Classification of Bacterial and Viral Childhood Pneumonia Using Deep Learning in Chest Radiography

Xianghong Gu[1], Liyan Pan[2], Huiying Liang[2], Ran Yang[1]
[1] School of Data and Computer Science, Affiliated Guangzhou Women and Children's Medical Center, Sun Yat-sen University, Guangzhou, China
[2] Institute of Pediatrics, Affiliated Guangzhou Women and Children's Medical Center, Sun Yat-sen University, Guangzhou, China
e-mail: guxh7@mail2.sysu.edu.cn; panly5@mail2.sysu.edu.cn; lianghuiying@hotmail.com; yangran@mail.sysu.edu.cn

ABSTRACT
Over decades, computer aided diagnosis (CAD) system has been investigated for detection of lung diseases based on chest X-ray images. Incited by the great success of deep learning, in this work, we propose a novel CAD system to identify bacterial and viral pneumonia in chest radiography. The method consists of two parts, lung regions identification and pneumonia category classification. First, left and right lung regions are segmented and extracted with a fully convolutional networks (FCN) model. The model is trained and tested on the open Japanese society of radiological technology database (JSRT, 241 images) and Montgomery County, Md (MC, 138 images) dataset. After segmentation, a deep convolutional neural network (DCNN) model is used to classify the target lung regions. Then, based on the DCNN model, features of the target lung regions are extracted automatically and the performance is compared with that of manual features. Finally, the DCNN features and manual features are fused together and are put into support vector machines (SVM) classifier for binary classification. The proposed method is evaluated on a dataset of Guangzhou Women and Children's Medical Center, China, with 4,513 pediatric patients in total, aged from 1 to 9 years old, during the period from 2003 to 2017. The performances are measured by different criteria: accuracy, precision, sensitivity, specificity and area under the curve (AUC), which is a comprehensive criterion. The experimental results showed better accuracy (0.8048 ± 0.0202) and sensitivity (0.7755 ± 0.0296) in extracting features by DCNN with transfer learning. The values of AUC varied from 0.6937 to 0.8234. And an ensemble of different kinds of features slightly improved the AUC value from 0.8160 ± 0.0162 to 0.8234 ± 0.0014.

CCS Concepts
• Computing methodologies → Classification and regression trees

Keywords
CAD system; deep learning; chest X-rays; fully convolutional networks; convolutional neural networks

ICMIP 2018, March 16-18, 2018, Guiyang, China

ACM ISBN 978-1-4503-6468-3/18/03…$15.00
DOI: https://doi.org/ 10.1145/3195588.3195597

1. INTRODUCTION
According to the report of World Health Organization, pneumonia killed 920,136 children under 5 years old in 2015, accounting for 16% of all the pediatric deaths [1]. Pneumonia is a form of acute respiratory infection that affects the lungs and can be caused by bacteria, virus or fungi [2]. Different causes need different therapies, and medication should be more elaborately prescribed to children than to adults to prevent potentially dangerous side effects. Even though differentiation of bacterial and viral pneumonia in children has been studied in medical area since 1988 [3], it usually depends on chest radiology and laboratory tests, especially sputum culture, which will take a few days. There is no explicit guideline to distinguish bacterial and viral pneumonia in computer aided detection and diagnosis field with image processing techniques. In addition, the presenting features of the two types of pneumonia in chest radiography are similar and confusing. To improve the accuracy and efficiency of diagnosis, developing a CAD system capable of autonomously diagnosing bacterial and viral pneumonia on chest X-rays is most desirable.

Considering the abundant noises caused by rib cage, clavicle, complex lung structure and subtle texture in chest radiography, segmentation of regions of interest (ROIs) or lung areas is a preliminary step for detecting and diagnosing lung disease. In this area, there are many techniques such as level sets [4] and active shape model [5] have been used to segment chest X-rays, like [6]. For instance, [7] adopts an intensity and discontinuity based method to detect the lung boundaries and extract statistical and geometrical features from lung regions. [8] applies a graph cut method to extract lung regions and then computes texture and shape features. Sometimes a hybrid of two or more lung segmentation techniques is also considerable.

