A Lung Disease Classification Based on Feature Fusion Convolutional Neural Network with X-ray Image Enhancement

Yue Cheng[*], Jinchao Feng[*] and Kebin Jia[*†]
[*] Faculty of Information Technology, Beijing University of Technology, Beijing, China
[†] Advanced Innovation Center for Future Internet Technology, Beijing University of Technology, Beijing, China
E-mail: ycheng12@emails.bjut.edu.cn Tel: +86-10-67396691
E-mail: fengjc@bjut.edu.cn Tel: +86-10-67396691
E-mail: kebinj@bjut.edu.cn Tel: +86-10-67392529

***Abstract*— With the explosive growth of lung diseases in patients, automatically detecting diseases and obtaining accurate diagnosis through the X-ray medical images become the new research focus in the field of computer science and artificial intelligence to save the significant cost of manual labeling and classifying. However, the quality of common radiograph is not satisfied for the most tasks, and traditional methods are deficient to deal with the massive images. Therefore, we present a feature fusion convolutional neural network (CNN) model to detect pneumothorax from chest X-ray images. Firstly, the pre-processed image samples are enhanced by two methods. Then, a feature fusion CNN model is introduced to combine the Gabor features with the enhanced information extracted from the images and implement the final classification. Comprehensive qualitative and quantitative experiments demonstrate that our proposed model achieve better results in multi-angle views.**

I. Introduction

There has been explosive growth in lung diseases among people over the past years, due to the aggravated environment pollution and unhealthy living habit of human beings, which can present as pneumonia, lung nodule, pneumothorax, edema, fibrosis and other lethal clinical symptom. Pneumothorax as a typical one of them is an abnormal collection of air in the pleural space and result in chest pain and shortness of breath [1]. Pneumothorax mostly caused by some chronic obstructive pulmonary diseases or physical trauma to the chest. It is also named as collapsed lung, which is similar to atelectasis but different. Atelectasis is the collapse or closure of a lung resulting in reduced or absent gas exchange [2]. Diagnosis of the lung disease can be difficult. A chest X-ray images is usually used to confirm its presence. The symptoms of these diseases showed in this kind of medical image can be vague and inconclusive, especially in most cases appearing similar characteristics that quite smaller for visual detection, such as shade, tissue swelling, or vessels gathering, which makes diagnosis hard and burdensome [3]. Along with the medical science development, a large amount of images for diagnosis and treatment have been accumulated from modern hospitals and laboratories for decades, which give the opportunity to exploit computer-aided diagnosis (CAD) techniques to improve the accuracy and efficiency of disease detection and identification [4]. In this case, the quality of acquired images and the precision of image representation are important for detection results [5].

Traditional CAD techniques combine image processing with machine learning classification methods to acquire the feature representation of X-ray images and category results of disease. As the input of the classifiers, the common feature extraction methods of image representation include principle component analysis (PCA), linear discriminant analysis (LDA), Gabor wavelet transformation, and local binary pattern (LBP) [6] to achieve either global features or local features for describing the contour and texture of the images. These methods have one common that using hand-crafted features to represent the lung tissues, which are often fail to reuse on new data or patterns. It is also important to choose the appropriate classifiers to have better performance dealing with these features. K-nearest neighbors (K-NN), naïve Bayesian, random forest and support vector machines (SVM) [7] are classical and frequently used methods. However, traditional image processing and machine learning techniques are inadequate for massive medical images processing.

Deep learning is firstly proposed by Geoffrey Hinton in 2006 [8], which becomes research focus of artificial intelligence rapidly. The primeval deep learning neural network is a deep belief network (DBN) that consists of a group of restricted Boltzmann machines (RBM) to transform adaptively the features from low level to higher complex approximation, which is automatically refine the features through training iterations to extract precise information of the images. Then, the proposal of convolutional neural network (CNN) makes huge progress of computer vision by remarkably improve the performance, which utilizes convolution filters in each layers to implement feature representation and defines a classifier at the top of the network [9].

