



Pulmonary DR Image Anomaly Detection Based on Deep Learning

Zhendong Song, Lei Fan, Dong Huang, and Xiaoyi Feng^(✉)

Northwestern Polytechnical University, Shaanxi, China
secszd@163.com, {dear, huangdong137}@mail.nwpu.edu.cn,
fengxiao@nwpu.edu.cn

Abstract. The morbidity and mortality in lung cancer is increasing which makes the diagnosis of abnormal lungs particularly important. Because of the advantages in DR image, this paper aimed at two problems in current medical image research: first, it is difficult to completely segment the lung of DR image only used traditional image segmentation methods. This paper replaces the padding in the U-net network model with zero padding to maintain the image size and apply it to the lung DR image segmentation, and finally uses the lung DR image dataset to fine-tuning. Secondly, the results of anomaly detection experiments show that the algorithm would get more complete segmentation of lung DR images. Secondly, because of the insufficient of training set, the idea of multi-classifier fusion is used. Combining Gabor-based SVM classification, 3D convolutional neural network, and transfer learning to achieve a more complete description of features and make full use of the classification advantages of multi-classifiers. The experimental results show that the classification accuracy of this algorithm is 6% higher than that of the Transfer-ImageNet algorithm, 5% higher than SVM, 15% higher than 3D convolutional neural network, and improved 2.5% compared with FT-Transfer-DenseNet3D algorithm.

Keywords: DR image · SVM · Transfer learning · 3D convolutional neural network · Lung field segmentation

1 Introduction

1.1 Research Background and Significance

With the changes in the global environment, the air quality is getting worse. At the same time, the improvement of living standards and the faster pace of work make life more convenient. The resulting irregularities in life is increasingly detrimental to people's health. According to the latest data, 546,259 cases of tracheal, bronchial and lung cancer deaths occurred in China in 2013, accounting for one-third of the number of deaths of this type of cancer worldwide. According to the World Health Organization's research, between 2008 and 2018, the incidence of cancer in the world is rising linearly. At the same time, the probability of patients surviving within 5 years is only

6–13%. The reason to this phenomenon is that 65% of patients are already in precancerous stage. If the symptoms of early stage lung cancer can be found through related techniques, there will be a great survival rate within 5 years. Therefore, the detection rate of early lung cancer can greatly improve the treatment of thousands of patients, and can also reduce the medical burden of patients with advanced cancer.

At present, the lung image can be said to be an extremely important and common way in the medical diagnosis methods. The doctor can find the abnormal part to carry out further diagnosis and treatment. DR imaging and CT tomography are the two main methods of pulmonary diagnosis. Among them, DR images have less radiation and higher image quality than other X-ray images [7], and the price is cheaper. Therefore, automatic reading through DR images will have a larger application space.

In terms of edge detection and contour extraction, Yao et al. proposed a method for segmentation of lung parenchyma from coarse to improve the segmentation accuracy [3]. In this method, 246 high-quality labeled DR images are used as data sets. Wang et al. proposed a DR chest slice contour extraction method based on fully convolution network [14]. Chernuhi et al. used edge detection and target extraction for DR images, proposed a target detection and edge extraction method for X-ray medical images, and represented the object boundaries with vectors [1]. Candemir et al. implemented an automatic lung field segmentation algorithm that has a good segmentation effect on non-rigid registration [6]. The overall accuracy of the JSRT dataset is 95.4%. Luo et al. considered that X chest radiographs have different regional features and edge features in pneumoconiosis, which makes the segmentation results not accurate enough [14]. An algorithm combining wavelet transform and Snake model algorithm is proposed. The experimental results show that the proposed method has a high recognition rate and stable recognition effect.

In terms of image enhancement and noise suppression, Feng et al. described an image enhancement technique CWGCE based on neighborhood contrast and wavelet transform coefficients, which can adaptively enhance edge details and effectively suppress noise [2]. Du et al. proposed a cell membrane optimization algorithm to calculate the spatial position of the optimal pixel in the image for the disadvantages of sharp noise, improper exposure, thick tissue and uneven distribution in the DR image [13]. Xu Yanli et al. constructed a Gaussian pyramid and a Laplacian pyramid, used a specific function to adjust its coefficients, and repeatedly extended the image. Finally, the results were added to reconstruct the original image to enhance the detail of image [12].