Recently, deep learning has gained wide attentions because of its ability of obtaining informative feature representation automatically and has been successfully applied to many fields. It also has been explored in medical image analysis for segmentation and classification tasks [9]. More concretely, fully convolutional networks (FCN) for semantic segmentation [10] is one of the end-to-end fashion deep learning methods that can be utilized for identifying ROIs in medical images [11]. The deep convolutional neural Network (DCNN) has been occupying the leading position due to its nice performance in image classification since it achieved remarkable progress in object recognition for natural images in ImageNet [12]. Brain tumor segmentation at magnetic resonance imaging [13], lung segmentation at computed tomography (CT) [6], lung nodules classification at CT [14], pulmonary tuberculosis auto

Table 1. Chest X-ray Datasets

Data Origin	No. of Images	File Type	Bit Depth	CR/DR	Resolution	Average Age (y)	Bacterial Infection (%)
Japanese society of radiological tech. (JSRT)	241	PNG	12 bit	CR	4096×4096	——	——
Montgomery County, Md (MC)	138	PNG	8 bit	CR	4020×4892	33.1±18.1	——
Guangzhou Women and Children's Medical Center	4,513	DICOM	16 bit	CR+DR	2248-3280×2428-4248	5.5±4.2	59.1 (2,665/4,513)

classification [15] and all other studies show a growing interest in applying deep neural networks in radiology. In this work, we propose a novel computer-aided method to distinct pediatric bacterial and viral pneumonia through chest radiography. First, at pre-processing stage, lung regions are identified and segmented based on FCN model. Once the lung regions are extracted, classification of DCNN method is applied. Then both deep learning features and traditional handcraft features of the lung regions are extracted and are evaluated and compared independently. To enhance the diagnosis performance, an ensemble method of two kinds of features is adopted. The datasets for segmentation are from the public JSRT database [16] and the MC dataset [17]. Classification image data comes from Guangzhou Women and Children's Medical Center.

The rest of this paper is organized as follows. In section II, our method is described in detail, which consists of four steps: lung regions detection and segmentation, classification and high-dimensional feature extraction by DCNN model, features estimation by manual crafted, feature fusing and final classification by SVM with radial basis function (RBF) kernel. The result of proposed method is thoroughly evaluated on our own dataset in section III. Concluding remarks is addressed in section IV.

2. MATERIALS AND METHODS

A. Datasets

In lungs segmentation and extraction stage, two public datasets were used as training and validation set to learn features in the FCN model. One dataset comes from the JSRT database which contains 241 X-ray chest images accompanied by lung masks which is resulted from manual segmentation. JSRT's image data is provided in Portable Network Graphics (PNG) format as 12-bit gray level images whose size is 4096×4096 pixels and is also available in DICOM format. The MC dataset has been collected in collaboration with the Department of Health and Human Services, Montgomery County, Maryland, USA. It has 138 frontal chest X-rays whose size is 4020×4892 pixels in PNG format as 8-bit gray level images.

After extracting ROIs of lung regions, effective image representation and quantification of the target ROIs are key for accurate analysis of chest images. The bacterial and viral childhood pneumonia sample images are offered by Guangzhou Women and Children's Medical Center. We obtained 4,513 frontal chest x-ray images from pediatric patients with age under 10 years old, which include 2,665 bacterial pneumonia images and 1,848 viral pneumonia images. All studied patients were confirmed with pathologic findings of sputum, original authors of the radiology reports. All raw data samples were chosen from and compared with both Sputum Culture System and Picture Archiving and Communication System (PACS) from the year 2003 to 2017. Examples of cross infection were excluded deliberately to screen out the conditions infected only by bacteria or by virus. Feature extraction and classification experiments were conducted on this dataset. Overall and detailed information of all datasets that we used is showed in Table I.

B. Methods

DCNN is a deep learning network with multiple hidden layers that learn hierarchies of filter kernels automatically from a large data set. Due to the lack of large datasets in medical area, there exists a challenge applying deep learning methods on medical image analysis. Hence a special technique "*transfer learning*" based on transferring features learned from one domain to a new domain was proposed to address the problem [18], [19]. Transfer learning using a convolution neural architecture has been proved effective and efficient in medical image feature extraction and classification [18]–[21]. It usually relies on training a large non-medical dataset such as ImageNet or a pre-trained model before starting transfer learning.