In this paper, a modified CNN model is proposed for pneumothorax/atelectasis X-ray images classification. Firstly, two kinds of image enhancement methods: multi-scale retinex (MSR) and histogram equalization are used respectively to equipoise the radiography for local contrast enhancement and luminance adjustment. Secondly, a five-layer feature fusion CNN structure is designed to combine enhanced images with feature map extracted from Gabor filters for complex features representation and implement diseases discrimination by optimizer with fine-tuning. Comparing with traditional machine learning methods and basic CNN model, the proposed method achieves better results.

The rest of this paper is organized as follows: Section II describes the image enhancement methods and deep structure in details with Gabor filter features. Section III introduces the evaluation methods and analyzes the experimental results of the different approaches. The conclusions are expressed in Section IV.

II. Deep Structure with Gabor Features Fusion

This section details the procedure of the proposed method. The overall flow-chart is illustrated in Fig. 1. Firstly, it is necessary to extract the region of interest (ROI) which is refer to the lung region out of the chest X-ray images. We manually segment the images and resize them into 64×64 pixels as pre-processing. Secondly, we employ image enhancement to improve local features and apply Gabor filters to capture the contour and shape information, and then combine two feature maps as the input of the neural network. Finally, the softmax classifier is set to get the results of the prediction labels after a fully connection layer.

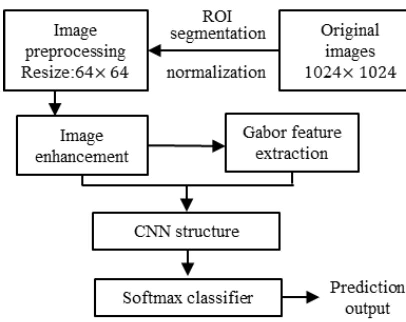


Fig. 1 The overall flow-chart of the proposed method

A. Image Enhancement

Fig.2 shows two radiograph of two lung diseases with partial fuzzy area, luminance disequilibrium, detail deficiency and other problems.

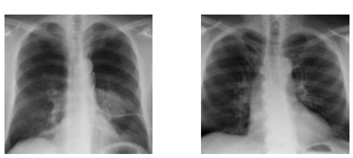
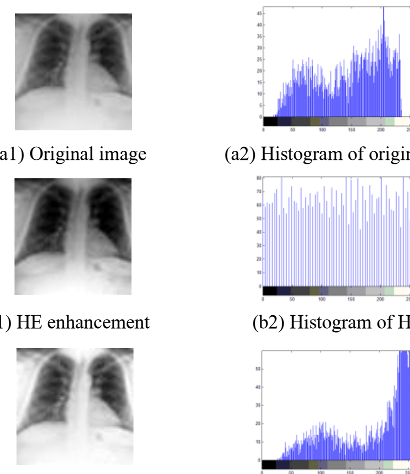


Fig. 2 The samples of radiography image. (left) Pneumothorax; (right) Atelectasis

The images enhancement will result in local contrast improvement and image sharpening with dynamic range compression. The information of hidden area gets enhanced. One common method is histogram equalization (HE) that implies mapping the given histogram distribution of original images to a wider and uniform distribution of intensity values, spreading over the whole range of gray scale to generate the contrast-enhanced images.

Histogram equalization will achieve better performance on the images with single peak in histogram distribution. For the images with lower luminance and wider dynamic region, a method based on multi-scale retinex (MSR) has been proposed in [10]. In MSR, three different Gaussian filter coefficients under three different deviations were calculated. The convolution operation was implemented between image distribution and Gaussian filters to get a weighted average of result mapped to gray range for visualization. The enhanced images comparing with original ones are shown in Fig. 3.



