.

(c). After training, true-positive, false-positive, true-negative, false-negative of the test set were recorded successively.

(d). Averaging the results of ten records to obtain the average sensitivity and average specificity of the final test results.

C. METRICS FOR EVALUATION

In medical terms, positive means that the patient has abnormal lesions or has the virus, while true-positive (TP) means that patients with abnormal lesions are tested according to a pre-defined test screening method, and they are correctly labeled as abnormal lesions. False-positive (FP) is defined as a medical misdiagnosis of a patient without abnormal lesions who are wrongly labeled as a patient with abnormal lesions or a sample with virus, which results in medical misdiagnosis. Negative means that the patient is normal, and true-negative (TN) is the patient without abnormal lesions and is diagnosed as normal. False-negative (FN) is defined as a condition in which a patient with an abnormal lesion is labeled as
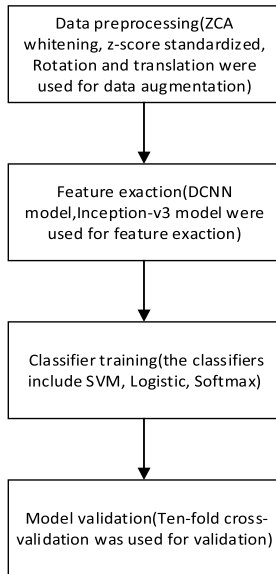
FIGURE 2. The process of method.

		Actual Value		
		True nodule	**False nodule**	**Total**
Predicted Value	**True nodule**	TP=19344 96.5%	FN=6219 27.6%	30221
	False nodule	FP=5528 3.5%	TN=16333 72.4%	17203
	Total	24872	22552	47424

FIGURE 3. Confusion matrix.

a normal sample, which is a condition that causes missed diagnosis.

Sensitivity, also known as true positive rate, is the probability that patients with abnormal lesions or viruses will be diagnosed as medically positive, and this value needs to be as large as possible. The higher the value, the more sensitive the diagnosis is. The formula of sensitivity is shown in (7):

$$SENSITIVITY = \frac{TP}{TP + FN} \qquad (7)$$

Specificity, also known as true-negative rate, is the probability that people without abnormal lesions are diagnosed as negative. Similarly, this value needs to be as large as possible. That is to say, the higher the value, the more accurate the diagnosis is. The formula of specificity is shown in (8):

$$SPECIFICITY = \frac{TN}{TN + FP} \qquad (8)$$

Accuracy is the ratio of the number of true samples to the total number of samples. The formula of accuracy is shown in (9):

$$ACCURACY = \frac{TN + TP}{T + F} \qquad (9)$$

Sensitivity and specificity are the most accurate and scientific evaluation criteria for medical image evaluation in medicine. The higher the sensitivity is, the lower the rate of missed diagnosis in this model; the higher the specificity is, the lower the misdiagnosis rate is. However, as sensitivity increases, specificity may decrease, so it is necessary to weigh sensitivity and specificity. The results of sensitivity, specificity and accuracy in this experiment are both experimental results with relatively high scores after many experiments.

D. RESULT

In the experiment of Pulmonary image classification based on DCNN model, we fine-tune the parameters of DCNN model for many times and input the features extracted from DCNN model into different classifiers(SVM, Logistic, Softmax) for classification. Finally, the experimental results of average sensitivity and average specificity were compared with other methods. Experimental results with SVM and results compared with other experimental results on JSRT dataset are shown in table 4.

In the next optimization experiment, this paper uses the structure of GoogleNet for reference and uses the Inception-v3 model for transfer learning. The Inception-v3 model is the same as the DCNN model that three different classifiers (SVM, Logistic, Softmax) are used in the experiment. The experimental results are shown in table IV. Moreover, the confusion matrix of inception-v3 model for result is shown in figure 3.

E. RESULT ANALYSIS AND DISCUSSION

As shown in table 4, It can be seen that compared with other non-hand-crafted feature extraction methods, DCNN model has higher specificity and sensitivity, while its effect is lower than some hand-crafted feature extraction methods. However, DCNN can automatically extract features so that it can save a lot of time and labor. Moreover, The sensitivity and specificity of Softmax, Logistic and SVM connecting the fine-tuning Inception-v3 model increased by about 2% compared with the experimental results of DCNN. Although the Inception-v3 model was trained on the ImageNet dataset, compared with the DCNN model without transfer learning, its average accuracy has been further improved and ran faster. Experiments that fine-tune the Inception-v3 model connected different classifiers have resulted in significantly better results. The classification errors mainly lie in:

(a).For migrating large-scale neural networks such as inception-v3, Because the network is only fine-tuned, the extracted features are not enough, which leads to some image classification errors.