Our experiments were conducted on a computer with a Linux operating system (Ubuntu14.04 LTS) and the Caffe deep learning framework (http://caffe.berkeleyvision.org; BVLC, Berkeley, Calif) [22], with CUDA7.5/cuDNN7.5 dependencies for graphics processing unit acceleration. The computer contained an Intel i7 4770k 3.40GHZ processor, and a CUDA-enabled Nvidia Quadro M2000 × 4GB graphics processing unit. As for time cost of deep learning training, it took around 500 milliseconds for each iteration and almost 4 hours to finish classification with the AlexNet-based network. Detailed method is described step by step as follows:

1) Image Segmentation: Fig. 1 demonstrates an overall workflow of our method. In the first step, an AlexNet-based fully convolutional networks model was applied for segmentation of lung regions. The FCN model was pre-trained by PASCAL VOC datasets (http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2012/) which is a challenge for objects recognition with several object classes in realistic scenes [23], containing 11,530 images categorized into 20 classes. The purpose of the segmentation step is to isolate the anatomical regions of interest from medical images. Then transfer learning was accomplished on the described FCN model with the JSRT (241 images) dataset and the MC set (138 images). The two datasets (379 images in total) were randomly split into 234 (61.5%) train set, 75 (20%) validation set, 70 (18.5%) test set. All these images were resized to a $500 \times 500 \times 3$ matrix and were augmented by using mirror flip, rotations of 90°,180° and 270°.

The architecture of AlexNet-based FCN model consists of eight convolutional layers. Different to previous networks, it changes the last three fully connected layers into convolutional layers. With the upsampling skills [10], the output layer makes it possible to achieve end-to-end or pixel-to-pixel prediction. Every pixel will be classified into either the positive class or the negative class. In our experiment, one pixel in positive class represents its existence within the lung regions observing from the image, and in negative class means existing outside of target areas. Since every pixel is predicted, the final output is an image with two masks of left and right lungs, where the output image is of the same size as its input. With the lung masks, ROIs can be easily extracted from original images using mathematical operation. Before starting training the network, our solver parameters were set as: 50 epochs; base learning rate 0.0001 and dropout ratio 0.5 for two dropout layers.

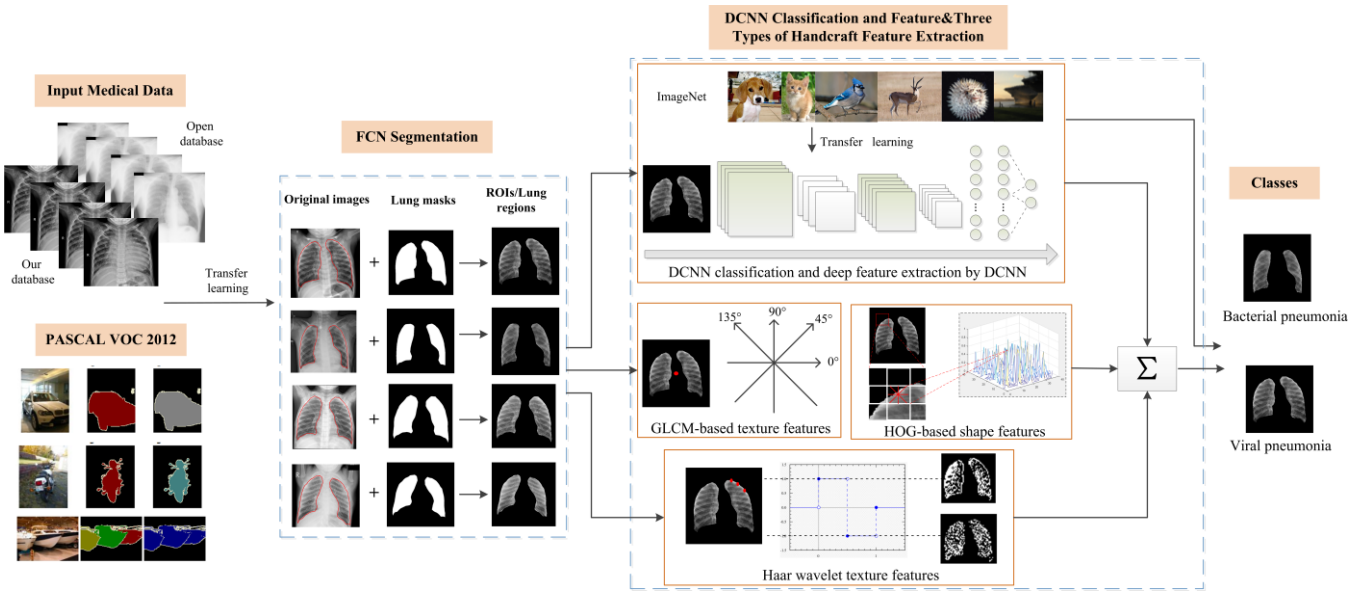

Figure 1. Overall workflow chart. The flow is divided into three parts. First, use an eight-layer FCN model with transfer learning to segment anatomical lung regions. And then train the target images using AlexNet DCNN model to classify test set and extract high-dimensional "fc7" layer features subsequently. GLCM and haar wavelet texture features, and HOG shape feature are meanwhile extracted, then features is classified by SVM classifier respectively and altogether.