Hong et al. divided 100 test samples into 50 normal and 50 abnormal, and used SVM classifier to train the training set of chest DR images to obtain decision function [4]. Khatami et al. proposed a three-step framework for multi-category X-ray image classification. The denoising techniques based on wavelet transform (WT) and Kolmogorov Smirnov (KS) statistics are first used to remove noise and insignificant features in the image; then unsupervised deep belief network (DBN) is used to learn

unlabeled features; The more descriptive features serve as input to the classifier [5]. Torrents-Barrena et al. used a method to automatically analyze X-ray images based on tissue density to determine the presence of normal and suspicious regions in the breast [14]. Luo Haifeng et al. proposed a method for feature extraction and classification by combining gray level co-occurrence matrix with BP neural network [10]. The average accuracy of the network can reach 68.3%. Li Bo et al. designed a relatively independent classification framework by combining the grayscale features, texture features, shape features and features extracted from the frequency domain with medical image segmentation methods. The accuracy of traditional medical image classification methods has been significantly improved [11]. Song Yuqing et al. proposed a classification method based on feature-level data fusion and decision-level data fusion based on the shortcomings of any feature that could not correctly express medical images [2].

In summary, medical image still has great prospects in the field of artificial intelligence. However, the main problems at present are: first, the DR medical images in the lungs are difficult to obtain, and there is no system for automatic labeling. Second, in the difficulty of image segmentation algorithms, the morphology of the lungs is different under different disease conditions. The manual extraction method such as template matching cannot find the edge lines in every cases. Thirdly, in the anomaly detection algorithm, manual extraction of features and the construction of classifiers through machine learning algorithms usually can be used because of small sample of the dataset. On the other side, the deep techniques have been used in many fields successfully and achieved promising results [15, 16]. Therefore, this paper studies the application of automatic labeling algorithm and deep learning used in lung field segmentation and anomaly detection.

2 Lung DR Image Segmentation

Because it is impossible to obtain clear and complete segmentation results while retaining the lung texture. Considering the good segmentation effect of U-net network for segmentation of medical cells under a magnifying glass, U-net network is used in this paper. First, gradation and histogram equalization for the low contrast of the lung DR image in some cases; second, minimize the lung image in order to preserve the lung information as much as possible.

ChinaSet_AllFiles dataset construct by the National Medical Library of Malan State and the Third People's Hospital of Shenzhen Medical College. A chest X-ray database created in Montgomery County, the NLM-MontgomeryCXRSet dataset created by the National Library of America and the Montgomery County Department of Health and Human Services in Maryland. DR_date data set collected by one hospital in Xi'an. CT images dataset which is published in Kaggle competition.

2.1 Data Preprocessing

Three data sets used in this paper including lung DR images, CT images, and ImageNet data sets. The overall DR image data included 700 cases of normal cases and 700 cases of tuberculosis, some of them are shown in Fig. 1.

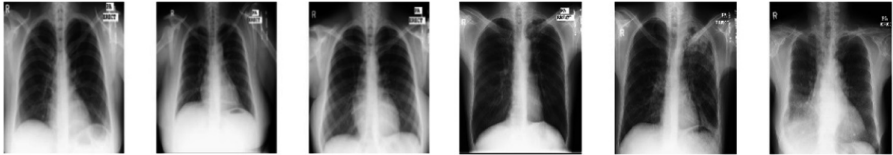


Fig. 1. Normal dataset and abnormal dataset sample

Data Enhancement

On the one hand, some of the annotated data is translated, rotated, and cut, to increase the quantity and diversity of data, as shown in Fig. 2. On the other hand, semi-supervising the unlabeled samples. After data enhancement processing, the sample size of the data has increased by more than 200.

Histogram Equalization

For each image, different gray levels have different total pixels. After histogram equalization, the input pixel distribution of the original input image can be changed, so that the pixel numbers of each gray level are different, so that the gray histogram of the image is smoothed.

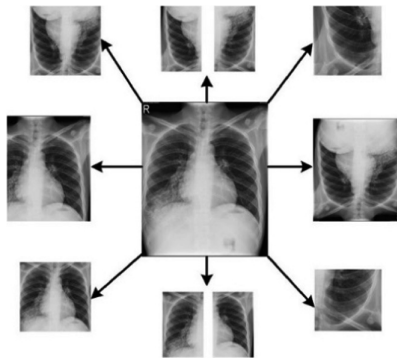


Fig. 2. Translated, rotated, and cut some of the annotated data

Suppose an image has a gray level of g , the range is $[(0, 1)]$, and passes through the mapping function s . For g in any range, according to the form of the transformation:

$$s = T(g), \quad 0 \leq g \leq 1 \tag{1}$$

The gray level of the lung DR image is set to a random number of $[0, 1]$, expressed as a probability density function (PDF). Probability density function based on random variables r and s , the probability density function can be obtained:

$$P_s(s) = P_r(r) \left| \frac{dr}{ds} \right| \tag{2}$$

The histogram of the image is obtained:

$$P_s(s)ds = P_r(r)dr \tag{3}$$

If $P_s(s) = s$, c is a constant and the histogram equalization formula is obtained:

$$\int_0^s sds = \int_0^r P_r(r)dr \Rightarrow s = \frac{1}{c} \int_0^r P_r(x)dx \tag{4}$$

As shown in Fig. 3, the second one is the histogram of the original image. It can be seen that the gray scale distribution of the whole image is concentrated at about 200, and the gray portion of the bright portion is less distributed; the third one shows the lung after the histogram equalization of the DR image, it can be seen from the histogram that the gray level of the entire image is still concentrated at 200, but the number of gray values of the bright portion is increased, and the overall gray scale distribution is enhanced, so that the image is enhanced. At this point, the image has a higher contrast, so the details can be displayed.

