## Letter to the Editor

# Deep Learning in Chest Radiography: Detection of Pneumoconiosis*

LI Xiao[1], LIU Chao Fei[2], GUAN Li[1], WEI Shu[2], YANG Xin[2], and LI Shu Qiang[1,#]

Pneumoconiosis, the predominant occupational disease in China and the world, is a pulmonary disease caused by the inhalation of inorganic dust, that is, particulate matter in the solid phase without living organisms[1]. As an irreversible, crippling, and even fatal disease, pneumoconiosis places a heavy burden on society. Industrial workers who are exposed to exhalable inorganic dust like asbestos, silica, and coal dust have a greater risk of developing pneumoconiosis. Early detection, diagnosis, and treatment are the key to a better prognosis. According to International Labour Organization (ILO) Guidelines, chest radiography is the most accessible and affordable radiological test available for the physical examination of workers exposed to dust and mass screening for pneumoconiosis[2]. However, the test is limited by the tremendous volume of images produced and the resultant burden on radiologists, resulting in low efficiency and poor stability. Additionally, radiographic interpretation is subjective and reliant on the personal experience of radiologists. Junior physicians may interpret the radiographs inaccurately, resulting in missed or delayed diagnoses. Therefore, a computer-aided diagnosis (CAD) scheme developed for accurate and fast detection of pneumoconiosis can effectively reduce the workload of radiologists and improve the efficiency of mass screening with chest radiography. The system can also provide a diagnostic standard or reference for junior radiologists, which serves as a useful tool for radiological training.

Deep convolutional neural networks are representational learning networks that can automatically extract features associated with a task. DL-based artificial intelligence systems can be applied to expedite and improve the interpretation of several imaging modalities, including radiography, computed tomography, and magnetic resonance, and will revolutionize radiological diagnostic systems. The application of DL algorithms to chest radiographs is currently being evaluated for various thoracic diseases, such as lung nodules, pulmonary tuberculosis, pneumonia, pneumothorax, and many other pectoral abnormities such as pleural effusions, pulmonary opacities, enlarged cardiac silhouette, and hilar prominence[3-5].

Large-scale studies regarding the diagnostic performance of pneumoconiosis by DL algorithms remain scarce. In this study, we created a chest radiograph database containing more than 230,762 digital radiography images from 21 hospitals nationwide in 16 provinces of China. A total of 34,997 chest radiographs from patients with pneumoconiosis or workers exposed to dust were selected from the database. The following 1,504 images were excluded: 1) lower-quality (fourth-class) chest radiographs, 2) not posterior–anterior chest radiographs, and 3) chest radiographs that do not meet technical requirements of the ILO Guidelines. The exclusion ensured that the technical requirements for pneumoconiosis are fulfilled, specifically the criterion that chest radiographs be obtained *via* posterior–anterior projection and a digital radiography technique. The remaining 33,493 films were all labeled. Finally, selected chest radiographs were randomly assigned to one of the following two datasets: a training dataset that consisted of 27,493 chest radiographs (9,011 positive and 18,482 negative images) to train the network and a test dataset (6,000 chest radiographs consisting of 1,966 positive and 4,034 negative chest radiographs) to validate the detection performance of the training network.

The test dataset was annotated by experienced radiologists in the same manner as the training set, and the positive rates (pneumoconiosis) for the two

*doi: 10.3967/bes2021.116*

*This study was supported by the National Natural Science Foundation of China: Molecular mechanism of 1-phosphosphingol and lysophosphatidic acid involved in silicosis [No: 81641119].

1. Peking University Third Hospital, Beijing 100000, China; 2. Beijing Tianming Innovation Data Technology Co., LTD, Beijing 100000, China

datasets were nearly the same, both exceeding 30%. Cases in the two datasets were different and exclusive to each dataset. Only images labeled as pneumoconiosis were defined as positive. Images labeled as other pulmonary diseases were excluded from the training dataset.

We use self-developed software for the annotations (http://anno.maxdiag.com/) and have formulated operation specifications to standardize the annotation process. In this study, chest radiographs evaluated were annotated by 10 occupational–disease–diagnosis experts and two professional radiologists blinded to each other's labels and annotations. Each radiograph has 3–5 annotated records. These records contain detailed information regarding the quality of the chest radiograph, characteristics of the pneumoconiosis opacities (i.e., shape, size, location, and distribution), and most importantly, diagnosis (i.e., normal, pneumoconiosis, or other pulmonary diseases). Following the ILO Guidelines[2], each lung field is divided into three zones, namely, upper, middle, and lower zones. The radiologists diagnose pneumoconiosis based on the profusion, shape, and size of the pneumoconiosis opacities. The profusion level of small opacities represents the concentration of small opacities in the affected zones of the lung, which can reflect the degree of pneumoconiosis. Profusion is determined by comparison with standard radiographs and recorded as categories 0, 1, 2, and 3 and subcategories 0/−, 0/0, 0/1, 1/0, 1/1, 1/2, 2/1, 2/2, 2/3, 3/2, 3/3, and 3/+. As a result, approximately 65.8% (22,026/33,493) of chest radiographs had no pneumoconiosis or other pulmonary diseases (labeled as normal). Pneumoconiosis was detected in 32.8% (10,977/33,493) of the images. Among them, pneumoconiosis of categories 1, 2, and 3 comprised 57.1% (6,268/10,977), 16.2% (1,776/10,977), and 26.7% (2,933/10,977) of the images, respectively. Additionally, 1.5% (490/33,493) of the radiographs were labeled as other pulmonary diseases.