2) DCNN classification and DCNN features: During classification stage, the possibility was explored to use the pre-trained DCNN model learned from ImageNet, the large scale of non-medical image database, for the task of medical image analysis. Transfer learning method was fulfilled on target lung regions with AlexNet [12] networks, which comprised of five convolutional layers, two fully connected layers following two dropout layers, and an output layer. All DCNN layers except the last and the penultimate layer were inherited from the trained model in AlexNet for fine-tuning. Of the 4,513 images in our own dataset (Guangzhou Women and Children's Medical Center dataset in Tabel I), 500 random images (11%) were selected for testing. Among the remaining 4013 images, they were split into 80% training set (3211 images) and 20% validation set (802 images). All input ROIs images were resize to 256×256 matrices.

After training 120 epochs, 500 images were tested on the output model. It was an examination of how well our model has learned from the medical dataset. In addition, features of the "fc7" layer (the penultimate layer) in the network structure was taken out to put into a SVM classifier. There are two purposes of DCNN features extraction where one is for the comparison of different features and another is for performance discrimination of DCNN classifier and SVM classifier. Results and analysis can be seen in section III.

The reason why we did not adopt deeper architecture such as 19 layers in VGGNet and 22 layers in GoogleNet was our preliminary experiments performed less satisfactory and cost more time than AlexNet architecture in this paper. The possible explanation for the unsatisfactory performance is due to limited data and overfitting.

3) Texture and shape features extraction: As shown in Fig. 1, three kinds of image features were extracted, i.e., gray level co-concurrence matrix (GLCM) based features [24], haar wavelet transform features [25] and histogram of oriented gradient (HOG) based features [26].

GLCM based features reflects the comprehensive information of statistical texture and is useful in detection and classification tasks. The GLCM is calculated by counting the number of times adjacent pixels having the same orientation. There are in total thirty-two GLCM features for each segmented ROI. Mean value and square root value are computed from eight base features: energy, entropy, homogeneity, contrast, correlation, angular second moment and inverse differential moment.

Wavelet analysis has been proved to be an effective way to generate multi-scale representations of images [25]. Eight level haar wavelets decomposition are selected for each ROI. Then high frequency coefficients in each sub-image are chosen as distribution characteristics of image. As a result, there are in total eight features in each ROI image.

HOG based features has been act as a global shape descriptor while not ignoring the local spatial arrangement of edges or gradients in the images. The main step is to compute cell and block descriptors after image gradient computation. The HOG-based features extraction part shows that the input image is divided into several rectangular cells. Histogram of gradients then is computed and projected into bins for nine directions with every two directions at a $40\,°$ angle. The background of images should be eliminated before computing the gradient to avoid missing value in matrix. When cells computing is done, a group of cells form a block that are normalized to an outcome vector G. We applied L2 norm as follows:

$$G_{out} = \frac{G_{in}}{\sqrt{\|G_{in}\|_2^2 + \varepsilon^2}} \qquad (1)$$

where G_{in} is the input vector and G_{out} is the output vector. And ε is a small regularization constant.

Once these texture and shape features was gained, classification was performed using SVM RBF kernel and 5-fold cross validation respectively. The comparison of results is showed in Table II