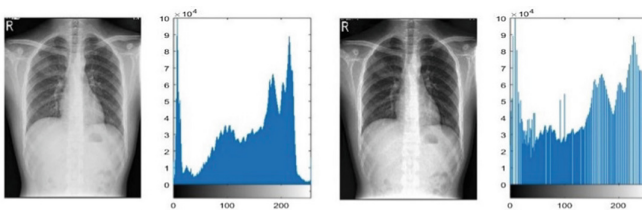


Fig. 3. The first and second pictures are the histograms and it's original image, and the last two are the histograms after the histogram equalization and it's image.

2.2 Lung Field Segmentation Based on U-Net Network

In this paper, the lung field segmentation of lung DR images based on U-net network is mainly divided into three processes. First, grayscale the image. Then perform histogram equalization on grayscale images to denoise and enhance image contrast. Finally, the segmentation result of the DR image is obtained by training the U-net network.

As shown in Fig. 4, the U-net network set a 2×2 pooling at layers with even sizes, which can keep the same size [9]. This architecture, the so-called “full convolutional network”, can get more accurate segmentation result with fewer training images after modified and extended. The main idea is to supplement the usual shrinking network by successive layers, where the pooling operator is replaced by the upsampling operator. Therefore, these layers increase the output resolution, combining high resolution features from the contracted path with the upsampled output. Finally, the convolutional layer can learn more accurate output based on this network learning. An important advantage in this architecture is that in the upsampling section, there are a large number of feature channels that allow the network to propagate context information to higher resolution layers. So the expanded portion and the contracted portion have symmetrical characteristics, forming a U-shaped structure. Since the full connection layer is not used in the entire network architecture, only the effective part of the convolution is used, so the result of the segmentation can only obtain the image pixel portion, it has the connections in the upper and lower pixel. By overlapping block strategies, it can seamlessly segmenting arbitrarily large images. In order to predict pixels in the boundary region of the image, the missing pixel portion is extrapolated by mirroring the input image. This tiling strategy can make the network used in large images, so that resolution is no longer limited by GPU memory.

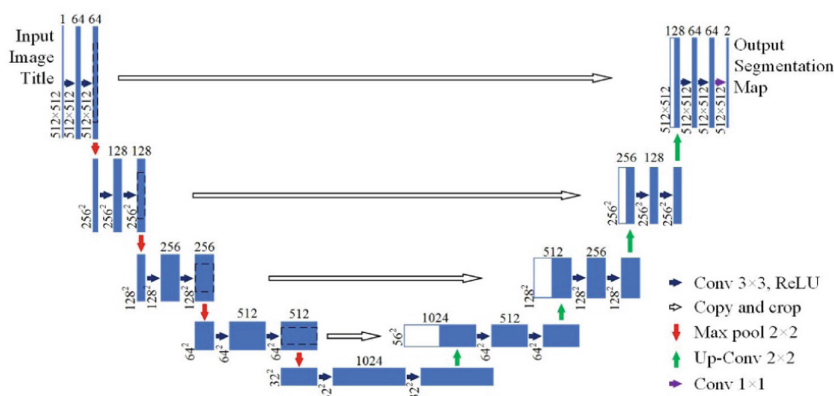


Fig. 4. U-net networks: Each blue box corresponds to a multi-channel feature map. The number of channels is indicated at the top of the box. The x-y size is located on the lower left edge of the box. The white box indicates the copied feature map. The arrows indicate different operations [8]. (Color figure online)



Fig. 5. The image of the lung DR image and the label is divided according to the Grab cut edge detection, then combined U-net network segmentation with the gradient information

The basic U-net architecture is shown in Fig. 5. The network has 23 convolution layers. It consists of a shrink path and an extension path. The entire network architecture is similar to a convolutional neural network architecture. Two such convolutional blocks are constructed from a 3×3 convolutional layer, a linear rectifying unit and a 2×2 pooling layer as a convolutional block, then a downsampling of one step. In each downsampling step, the number of feature channels is doubled. Each step in the extended path upsamples the feature map and then performs a 2×2 up-convolution, halving the number of feature channels in tandem with the corresponding cropped feature map from the shrink path. Finally, a 1×1 convolution is applied to the 64-component feature vector to obtain the corresponding class.