(a1) Original image (a2) Histogram of original image

(b1) HE enhancement (b2) Histogram of HE

(c1) MSR enhancement (b2) Histogram of MSR

Fig. 3 The image enhancement of the original images

B. Feature Fusion CNN model

Gabor wavelet transformation is very similar to visual stimulation response in human visual system. It has good properties to grab spatial domain and frequency domain information of the targets. Gabor filters are sensitive for images contour through the variety of directions and scales, but insensitive to illumination changes, which is beneficial for global feature extraction from the high contrast X-ray images. Combing Gabor features with the enhanced local features after image enhancement can refine the representation of the original images and improve efficiency for CNN structure.

A Gabor filter is a function obtained by modulating the amplitude of a sinusoid with a Gaussian function [11], which is defined as follows:

$$h(x,y) = g(x,y)\cdot \exp(2\pi j\omega x) \tag{1}$$

The feature fusion model firstly merges the Gabor feature maps h(x) with enhanced original matrixes O(x) as two channels of new image samples. Therefore, the convolution process in CNN can be formulated as:

$$\int_{-\infty}^{+\infty} (O(\tau)+h(\tau))\cdot g(x-\tau)d\tau \tag{2}$$

In every CNN layer, the model implements image convolution separately on each channel by depthwise filters and then combines all features by pointwise filters to generate the output channels. The separable convolution use the convolution filters identify with the input depth instead of the output depth to get feature maps, which significant reduces the magnitude of parameters. Then the model sets kernels in accord with output channels to add up each maps. These kernels are defined as depth multiplier to manage the mapping between input and out channels. The process is shown in Fig. 4, where the depth multiplier is set as one.

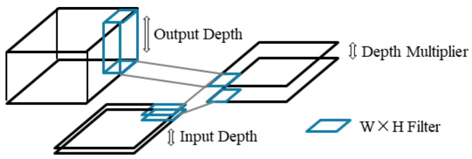


Fig.4 Feature fusion structure

The CNN structure is designed for five convolutional layers with a fully connected layer and softmax classifier to achieve the prediction labels, linked by rectified linear units (ReLUs) [12] activation. The model also includes 2×2 max-pooling for the first two convolutional layers. Each convolutional layer consists of 32 output filters with kernel size of 3×3. The fully connected layer is set of 1024 nodes. Fig. 5 shows the overview of the CNN structure to implement the back-propagation of the gradient in the optimization iterations. We use Adadelta optimizer [13] to minimize the binary cross-entropy loss function with L2 regularization:

$$\begin{aligned} J(\theta) &= C(y', y) = C(f(x;\theta), y) \\ &= -y \cdot \log f(x;\theta) - (1-y) \cdot \log(1 - f(x;\theta)) \end{aligned} \tag{3}$$

where y' indicates the predicted categories of input samples x. θ represents all parameters of the model we trained for. Y is the symbol of the ground truth of sample categories.

As defined above, the proposed Gabor features fusion CNN structure (GF-CNN) is settled for lung diseases pneumothorax and atelectasis classification.

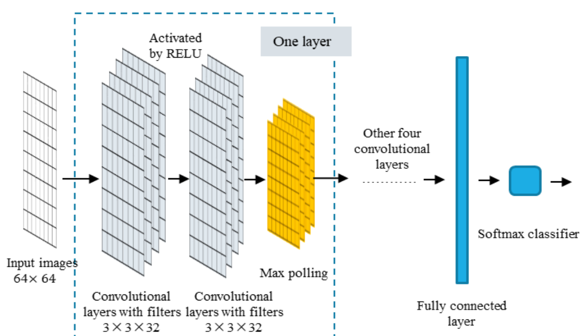


Fig.5 CNN structure

III. Experiment Results

A. Experiment Set

Clinical dataset [14] of chest X-rays published online is used for experimental validation, which consists of 14 lung diseases. For comparing with traditional machine learning methods, we use 2000 images samples (either pneumothorax or atelectasis) to reduce the influence of massive data. The experiment sets 85% data as training samples and the rest 15% as validation samples.