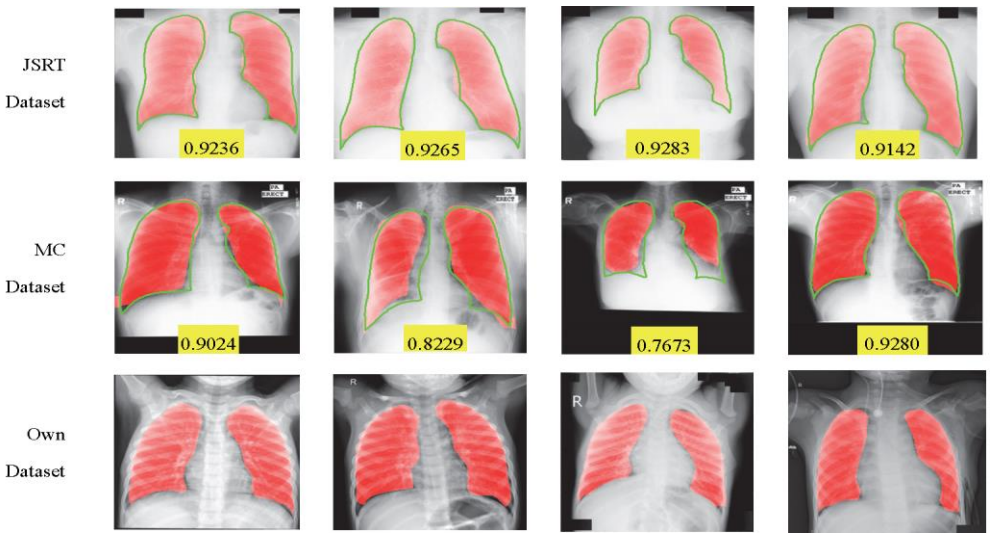

Figure 2. Comparison of segmentation results on the JSRT, MC dataset and our own dataset. At first two rows, the green contour aroud the lungs is ground truth from JSRT and MC manual mask, which is absent in our own dataset at the third row. The value showed on each image is the DSC, which quantilizes the diffrences between ground truth and our predicted mask. For JSRT test set, DSC ranges from 0.9142 to 0.9657 while DSC ranges from 0.7637 to 0.9548 for the MC test set. For all rows, the red masks above lung is a predicted mask segmented by our FCN model.

4) Ensemble of all features: In this step, three types of texture and shape features were spliced into one handcraft feature vector. To exploit the diagnostic potential of both DCNN and handcraft features, the two kinds of feature sets were combined and classified by SVM with RBF kernel. Four measurements were examined: accuracy, precision, specificity, sensitivity, defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

where true positive (TP) is the number of bacterial samples which are correctly classified and true negative (TN) is the number of viral samples that are classified rightly. The false positive (FP) and false negative (FN) are bacterial samples and viral samples are wrongly classified. Another important performance criterion is AUC, varying between 0 and 1, which is a comprehensive standard to evaluate the performance of a classifier. And the closer is AUC value to 1, the better performance it has. For all features, values are standardized: each column has its mean subtracted, and is divided by its standard deviation, as shown in Table III.

Table 2. Comparison of handcraft features

Category	Dimension of features	AUC	CI[a]
GLCM-based Feature	32	0.7060	[0.6388,0.7209]
Haar Wavelet Feature	8	0.6769	[0.6584,0.6931]
HOG-based Feature	274	0.6930	[0.6892,0.7145]

a. CI (Confident Interval) is an interval estimate computed from observed data.

Table 3. Comparison of different methods for classification

	Accuracy	Precision	Specificity	Sensitivity	AUC
GLCM Features	0.7060 ±0.0672	0.6737 ±0.0087	0.8980 ±0.0062	0.6378 ±0.0058	0.7280 ±0.0042
Wavelet Features	0.6769 ±0.0100	**0.8886 ±0.0319**	0.8779 ±0.0205	0.5612 ±0.0065	0.6937 ±0.0072
HOG Features	0.7511 ±0.0127	0.7320 ±0.0678	0.8651 ±0.0664	0.5714 ±0.0617	0.7209 ±0.0036
Handcraft Features	0.7640 ±0.0330	0.7864 ±0.0526	0.8848 ±0.0387	0.6213 ±0.0482	0.7200 ±0.0060
DCNN only	0.7360 ±0.0023	0.7806 ±0.0023	0.7072 ±0.0023	0.6322 ±0.0023	0.7384 ±0.0023
DCNN Features	**0.8048 ±0.0202**	0.7406 ±0.0452	0.8207 ±0.0507	**0.7755 ±0.0296**	0.8160 ±0.0162
Fused Features	0.7692 ±0.0122	0.8308 ±0.0155	**0.9267 ±0.0301**	0.5567 ±0.0379	**0.8234 ±0.0014**