The original image and its segmentation map are used as labels for the training network, and implement TensorFlow's stochastic gradient descent. There are two improvements compared with original network. On the one hand, changing the no-fill in the original convolution network to zero padding keeps the same image boundary width, and the step size is changed to 1. On the other hand, the original network training set is the cell medical image of the human body under the microscope. Since the morphological characteristics of the cells are different from the DR images, the parameters after training in the cell medical image are retained.

2.3 Segmentation Result

In view of the small amount of training data, applying elastic deformation to make data enhancement. Allowing the network to learn the invariance of this deformation without

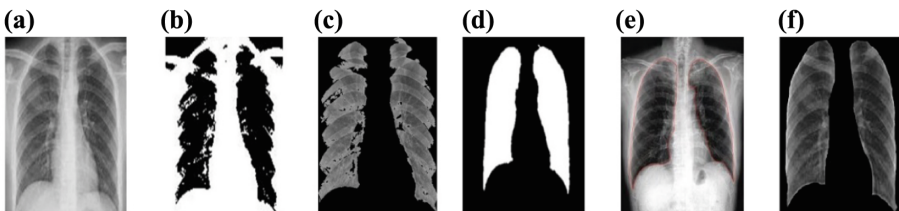


Fig. 6. (a) original image, (b) result based on Otsu; (c) result based on Watershed algorithm; (d) the result based on CNN; (e) the result based on Crab cut; (f) the result based on U-net

having to see these transformations in the annotated image corpus, which is especially important in biomedical segmentation because deformation is the most common change in tissue and can effectively simulate real Deformation.

In order to increase the CPU usage, large input slices are used in the case of large quantities, thereby reducing the batch size to a single image. Therefore, the use of high momentum (0.99) allows a large number of training samples used in the previous iteration to participate in the update in the current optimization step.

As shown in Fig. 6, the original lung texture is preserved and a considerable part of the edge features are also lost in detail; Automatic segmentation of the lung field based on convolutional neural networks shows that the contours of the lungs can be displayed evenly, but the texture of the lungs is not preserved, and small lung nodules will be lost after segmentation, and it is not easy to detect early abnormal state; Grab cut segmentation result shows that the contour of the lungs is very good, and the overlap rate reached 95.9%. The segmentation result of the Grab cut is added to the network training as a training set, and the segmentation is performed in combination with the gradient information. The results show that the contrast of the lungs is deeper, the contours of the lungs are more pronounced, and the lungs can be completely segmented without losing edge information.

3 Anomaly Detection Algorithm for Lung DR Images

This chapter is aimed at the shortcomings of the training set samples in the DR image anomaly detection and the low classification accuracy. Considering the advantage that the model fusion can combine the different classifiers to improve the classification accuracy, based on the improved bagging idea, combined SVM, 3D CNN, and transfer learning. First, establish the Gabor feature matrix, select the radial basis kernel function as the kernel function, train the SVM classifier. Second, use the Inception-v3 model as the transfer object on ImageNet dataset, and fine tune on the DR image data; thirdly, constructing a 3D CNN training on the preprocessed CT data set, and testing on the DR image dataset. Finally, combined these algorithms.

3.1 Model Fusion

Using the idea of bagging to construct three classification models and carry out a certain degree of fusion, and improve on the basis of bagging, first bagging is to form multiple data sets by sampling the same sample set multiple times. The set is set to three relatively large data sets, including the CT dataset, the DR dataset, and the JPG format DR data, and the data of three different data formats are combined to construct the classifier. Selecting a combination strategy of relative majority voting to combine the advantages of three different classifiers to extract different features; secondly, the relative majority vote in bagging randomly selects a result when the results of multiple

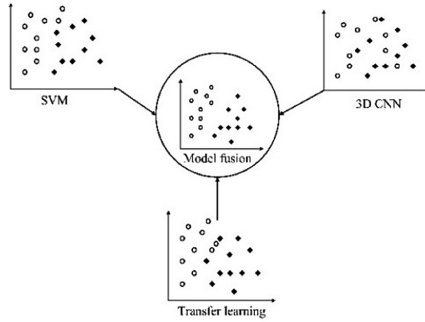


Fig. 7. Model fusion. The decision of the classifier will be combined according to the respective results of these component classifiers using a relative majority voting method.

classifiers are different, so it is proposed in this case Samples that did not receive a majority of the voting results were retrained and predicted again. Get the final abnormal diagnosis result.

