The Multi-Task Deep Model-Based Pneumonia Detection

Xingfang Ai¹, Zhongliang Li², Min Cao³* and Xuechen Li²*

¹ Information Engineering College, Jiangmen Polytechnic, Jiangmen, Guangdong, 529000, China
² College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong, 518000, China
³ Radiology Department, Guangzhou Panyu Sanatorium, Guangzhou, Guangdong, 510000, China
*Corresponding author’s e-mail: timlee@szu.edu.cn

Abstract. As a common lung disease, pneumonia affects millions of people worldwide each year and is often detected during physical examination using chest X-ray images (CXRs), which are usually diagnosed by radiologists. This time-consuming task often leads to fatigue-based diagnostic error and cannot be performed in countries or areas lack of radiologists. In this work, we proposed a multi-task learning model: Mt-pnet. We learn from the experience of imaging doctors in the diagnosis of pneumonia, let the network pay attention to the patient's gender and age information in the diagnosis of pneumonia. The combined information is helpful for the network to learn robust image features for pneumonia diagnosis. At the same time, we also found that using multi-task model is helpful to improve the recall of patients with pneumonia, which has important clinical significance.

1. Introduction

As a common lung disease, pneumonia affects millions of people worldwide each year and is often detected during physical examination using chest X-ray images (CXRs), which are usually diagnosed by radiologists. This time-consuming task often leads to fatigue-based diagnostic error [1] and cannot be performed in countries or areas lack of radiologists. Thus, a computer system that can diagnose CXRs effectively could provide substantial benefit from improved workflow prioritization to clinical decision.

Machine learning based methods have achieved tremendous success in medical image analysis, such as CT [2], MRI [3] and CXR [4]. Previous work has proved that machine learning methods, especially deep learning methods[5] are promising to be applied in the diagnosis of glaucoma[6], cancer[7], acute brain hemorrhages [8].

In recent years, image classification based on the convolution neural network (CNN) have attracted increasing interests in different communities and demonstrated impressive performance. Take AlexNet [9] and VGG16 [10] as examples, these classic networks achieved excellent results in the ImageNet Large Scale Visual Recognition Competition [11], and were widely employed in medical image classification areas. Since the release ChestX-ray14 dataset [12] by National Institutes of Health in 2017, more studies have been proposed using deep learning for chest radiograph diagnosis [12, 13]. Wang et al. [12] presented a CNN based lung disease classification method, which can detect eight lung diseases, i.e. atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia and pneumothorax. They tested four classic models, i.e. AlexNet, VGG, Google and ResNet, and the experimental result showed...
that ResNet achieved the best performance. Yao et al. [13] employed DenseNet for 14 lung diseases classification. Rajpurkar et al. [14] presented a 121-layes DenseNet (CheXNet) for the same purpose and achieved the-state-of-the-art performances. However, the labels of ChestX-ray14 dataset were generated from the associated radiological reports using natural language processing, are thus very noisy. On the other hand, ChestX-ray14 dataset comprises 108,948 frontal-view CXRs with common thoracic diseases e.g. atelectasis, effusion and nodule, is thus not collected especially for pneumonia classification task.

In the actual clinical diagnosis, radiologists not only read the chest films of the patients, but also analyze their gender and age information to get the final diagnosis results. I. M. Baltruschat et al. [15] put the chest X-ray combined with the gender and age information of the case into the network learning. Through a large number of experiments, it is found that the introduction of non-image information can improve the generalization performance of the model.

In this work, we proposed a pneumonia diagnosis model based on multi-task learning. The gender and age information of patients and whether chest X-ray contains pneumonia are taken as the prediction task of the network, which can make the network to learn the characteristics of pneumonia, and extract the characteristics related to the age and gender of the patients at the same time. These can improve the learning difficulty of the network and further improves the robustness of the network.

2. Material and Methods

2.1. Datasets

In this work, 10784 CXRs were collected at the Shenzhen No.2 People’s Hospital, Guangdong, Shenzhen, China. The dataset, named as “MT pneuxray”, were used for training and evaluating the proposed Mt-pnet. All the chest X-ray images are labeled by expert radiologists as “Pneumonia” or “Normal”. The resolution of the original image is about 3000 × 3000. A total of 2882 DICOM files were collected, including the patient’s chest X-ray images, age, gender and whether he has pneumonia information.