3. RESULTS AND ANALYSIS

For segmentation stage, a FCN model was used with transfer learning to reduce the noises such as shadows of skull, clavicles, backbone and so on. And Fig. 3 shows the training and validation loss during 50 epochs of FCN deep learning model. It reveals a quick convergence to low loss at 0.1 and then the loss goes down at a slow rate. Based on this trained model, 70 test images including 45 images from JSRT database and 25 images from the MC set, were tested. The average computation time for each test image was around 781 milliseconds. Since the ground truths were marked by radiologists in both JSRT and MC dataset, dice similarity coefficient (DSC) values were computed to measure the ability of our segmentation model:

$$DSC = \frac{2*TP}{2*TP + FP + FN} \qquad (6)$$

where TP represents that pixel is predicted rightly, and FP counts when pixel is predicted wrongly to positive one. And FN is defined similarly. The output shows in Fig. 2. In JSRT test set, the average DSC is 0.92, while the average DSC is 0.90 in MC test set. It means that our FCN model performed well on the test set of JSRT and MC. And ROIs of lung regions was then extracted from our own dataset with the well-trained FCN model. The images with red masks at the third row in Fig. 2 are the segmented results.

To accomplish the classification, an AlexNet pre-trained DCNN model was trained on Guangzhou Women and Children's Medical Center dataset. The training curve is showed in Fig. 4. The test accuracy reached 0.99 and the validation loss is relatively lower than training loss over time. The best performed model at the 120 epoch was used on our dataset. Once the transfer learning process is complete, the neurons in the penultimate layer of the network were extracted as feature representation, which is 4096 dimensions, for the tested radiography.

Next, the DCNN features and handcraft features were extracted, and then were evaluated by SVM individually and fused together. Handcraft features were described in Table II and the classification results were summarized in Table III. The AUC ranges from 0.6937 to 0.8234 and the accuracy as well as precision is almost lower than 0.8. Neither classification with handcraft features nor only by DCNN model achieves a significant classification result. Sensitivity is the lowest in the result and means that a part of viral pneumonia is classified to bacterial pneumonia. Features seems not to generate sufficient representations from medical images. One possible reason for the unsatisfactory value of accuracy and sensitivity is for unbalance of our data. Another reason may be overfitting during our training process. Although the overall result is lower than we expected, it shows that DCNN features has better performance than other features. The AUC for DCNN can reach 0.8160, which is obviously higher than 0.7640 for handcraft features. Besides, the accuracy was slightly improved from 0.7360 to 0.8048 after employing SVM. An ensemble of DCNN features and handcraft features also improves the performance slightly because its AUC is the highest at 0.8234.

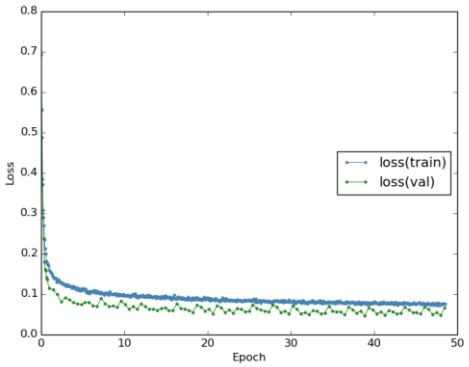

Figure 3. FCN training loss and validation loss with 50 epochs. The traits of loss reach quick convergence at 0.1 and then gets closer to 0 slowly. The validation loss curve going under the train loss curve means there is no appreciable overfitting during our training. The best performing model at 50 epoch was used on our own dataset to segment lung regions.

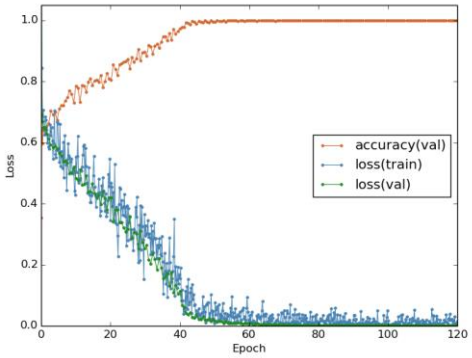

Figure 4. Training curve of DCNN classifier. The x-axis is the number of epochs, while the y-axis is the loss rate calculated by loss function of the neural network model. The smaller the loss, the better the ability to model the training set. In the initial stage, the loss is reduced from 0.8 to 0.6 and then slowly converges to 0.1 until steady.