Since there are 3 independent data sets, DR data D1 in JPG format, DR data D2 in DIM format, and CT data set D3 in lung. The classifier C1 established by the machine learning SVM is constructed with the three different kinds of data, the classifier C2 established by the transfer learning algorithm and the classifier C3 established by the 3D CNN are generated, finally, The decision of the classifier will be combined according to the respective decision results of these component classifiers using a relative majority voting method. The different training models have different weights under the same classification model. Therefore, it is possible to integrate the connection between different medical data and models to a certain extent to achieve abnormal diagnosis, as shown in Fig. 7.

3.2 SVM-Based Pulmonary Abnormality Detection

The detection algorithm is based on SVM. First, input the original image, normalize the image to a 512×512 matrix size as input. Second, through the threshold segmentation, corrosion, connectivity processing, to obtain the pre-processed image. Third, establish a total of 24° Gabor feature matrices in the four directions of 0° , 45° , 90° , 135° and six scales 7, 9, 11, 13, 15, 17. The frequency determines the wavelength of the Gabor filter, and the direction is determined. The size of the Gaussian window. Fourth, the feature matrix established by the training set sample is taken as the input of the radial basis kernel function map. The penalty parameter of the error term is used as the trade-off between the smoothing of the control decision boundary and the correct classification training point. Set it to 0.01 for SVM classifier training. Finally,

according to the test set picture, the Gabor feature matrix with corresponding dimensionality reduction is established as the classifier input for class prediction, and finally the prediction result is output.

3.3 Pulmonary Abnormality Detection Based on Transfer Learning

Network Construction

Transfer learning can extract and leverage relevant features from existing data to complete new learning tasks. It is possible to adapt a trained model to a new problem with a simple fine-tuning process.

The transfer learning algorithm in this paper uses the trained Inception-v3 model on the ImageNet dataset to solve the problem of abnormal classification of lung DR images. The entire network is centered on the Inception-v3 model. By retaining all the convolutional layer (bottleneck layer) parameters in the trained Inception-v3 model, the DR image is fine-tuned on the latter m layer, replacing only the last layer of the full connection layer, set the number of categories to 2.

Transfer Learning Model

This transfer learning uses Inception-v3 for model migration. The Inception-v3 model is pre-trained on the ImageNet dataset to obtain the pre-training parameters of the network. The parameters in the pre- n -layer region of the pre-training model are transferred to another model. The same position, while the convolutional neural network m layer to fine-tune the DR data set, extract the deep features of the network and then perform anomaly diagnosis, as shown in Fig. 8.

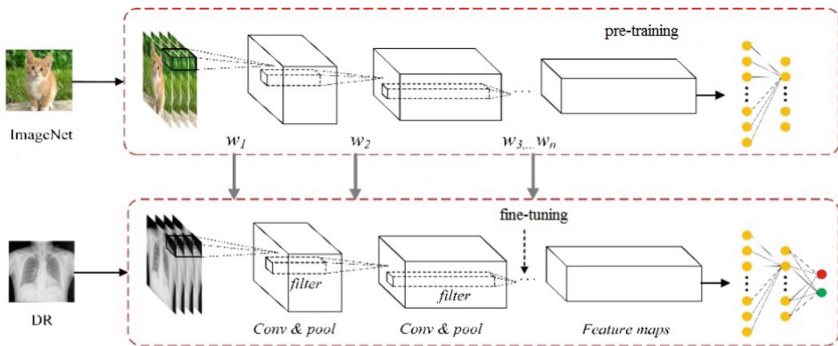


Fig. 8. Transfer learning.

3.4 Pulmonary Abnormality Detection Based on 3D Convolutional Neural Network

The 3D convolutional neural network is built on the CT dataset and anomaly detection is performed. The experimental results show that the feature extraction by convolutional neural network can extract the deep features of the original medical image and assist the anomaly detection. Anomaly detection can be realized by combining deep feature association between CT image data and DR image data.

Model Building

3D convolutional neural network architecture for lung abnormality detection using CT image data sets is described in this section. The 3D convolutional neural network consists of a hardwire layer, 3 convolutional layers, 2 downsampling layers, and a fully connected layer. A 7D image consists of a 3D convolution kernel. As shown in Fig. 49, seven 20×20 frames centered on the current frame are used as inputs to the 3D convolutional neural network model. First, a set of hardwire layers is used to generate multiple channels from the input frame, and 33 feature maps are obtained in the second layer of five different channels. A convolution with a kernel size of $3 \times 3 \times 3$ (space dimension of 3×3 , time dimension of 3) is applied to each channel. Two different convolution operations are used at each location to increase the number of feature maps, and two sets of feature maps are generated in the C2 layer, each group consisting of 23 feature maps. Thereafter, a 3D pooling layer with a kernel size of $2 \times 2 \times 2$ is applied to each feature map in the C2 layer, and spatial resolution is reduced by the same number of feature maps. Next, applying a 3D convolution with a kernel size of $3 \times 3 \times 3$ on each channel of S3, 13 feature maps of size 22×22 can be obtained. Similarly, the 3D pooling layer, with a kernel size of $2 \times 2 \times 4$ applied to each feature map in the C4 layer, can obtain the same number of feature maps with a size of 11×11 as the C4 layer. Finally, a dropout layer with a ratio of 0.8 is added after the normalized S5 layer is 28×28 size. After the completion of the neural network training, a SoftMax classification layer with a learning rate of 0.001 was added for classification (Fig. 9).