We summarize the data distribution of pneumonia and gender in MT pneuxray dataset in Table 1, and the distribution of age in Fig. 1. We can see that the categories of pneumonia and gender tags in MT pneuxray dataset are relatively balanced. For age distribution, there are more data for young and middle-aged people, and less data for the elderly and children. Combined with medical knowledge, we divide the age data into four categories as tags by taking the age of 14, 25 and 65 as the separation point.

Table 1. The data distribution of MT pneuxray Dataset

|              | Number | Total |
|--------------|--------|-------|
| Pneumonia    | 1683   | 2882  |
| Normal       | 1199   |       |
| Male         | 1492   | 2882  |
| Female       | 1390   |       |
2.2. Mt-pnet network architecture
The single task learning model only considers the pneumonia related features learned from the input chest film, and let the network learn the simple mapping relationship between the image and the disease label. The learned model thus lacks robustness in the pneumonia diagnosis task. In this work, we proposed a multi-task learning network Mt-pnet. Considering that radiologists will diagnose pneumonia with non-image information such as patient's gender and age, as shown in Fig. 2, Mt-pnet network extracts the chest image, age, gender and pneumonia information from DICOM file. The main task of network learning is pneumonia diagnosis, and the auxiliary task is patient age and gender classification. Such a multi-task structure can not only let the network learn the image features related to pneumonia, but also learn the patient's age and gender features, so that the network can extract more robust features for pneumonia diagnosis.

2.3. Loss function
Mt-pnet network is a multi-task model, which has three output branches: the main task is pneumonia diagnosis, and the auxiliary task is patient gender and age recognition. Considering that if the age tag is
directly used as the learning goal, it will be a classification problem of hundreds of categories, and the learning difficulty will be greatly increased. Therefore, combined with medical knowledge, we take the 14, 25 and 65 year-old as the separation point, and divide the age data into four categories as tags for learning.

Mt-pnet includes three learning objectives, i.e. pneumonia recognition target $\ell_s$, gender identification target $\ell_g$ and age recognition target $\ell_a$. We set different learning weights for each learning goal, which are $\lambda_s$, $\lambda_g$ and $\lambda_a$. Pneumonia and gender recognition are two classification tasks, which are optimized by binary cross entropy loss function. Age classification is a four classification task, which is optimized by standard cross entropy loss function. The overall learning goal is shown in equation 1.

$$ L = \lambda_s \ell_s + \lambda_g \ell_g + \lambda_a \ell_a $$

(1)

2.4. Network training

The system environment of this experiment is Ubuntu 16.04.4. Keras framework is employed to realize depth model. The network is trained on a Tesla P100 GPU. The learning rate of the network is set to 0.1; the activation function is relu; the batch size is 32; and the training epoch is 200. The objective function of network learning is binary cross entropy and standard cross entropy loss function, and Adam optimizer is used for network optimization. In the network training, we read the training data from the pickle file, load the training image according to the image file path of the first column of the pickle file, and train the network according to the gender, age and pneumonia label of the corresponding column to learn the robust deep pneumonia diagnosis features.

3. Experimental results

3.1. Evaluation metrics

We use Accuracy and Recall and AUC to quantitatively evaluate the classification performance of the model under different parameter settings.

$$ \text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} $$

(2)

$$ \text{Recall} = \frac{TP}{TP + FN} $$

(3)

3.2. Comparison with single task network

In Fig. 3 we compared the performance of the two learning strategies on accuracy, recall and AUC under the setting of training data and test data partition ratio of 8:2, 5:5, 3:7 and 1:9. We found that in the case of training-test data ratio of 8:2 and 5:5, the multi-task model has no obvious advantages in pneumonia diagnosis, and the accuracy rate of pneumonia diagnosis is slightly lower than that of single task model. However, with the decrease of training data, the advantages of multi-task learning methods are gradually reflected. The performance of the multi task model is better than that of the single task model when the training-test data ratio is 3:7 and 1:9.