We pre-process the images to capture lung-only area by using a rectangle of size 2.2d×1.2d for location and resize them to 64×64 matrixes, where d is the distance between shoulders. Dimensionality reduction and mean normalization are implemented during the experiments.

For feature extraction, images are convolved by 40 Gabor filters with five scales and eight orientations, then down sampling to 4096 dimension in keeping with the original images for subsequent feature fusion. The weights of neural network are initialized by a normal distribution proposed by Kaiming He [15]. The model sets mini-batches of size 8 and initial learning rate of 1.0 for the optimizer.

B. Results and Analysis

Four criteria including recognition accuracy of all samples, precision, recall and F-1 score of the positive (pneumothorax) samples prediction are chosen to validate the enhancement method HE and MSR. F-1 score is an equilibrium value of precision and recall. Another value is to calculate the area under the curve (AUC) of receiver operating characteristic (ROC) from the positive samples, shown performance of the proposed method GF-CNN. We set the same CNN structure without feature fusion process and a traditional machine learning method K-nearest neighbors (K-NN) classification with Gabor feature extraction as comparing algorithms. The results of image enhancement are shown in table I.

Table I
THE EVALUATION RESULTS OF IMAGE ENHANCEMENT ALGORITHMS

Classification methods	Criteria	Image enhancement		
		None	HE	MSR
Gabor Features with KNN	Accuracy (%)	67.78	68.33	67.78
	Precision	0.74	0.72	0.73
	Recall	0.54	0.59	0.57
	F-1 score	0.63	0.65	0.64
5-layer CNN	Accuracy (%)	76.11	77.78	76.11
	Precision	0.78	0.78	0.84
	Recall	0.72	0.78	0.64
	F-1 score	0.75	0.78	0.73
	AUC	0.79	0.81	0.83

Table I gives the comparison between image enhancements with original images in two classification methods. For the accuracy of all samples, the HE and MSR have slightly advantages of 1% from none-enhanced images. The CNN structure makes a great contribution to improve the accuracy over 10% from the 67.78% of Gabor features with KNN methods. F-1 score reflects the comprehensive performance of precision and recall, which has a little improved of either HE or MSR methods. Moreover, the AUC value is shown that HE achieves 0.02 improvement from the 0.79 of none-enhanced method and MSR gets another 0.02 progress over the HE methods. The results indicate that the image enhancement make improvement of the method performance.

For further evaluating the proposed GF-CNN method, the experiments set deep residual learning ResNet56 [16] as another comparison with the similar parameters setting.

Table II
THE EVALUATION RESULTS OF CNN ALGORITHMS

Image enhancement	Criteria	CNN models		
		ResNet56	CNN	GF-CNN
HE	Accuracy (%)	75.56	**77.78**	76.11
	Precision	0.85	0.78	0.74
	Recall	0.62	0.78	0.81
	F-1 score	0.72	0.78	0.77
	AUC	0.80	0.81	0.78
MSR	Accuracy (%)	71.67	76.11	**77.78**
	Precision	0.73	0.84	0.76
	Recall	0.69	0.64	0.81
	F-1 score	0.71	0.73	**0.79**
	AUC	0.74	**0.83**	0.82

Table II demonstrates the different performance of the ResNet56 and proposed CNN model with or without Gabor fusion. The two CNN networks share the same basic structure but with different channels of the input images. Considering comprehensively of the criteria, GF-CNN based on MSR method achieves the better performance with the highest accuracy of 77.78% and F-1 score of 0.79, also the second best AUC value of 0.82. Comparing with CNN method, GF-CNN performs similar results plus 0.06 improvement of F-1 score on MSR enhancement. The proposed CNN methods both show better results than ResNet56, especially on MSR algorithm. Considering image enhancement methods can be regard as feature representation as well, the images expression after MSR combining with the Gabor features is more suitable for CNN classification rather than HE features. The reason why GF-CNN does not achieve better results on HE method, is possibly for losing the key HE features through Gabor filters.