After the 3D convolutional neural network is built, the lung CT image is used as the training set, and the preprocessing is performed by unified. After the lung field segmentation, the training is performed as the input of the 3D convolutional neural network, and the training parameters are obtained. Finally, the lung DR image is obtained. The data is used as a test set for abnormal diagnosis.

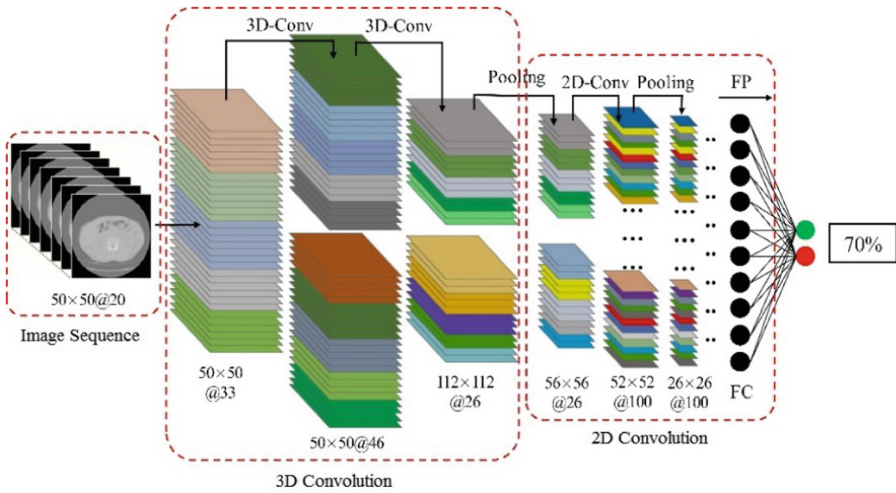


Fig. 9. The structure of 3D CNN

4 Experimental Result

4.1 SVM

The training data is divided into two types of samples, that is, the image in which no lesion occurs in the lung is a positive sample, the label is 1; the abnormal lesion has a negative sample, and the label is -1 . Firstly, the lung image is preprocessed, the lung field is segmented, and the Gabor feature is extracted as the feature matrix. After Gabor feature extraction for each image, PCA is performed on 24 different feature matrices of 6 angles and 6 directions in the Gabor feature. Dimensionality reduction, and finally a matrix vector representation containing the largest information entropy is obtained, as shown in Fig. 43. After training the SVM classifier with the training set feature matrix vector as input, the feature matrix vector constructed by the test set is classified, and the radial basis kernel function is trained by 10-fold cross-validation to predict the result.

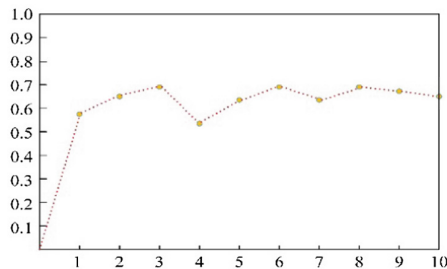
Table 1 shows the results of the existing research. In the case where the sample size is also small, the feature extraction and LS-SVM are combined to identify the pulmonary nodules. The final average accuracy of the experiment is about 60%. Figure 44 the results of the SVM experiment are the results of the algorithm used in this paper, and the graph (a) is the curve of the test accuracy as the number of iterations increases. It can be seen from the figure that the accuracy rate starts to decrease after the 9th iteration. After 10 iterations, the accuracy basically converges, the overall average accuracy rate reaches 80%, and the highest accuracy rate can reach 85%. From the ROC curve, you can see Out, with the increase of the number of iterations, when the false positive is 1, the true positive can reach 89%, and the diagnosis result is ideal. It can make the true positive reach the ideal value when the false positive is 1.

Table 1. Pulmonary nodule detection based on feature extraction and SVM classification.

Number	Kernel function parameter	Penalty factor	Number of samples	Accuracy
1	0.1	10	20	55%
2	0.1	100	20	55%
3	0.1	1000	20	60%
4	0.2	10	20	60%
5	0.2	100	20	65%
6	0.2	1000	20	65%

4.2 3D Convolutional Neural Network

In the experiment, CT images and DR images formed by 1500 patients were used as training data sets and test data sets, and a file containing the data tags was generated to act on the 3D convolutional neural network model of lung image recognition.