At the same time, in terms of the recall of pneumonia patients, the multi-task model has a significant improvement compared with the single task model. As can be seen from figure 3, the multi-task model pays attention to the gender and age information related to pneumonia diagnosis during learning, which improves the sensitivity of the network to patients with pneumonia. In the case of using only 10% data for training, the recall rate of pneumonia in multi task model reached 81.12%, which was much higher than that in single task model (67.13%).

We found using multi-task learning strategy did not significantly improve the diagnosis performance of pneumonia in the case of sufficient training data. On the contrary, because for the multi-task learning strategy, the network needs to predict the gender and age of patients while the network diagnosis of pneumonia, which increases the difficulty of network training, and leads to the lower performance of the network in the diagnosis of pneumonia in the main task. However, with the decrease of training data, the performance of Mt-pnet is higher than that of single task model. We can see that Mt-pnet model has
achieved better performance in pneumonia diagnosis task when the proportion of training data is reduced to 10% of the whole dataset. We can see that the recall of single task model is only 67.13%, while the recall rate of multi-task model reaches 81.12%, and AUC is 0.7514, which is better than 0.7481 of single task model. It further shows that multi-task learning strategy can help the network extract more robust image features, thus improving the robustness of computer-aided pneumonia diagnosis system.

Figure 3. The performance comparison of single and multi-task model under different data partition

4. Conclusion
In this work, we proposed a multi-task learning model: Mt-pnet. We learn from the experience of imaging doctors in the diagnosis of pneumonia, let the network pay attention to the patient's gender and age information in the diagnosis of pneumonia. The combined information is helpful for the network to learn robust image features for pneumonia diagnosis. At the same time, we also found that using multi-task model is helpful to improve the recall of patients with pneumonia, which has important clinical significance. Early screening of pneumonia patients for early treatment reflects the practical value of computer-aided pneumonia diagnosis system.

Acknowledgments
The work is supported by National Natural Science Foundation of China (Grant No. 61702337) and Science and Technology Project of Guangzhou, Panyu, (Grant No. 2019-Z04-48).

References
[1] J. H. Austin, B. M. Romney, and L. S. Goldsmith, "Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect," Radiology, vol. 182, p. 115, 1992.
[2] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features," IEEE Trans Med Imaging, vol. 27, pp. 1668-1681, 2008.
[3] G. Mohan and M. M. Subashini, "MRI based medical image analysis: Survey on brain tumor grade classification," Biomedical Signal Processing and Control, vol. 39, pp. 139-161, 2018.

[4] U. Avni, H. Greenspan, M. Sharon, E. Konen, and J. Goldberger, "X-ray image categorization and retrieval using patch-based visualwords representation," in Biomedical Imaging: From Nano to Macro, 2009. ISBI ’09. IEEE International Symposium on, 2009.

[5] D. Shen, G. Wu, and H. I. Suk, "Deep Learning in Medical Image Analysis," Annual Review of Biomedical Engineering, vol. 19, pp. 221-248, 2017.

[6] "Learning ECOC Code Matrix for Multiclass Classification with Application to Glaucoma Diagnosis," Journal of Medical Systems, vol. 40, p. 78, 2016.

[7] S. Mitra and B. U. Shankar, "Medical image analysis for cancer management in natural computing framework," Information Sciences, vol. 306, pp. 111-131, 2015.

[8] S. P. Singh, J. Ker, Y. Bai, J. Rao, and L. Wang, "Image Thresholding Improves 3-Dimensional Convolutional Neural Network Diagnosis of Different Acute Brain Hemorrhages on Computed Tomography Scans," Sensors, vol. 19, pp. 2167-, 2019.

[9] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in NIPS, 2012.

[10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Computer Science, 2014.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, pp. 211-252, 2015.

[12] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[13] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, and K. Lyman, "Learning to diagnose from scratch by exploiting dependencies among labels," arXiv preprint arXiv:1710.10501, 2017.

[14] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, et al., "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," arXiv preprint arXiv:1711.05225, 2017.

[15] I. M. Baltruschat, H. Nickisch, M. Grass, T. Knopp, and A. Saalbach, "Comparison of deep learning approaches for multi-label chest X-ray classification," Scientific reports, vol. 9, p. 6381, 2019.