(b).Only three classifiers are selected to compare the result in this paper, and the selection of classifiers will affect the classification error.

TABLE 4. Comparison of experimental results of different models.

Feature extraction method	Classification model	Sensitivity	Specificity	Accuracy
Automatic Localization of Heart and Lung Shadows [30]	SVM	77.1 %	76.4%	76.80%
Hand-crafted feature with oriented gaussian derivatives filter [31]	Fuzzy clustering method (FCM)	92.79%	97.07%	94.70%
Watershed segmentation [32]	SVM	55.74 %	35.81%	59.80%
Hand-crafted feature with a simple multiscale method [1]	SVM	97.74 %	99.37%	98.50%
DCNN model	Softmax	92.91%	78.39%	84.70%
DCNN model	Logistic	94.30%	79.31%	85.80%
DCNN model	SVM	91.11%	77.27%	83.30%
Fine-tuning Inception-v3 model	Softmax	94.71%	80.09%	86.40%
Fine-tuning Inception-v3 model	Logistic	95.41%	77.78%	85.10%
Fine-tuning Inception-v3 model	SVM	92.96%	79.83%	85.70%

IV. CONCLUSION

This paper proposed a method of lung image classification based on inception-v3 transfer learning in CT images, and the method was compared with other methods. In particular, the method of lung image classification based on migration learning can achieve higher accuracy. Moreover, the neural network model based on transfer learning performs better in pulmonary image classification on JSRT database than the model based on original DCNN. Fine-tuning Inception-v3 transfer learning can effectively improve the accuracy of lung image classification, so the model for pulmonary imaging on the JSRT dataset can provide an effective and rigorous computer-assisted diagnostic when lung image data is insufficiency.

However, if the network selection of transfer learning is inappropriate, the problem of negative transfer may occur, which will lead to the decrease of accuracy and the increase of training time. Therefore, how to select the appropriate network better for lung image tasks is the further research direction. Moreover, this method is only for lung tasks and is not suitable for other medical images so that the research for other medical images can be further studied.

For the problem of the large gap between the specificity and sensitivity of deep learning model in this study, we found that ensemble learning has a good performance in medical images from many pieces of literature, Their sensitivity and specificity are both high. Therefore, we will make further research on ensemble learning in future research work.

REFERENCES

[1] P. C. S and H. S. C, "Automatic detection of pulmonary tuberculosis using image processing techniques," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Chennai, India, Mar. 2016, pp. 798–802.

[2] M. A. Bruno, E. A. Walker, and H. H. Abujudeh, "Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction," *Radiographics*, vol. 35, no. 6, pp. 1668–1676, 2015.

[3] M. Woźniak and D. Połap, "Bio-inspired methods modeled for respiratory disease detection from medical images," *Swarm Evol. Comput.*, vol. 41, pp. 69–96, Aug. 2018.

[4] Q. Ke, J. Zhang, W. Wei, D. Połap, M. Woźniak, L. Kośmider, and R. Damaševičius, "A neuro-heuristic approach for recognition of lung diseases from X-ray images," *Expert Syst. Appl.*, vol. 126, pp. 218–232, Jul. 2019.

[5] P. P. R. Filho, A. C. da Silva Barros, G. L. B. Ramalho, C. R. Pereira, J. P. Papa, V. H. C. de Albuquerque, and J. M. R. S. Tavares, "Automated recognition of lung diseases in CT images based on the optimum-path forest classifier," *Neural Comput. Appl.*, vol. 31, no. 2, pp. 901–914, Feb. 2017.

[6] M. Woźniak, D. Połap, G. Capizzi, G. L. Sciuto, L. Kośmider, and K. Frankiewicz, "Small lung nodules detection based on local variance analysis and probabilistic neural network," *Comput. Methods Programs Biomed.*, vol. 161, pp. 173–180, Jul. 2018.

[7] X. Li, L. Shen, and S. Luo, "A solitary feature-based lung nodule detection approach for chest X-ray radiographs," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 2, pp. 516–524, Mar. 2018.

[8] C. Supanta, G. Kemper, and C. del Carpio, "An algorithm for feature extraction and detection of pulmonary nodules in digital radiographic images," in *Proc. IEEE Int. Conf. Automat./XXIII Congr. Chilean Assoc. Autom. Control (ICA-ACCA)*, Concepcion, Chile, Oct. 2018, pp. 1–5.

[9] J. Mukherjee, A. Chakrabarti, S. H. Shaikh, and M. Kar, "Automatic detection and classification of solitary pulmonary nodules from lung CT images," in *Proc. 4th Int. Conf. Emerg. Appl. Inf. Technol.*, Kolkata, India, Dec. 2014, pp. 294–299.