**Fig. 10.** The result trained on 3D CNN

First, define the weights and offsets of the convolutional neural network. The first convolution layer is set to $3 \times 3 \times 3$ patches, 1 channel, 32 feature maps; the second convolution layer is set to $3 \times 3 \times 3$ patches, 32 channels, 64 features. The full connection layer has a weight of 1024 and an output weight of 1024. The offset of the first convolutional layer is 32, corresponding to 32 feature maps, the offset of the second convolutional layer is 64, the offset of the fully connected layer is 1024, and the offset of the output is 2. Then define the fit function, the activation function and the loss formula, and optimize them with the Adam Optimizer optimizer. The CT scan data is trained as a training set, and the DR image data set is tested to obtain an accurate result curve, as shown in figure. The results show that after 10 iterations, the 3D convolutional neural network model test on the DR dataset achieve 70% accuracy (Fig. 10).

4.3 Transfer Learning

Firstly, the ImageNet data set is pre-trained on the Inception-v3 model, the trained parameter model is saved, and the parameters of the first 90 convolutional layers are frozen, and the next six convolutional layers are The DR image data set is fine-tuned, and the number of output nodes (number of categories) of the last layer is set to 2, and the final classification result is obtained.

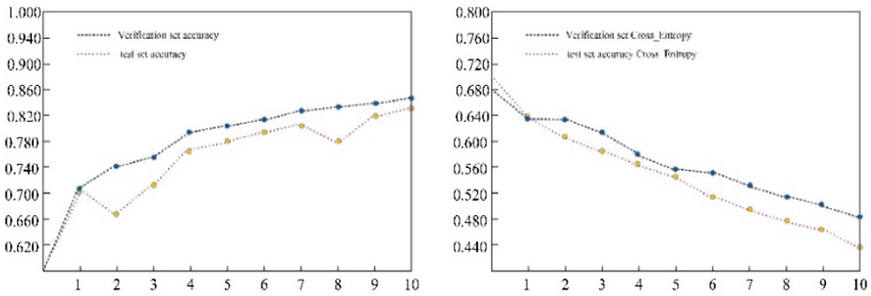


Fig. 11. The result of transfer learning. The x-axis represents the batch_size in units of 100, that is, the iteration number is 10 times, the y-axis represents the cross entropy loss value and the accuracy of the test set, and the red curve and the blue curve represent the result distribution curves of the verification set and the test set (Color figure online)

This paper first proposes the application of migration learning to DR medical image anomaly detection. Figure 11 shows the loss function and accuracy curve of the migration learning network after training. It can be seen the accuracy tends to converge after 10 iterations, and the cross entropy loss function is decreasing, eventually reaching 45%, and the test accuracy of the test sample can reach about 80%.

4.4 Model Fusion

The model fusion results are obtained by model fusion of the above three classifier models, as shown in Table 2. The results show that, first, after the model fusion, the accuracy of the test set of the whole anomaly detection is 5% higher than that of the base classifier, and the final accuracy can reach 85%. The classification accuracy rate is 5% higher than SVM, 15% higher than 3D convolutional neural network, and 6% higher than Transfer-ImageNet; AUC value is 4% higher than SVM. The 3D convolutional neural network increased by 7% compared to Transfer-ImageNet by 9%. Second, the classification accuracy of the SVM algorithm in the experimental results is higher than that of the 3D convolutional neural network, and it is not much different from the Transfer-ImageNet. This is mainly because the classifier has only 700 samples of positive and negative samples on the small sample data. When classifying, the SVM can still map the positive and negative samples according to the distribution of the sample to the corresponding high-dimensional space to find the optimal classification plane. However, in the deep convolution network, the size of the data is proportional to the classification effect. The larger the amount, the higher the classification accuracy. Thirdly, the model migration based on Inception-v3 can achieve fine classification results with SVM after the parameter is fine-tuned, indicating that migration learning

Table 2. Comparison of experimental results in this paper

Algorithm	Accuracy on test set	AUC	Accuracy on verification set
SVM	0.80	0.82	0.88
3D CNN	0.70	0.81	0.85
Transfer-ImageNet	0.79	0.80	0.83
Model fusion	0.85	0.86	0.92

Table 3. Comparison of existing experimental results

Algorithm	Accuracy
SVM	0.800
3D CNN	0.700
Transfer-ResNet3D31	0.782
FT-Transfer-ResNet3D31	0.821
Transfer-DenseNet3D31	0.797
FT-Transfer-DenseNet3D31	0.825
Transfer-Image	0.790
Model fusion	0.850

can achieve better classification effect when dealing with small sample data. And have better generalization ability for lung DR images in different situations. Fourth, the model fusion improves the final classification accuracy by 5% on the basis of the base classifier by combining the feature advantages of different classifiers.