IV. CONCLUSIONS

In this paper, we present a feature fusion CNN model to distinguish pneumothorax/atelectasis on chest X-ray images. Firstly, a ROI segmentation method and normalization are exploited to reduce the complexity of computation. Secondly, a feature fusion CNN model is introduced to combine Gabor features with enhanced features of the original images to obtain the contour information and local detailed information respectively for lung disease classification. The experiment results demonstrate that our proposed model performs better with enhancement methods and get batter results on feature fusion models.

In the future work, we will try to combine the enhanced images with more other features extracted from the original samples to achieve more accurate image representation and develop more efficient fusion models for classification.

ACKNOWLEDGMENT

This paper is supported by the Project for the National Natural Science Foundation of China under Grants No. 61672064, 81370038, the Beijing Natural Science Foundation under Grant No. 4172001, the Beijing Nova Program under Grant No. Z141101001814107, the China Postdoctoral Science Foundation under Grants No. 2016T90022, 2015T80030, 2015M580029, the Science and Technology Project of Beijing Municipal Education Commission under Grants No. KZ201610005007, and Beijing Laboratory of Advanced Information Networks under Grants No. 040000546617002.

REFERENCES

[1] I Pongrac, S Midzić, M Roglić. “Spontaneous pneumothorax,” *BMJ (Clinical research ed.)*, vol. 302, no. 6776, pp. 595 348, 1981.

[2] M E Wedding, B A Gylys. *Medical Terminology Systems: A Body Systems.* Philadelphia, Pa: F. A. Davis Company. ISBN 0-8036-1289-3, 2005.

[3] R Uppaluri, E A Hoffman, M Sonka, et al. “Computer recognition of regional lung disease patterns,” *American Journal of Respiratory and Critical Care Medicine*, vol. 160, no. 2, pp. 648–654, 1999.

[4] A Kumar, J Kim, W Cai, and D Feng. “Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data,” *J Digit Imaging*, vol. 26, no. 6, pp. 1025–1039, 2013.

[5] H Greenspan, B V Ginneken, R M Summers. “Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1153-1159, 2016.

[6] M Anthimopoulos, S Christodoulidis, A Christ, S Mougiakakou. “Lung pattern classification for interstitial lung diseases using a deep convolutional neural network,” *IEEE Transaction on Medical Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[7] R O Duda, P E Hart, D G Stork. *Pattern classification, 2nd* ed. Wiley-Interscience, 2001.

[8] G E Hinton, and R R Salakhutdinov. “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504 – 507, July 2006.

[9] Y Bengio, A Courville, and P Vincent. “Representation learning: A review and new perspectives,” *IEEE Transaction on Pattern Analysis*, vol. 35, no, 8, pp. 1798–1828, 2013.

[10] T Li, V Asari. “Modified luminance based MSR for fast and efficient image enhancement,” *Applied Imagery Pattern Recognition Workshop*, *Proceedings. 32nd Applied Image Pattern Recognition Workshop*. Washington, DC, USA, pp. 174-179, 2003.

[11] T S Lee. “Image representation using 20 Gabor wavelets,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 959-971, 1996.

[12] A Krizhevsky, I Sutskever, and G E Hinton. “Imagenet classification with deep convolutional neural networks,” *in Adv. Neural Inf. Process. Syst*., vol.25, pp.1097–1105, 2012.

[13] M D. Zeiler. “ADADELTA: An adaptive learning rate method,” http: //arxiv.org/abs/1212.5701, 2012.

[14] X S Wang, Y F Peng, L Lu, et al. “Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2097-2106, 2017.

[15] K M He, X Y Zhang, S Q Ren, J Sun. “Delving deep into rectifiers: surpassing human-level performance on ImageNet classification,” *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1026-1034, 2015.

[16] K M He, X Y Zhang, S Q Ren, J Sun. “Deep Residual Learning for Image Recognition,” *2016 IEEE Conference on Computer Vision & Pattern Recognition (CVPR)*, pp. 770-778, 2016.