[10] X. Jiang, Y. Pang, M. Sun, and X. Li, "Cascaded subpatch networks for effective CNNs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 7, pp. 2684–2694, Jul. 2018.

[11] L. N. A and J. J. B, "A computer aided diagnosis for detection and classification of lung nodules," in *Proc. IEEE 9th Int. Conf. Intell. Syst. Control (ISCO)*, Coimbatore, India, Jan. 2015, pp. 1–5.

[12] H. C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. M. Summers, "Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2497–2506.

[13] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, New York, NY, USA, Apr. 2015, pp. 294–297.

[14] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, Apr. 1980.

[15] Y. LeCun, B. Boser, J. S. Denker, Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Stateline, NV, USA, 2012, pp. 1097–1105.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, Apr. 2015, pp. 1–14.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. A. Vanhoucke, and V. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[19] D. Li, K. He, J. Sun, and K. Zhou, "A geodesic-preserving method for image warping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 213–221.

[20] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 8697–8710.

[21] F. Agostinelli, M. Hoffman, P. Sadowski, and P. Baldi, "Learning activation functions to improve deep neural networks," 2014, *arXiv:1412.6830*. [Online]. Available: https://arxiv.org/abs/1412.6830

[22] E. Yazan and M. F. Talu, "Comparison of the stochastic gradient descent based optimization techniques," in *Proc. Int. Artif. Intell. Data Process. Symp. (IDAP)*, Malatya, Turkey, Sep. 2017, pp. 1–5.

[23] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu, "Boosting for transfer learning," in *Proc. 24th Int. Conf. Mach. Learn.*, Corvalis, OR, USA, Jun. 2007, pp. 193–200.

[24] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," 2015, *arXiv:1502.02791*. [Online]. Available: https://arxiv.org/abs/1502.02791

[25] Z. Cao, M. Long, J. Wang, and M. I. Jordan, "Partial transfer learning with selective adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 2724–2732.

[26] P. P. Busto and J. Gall, "Open set domain adaptation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 754–763.

[27] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 248–255.

[28] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826.

[29] J. Shiraishi, S. Katsuragawa, and J. Ikezoe, "Development of a digital image database for chest radiographs with and without a lung nodule," *Amer. J. Roentgenol.*, vol. 174, no. 1, pp. 71–74, Jan. 2000.

[30] S. Candemir, S. Jaeger, W. Lin, Z. Xue, S. Antani, and G. Thoma, "Automatic heart localization and radiographic index computation in chest x-rays," *Proc. SPIE*, vol. 9785, Mar. 2016, Art. no. 978517.

[31] B. van Ginneken, S. Katsuragawa, B. M. T. H. Romeny, K. Doi, and M. A. Viergever, "Automatic detection of abnormalities in chest radiographs using local texture analysis," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 139–149, Feb. 2002.

[32] P. Campadelli, E. Casiraghi, and D. Artioli, "A fully automated method for lung nodule detection from postero-anterior chest radiographs," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1588–1603, Dec. 2006.

DELEI CHEN received the B.S. degree in biology from Zhengzhou University, Zhengzhou, China, in 2016. He is currently pursuing the master's degree in computer science with Huaqiao University, Xiamen. His main research interests include data mining and machine learning.

LIN HAO is a Medical Master Student with Tongji Hospital, Tongji University, Shanghai, China. His research interests include pathophysiology and interventional treatment of CHD, image reconstruction of coronary stent, and 3-D visualization of coronary stent.

XUEBO LIU received the M.D. and Ph.D. degrees from the Medical College, Fudan University, Shanghai, China. He is currently a Full Professor with Tongji Hospital, Tongji University, Shanghai. His research interests include interventional treatment of CHD, image reconstruction of coronary stent, and 3-D visualization of coronary stent. He has published more than 60 articles in scientific journals in these related areas.

YU ZENG received the B.S. degree in software engineering from Huaqiao University, China. He has been interested in deep learning for three years. His research interests include interdisciplinary study of medical science and deep learning.

JIANWEI CHEN is currently a Professor with the Department of Mathematics and Statistics, San Diego State University, USA. His research interests include data mining and deep learning.

CHENG WANG received the Ph.D. degree in mechanics from Xi'an Jiaotong University, China, in 2012. He is currently an Associate Professor and a master's degree Supervisor with Huaqiao University, China. His research interests include signal processing and artificial intelligence.

GUOKAI ZHANG is currently pursuing the Ph.D. degree with the College of Software Engineering, Tongji University, Shanghai, China. His current research interests include deep learning, object detection, and medical image analysis.

• • •