Table 3 shows a comparison of the existing related methods with the results of this paper. The data shows that the existing migration learning algorithm migrates the model parameters on the ResNet3D and DenseNet3D networks by pre-training the relevant data on the 2D convolutional neural network, and the accuracy rate can reach 82.5%. The migration learning was based on the Inception-v3 model, pre-training on the ImageNet dataset, fine-tuning through the lung DR image, and model fusion with the SVM, 3D convolutional neural network. The experimental results show that the accuracy of the model fusion is 6% higher than that of the Transfer-ImageNet algorithm, 5% higher than SVM, and 15% higher than 3D convolutional neural network, and FT-Transfer-DenseNet3D. The algorithm is up 2.5%. This is an increase of 2.9% compared to the FT-Transfer-ResNet3D algorithm.

5 Summary

Based on the relevant theories of deep learning and the starting point, this paper combines traditional machine learning and deep learning to construct the network model. The final anomaly classification accuracy is 6% higher than that of Transfer-

ImageNet. FT-Transfer-DenseNet3D increased by 2.5%, SVM increased by 5%, and 3D convolutional neural network increased by 15%.

There is still a great distance from the research stage to practice. For the future development direction of this research, we can improve from the improvement of DR image dataset, lung field segmentation of lung DR image and lung abnormality detection. Firstly, combined with the doctor's accurate judgment to obtain high-quality annotated data, it is expected to establish a sufficiently complete database; secondly, use a deeper network on larger data; finally, classify and normalize the classification types of lung abnormalities, and convert the two classification networks into multi-class networks, which can detect abnormal types.

References

1. Chernuhin, N.A.: On an approach to object recognition in X-ray medical images and interactive diagnostics process. In: Computer Science and Information Technologies (CSIT), pp. 1–6 (2013)
2. Feng, J., Xiong, N., Shuoben, B.: X-ray image enhancement based on wavelet transform. In: Asia-Pacific Services Computing Conference (APSCC 2008), pp. 1568–1573 (2008)
3. Yao, F., Jun-Feng, W., Lin, G., et al.: Lung field segmentation based on ASM and GrabCut on DR image. *Mod. Comput.* **4**, 25–30 (2015)
4. Hong, S., Tian-yu, N., Yan, K., et al.: Chest DR image classification based on support vector machine. In: Education Technology and Computer Science (ETCS), vol. 1, pp. 170–173 (2010)
5. Khatami, A., Khosravi, A., Nguyen, T., et al.: Medical image analysis using wavelet transform and deep belief networks. *Expert Syst. Appl.* **86**, 190–198 (2017)
6. Pan, S.J., Yang, Q.: A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**(10), 1345–1359 (2010)
7. Zhang, L., Lu, L., Noguees, I., et al.: DeepPap: deep convolutional networks for cervical cell classification. *IEEE J. Biomed. Health Inform.* **21**(6), 1633–1643 (2017)
8. DeSantis, C.E., Ma, J., Goding Sauer, A., et al.: Breast cancer statistics, 2017, racial disparity in mortality by state. *CA: Cancer J. Clin.* **67**(6), 439–448 (2017)
9. Szegedy, C., Vanhoucke, V., Ioffe, S., et al.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)
10. Wu, J., Li, H., Xia, Z., et al.: Screen content image quality assessment based on the most preferred structure feature. *J. Electron. Imaging* **27**(3), 033025 (2018)
11. Fang, M., Yin, J., Zhu, X.: Transfer learning across networks for collective classification. In: International Conference on Data Mining (ICDM), pp. 161–170 (2013)
12. Pan, W.: Transfer learning in collaborative filtering (2012)
13. Saxena, A.: Convolutional neural networks: an illustration in TensorFlow. *Crossroads* **22**(4), 56–58 (2016)

14. Diba, A., Fayyaz, M., Sharma, V., et al.: Temporal 3D ConvNets: new architecture and transfer learning for video classification. In: *Computer Vision and Pattern Recognition*, vol. 9, no. 11, p. 76 (2017)
15. Xia, Z., Lin, J., Feng, X.: Trademark image retrieval via transformation-invariant deep hashing. *J. Vis. Commun. Image Represent.* **59**, 108–116 (2019)
16. Xia, Z., Hong, X., Gao, X., Feng, X., Zhao, G.: Spatiotemporal recurrent convolutional networks for recognizing spontaneous micro-expressions. *IEEE Trans. Multimed.* (2019, online)