4. CONCLUSION

This paper presents an automatic bacterial or viral pneumonia diagnosis method in chest radiographs based on deep learning models and handcraft features. Two public datasets, JSRT and MC, were used for segmentation. And the method was tested on a dataset of Guangzhou Women and Children's Medical Center. The results showed better accuracy (0.8048 ± 0.0202) and sensitivity (0.7755 ± 0.0296) in extracting features by DCNN with transfer learning. And an ensemble of all kinds of features helped slightly improve the performance of classification, increasing AUC value from 0.8160 ± 0.0162 to 0.8234 ± 0.0014.

Our method is a general approach that is also applicable to other medical classification task. Future work entails further tuning of DCNN with chest radiography and feature extraction. And more universal features will be further studied in the future.

5. ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61273043), the Natural Science Foundation of Guangdong and the Fundamental Research Funds for the Central Universities of China.

6. REFERENCES

[1] WHO., Global Pneumonia Report 2015.: World Health Organization, 2015.

[2] A. Mcluckie, "Respiratory disease and its management," Springer, vol. 57, no. 5, pp. 469–469, 2009.

[3] F. A. L. Bettenay, J. F. D. Campo, and D. B. Mccrossin, "Differentiating bacterial from viral pneumonias in children," Pediatric Radiology, vol. 18, no. 6, pp. 453–454, 1988.

[4] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," IEEE Transactions on Image Processing, vol. 19, no. 12, p. 3243, 2010.

[5] G. B. Van, A. F. Frangi, J. J. Staal, T. H. R. Bm, and M. A. Viergever, "Active shape model segmentation with optimal features," IEEE Transactions on Medical Imaging, vol. 21, no. 8, pp. 924–33, 2002.

[6] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. Mcdonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 577–90, 2014.

[7] S. Khobragade, A. Tiwari, C. Y. Patil, and V. Narke, "Automatic detection of major lung diseases using chest radiographs and classification by feed-forward artificial neural network," in IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, 2017, pp. 1–5.

[8] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, and S. Antani, "Automatic tuberculosis screening using chest radiographs," IEEE Transactions on Medical Imaging, vol. 33, no. 2, p. 233, 2014.

[9] G. Litjens, T. Kooi, B. E. Bejnordi, S. Aaa, F. Ciompi, M. Ghafoorian, V. D. L. Jawm, G. B. Van, and C. I. Snchez, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, no. 9, p. 60, 2017.

[10] S. E. Long, J. and T. Darrell, "Fully convolutional networks for semantic segmentation," in IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[11] J. Wang, J. D. Mackenzie, R. Ramachandran, and D. Z. Chen, "A deep learning approach for semantic segmentation in histology tissue images," in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016, pp. 176–184.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in International Conference on Neural Information Processing Systems, 2012, pp. 1097–1105.

[13] M. Havaei, A. Davy, D. Wardefarley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks." Medical Image Analysis, vol. 35, pp. 18–31, 2017.

[14] K. L. Hua, C. H. Hsu, S. C. Hidayati, W. H. Cheng, and Y. J. Chen, "Computer-aided classification of lung nodules on computed tomography images via deep learning technique," Oncotargets & Therapy, vol. 8, p. 2015, 2015.

[15] P. Lakhani and B. Sundaram, "Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks," Radiology, pp. 574–582, 2017.

[16] J. Shiraishi, J. Katsuragawa, Sikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists detection of pulmonary nodules," in AJR 174, 2000, pp. 71–74.

[17] S. Jaeger, S. Candemir, S. Antani, Y. X. Wng, P. X. Lu, and G. Thoma, "Two public chest x-ray datasets for computer-aided screening of pulmonary diseases." Quantitative Imaging in Medicine & Surgery, vol. 4, no. 6, p. 475, 2014.

[18] H. C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. Summers, "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp.1285–1298, 2016.

[19] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1717–1724.

[20] Y. Bar, "Deep learning with non-medical training used for chest pathology identification," in SPIE Medical Imaging, 2015, p. 94140V.

[21] H. Greenspan, B. V. Ginneken, and 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.

[22] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," pp. 675–678, 2014.

[23] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge 2012 (voc2012) results." [Online]. Available: http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html

[24] C. Gao and X. Hui, "Glcm-based texture feature extraction," Computer Systems & Applications, 2010.

[25] S. Chaplot, L. M. Patnaik, and N. R. Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network," Biomedical Signal Processing & Control, vol. 1, no. 1, pp. 86–92, 2006.

[26] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in CVPR, 2005, pp. 886–893.