To improve the accuracy of the models and reduce information loss from the images, four kinds of images with different picture formats or resolution ratios, namely, JPEG-256, JPEG-512, DICOM-256, and DICOM-512, were used to train the models. The accuracy values for each picture format were 0.87, 0.89, 0.88, and 0.90, respectively. Therefore, instead of using JPEG files, 512 × 512 high-resolution images from the DICOM dataset were used as inputs for the model to minimize information loss.

Moreover, a self-developed software (http://anno.maxdiag.com/) and SPSS statistical software (IBM SPSS Statistics, Armonk, NY, USA) were used for statistical analysis. The true-positive (the number of pneumoconiosis cases identified as pneumoconiosis), false-positive (FP, the number of normal cases judged as pneumoconiosis), true-negative (the number of normal cases judged as normal), false-negative (the number of pneumoconiosis cases judged as normal; this was the number of missed cases), and sensitivity (the proportion of cases correctly diagnosed as pneumoconiosis) rates were calculated to determine the diagnostic power.

The system is designed for pneumoconiosis screening; thus, high sensitivity is necessary to reduce missed diagnoses. To improve sensitivity, a resampling technique was employed in model training to improve the suitability of the model for the physical examination of workers exposed to dust and pneumoconiosis screening. In the final evaluation results for the test dataset, the sensitivity of the DL algorithm was 0.99, the specificity was 0.88, and the accuracy was 0.92. The area under the curve (AUC) of the DL algorithm was 0.96. This performance profile is appropriate for screening low-prevalence findings. The extremely high sensitivity makes the model an optimal tool for pneumoconiosis screening in an occupational health examination, as it can efficiently identify nearly all potential cases of pneumoconiosis. The identified cases can then be submitted to a human radiologist for further image review and diagnosis. With this system, the workload of radiologists will significantly decrease without compromising the accuracy of diagnosis. Additionally, because of its high AUC, the model can aid in the re-evaluation of radiographic interpretations by human radiologists and help avoid missed or false diagnoses caused by subjectivity or fatigue. The models can also help junior radiologists or physicians improve their skills in radiographic interpretation.

Various automated and semi-automated CAD systems have been developed for the detection of pneumoconiosis over time, with the methods used changing from traditional image-analysis techniques to machine-learning and DL approaches. Several decades ago, CAD systems were explored for the detection of pneumoconiosis in chest radiographs using texture analysis[6]. Yu et al. adopted a modified approach in their study, with accuracy and AUC of 87.7%–89.2% and 0.948–0.978, respectively[7]. Support vector machine (SVM) was applied as a

region-level classifier, incorporating the prediction results for six regions into the final classification. The lung field was first subdivided into six regions and subjected to further segmentation based on a combination of Otsu's threshold method and morphological reconstruction. A computerized scheme with 22 wavelet-based energy texture features and SVM was developed for the detection of abnormal regions among normal sections, resulting in high-opacity recognition accuracy of up to 83.3%[8]. All of the above-mentioned methods are non-neural networks, which require manual feature extraction and additional time. Consequently, they never became mainstream methods in the machine-learning field. Okumura et al.[9] combined rule-based plus artificial neural network (ANN) methods with three new enhanced techniques to distinguish abnormal from normal regions. The diagnostic power achieved with this approach was high for severe pneumoconiosis, with AUC values of 0.93 ± 0.02, but moderate for early pneumoconiosis, with AUC values of 0.72 ± 0.03. In 2017, the authors improved the technique by developing a CAD system based on a three-stage ANN method, which achieved a satisfactory diagnostic power for both cases of severe pneumoconiosis and early pneumoconiosis, with AUCs of 0.89 ± 0.09 and 0.84 ± 0.12, respectively[10]. However, ANN is a shallow neural network that is incapable of efficient representational learning to obtain satisfactory accuracy, which compromises its potential for broader applications.

Compared with previous reports, our study has the following advantages. First, we used deep CNNs to train the model, which employs representational learning and can automatically extract features associated with a task. Additionally, DL is labor-saving and can significantly improve the accuracy of detection compared with traditional methods. Second, the development of a reliable classifier for clinical use mainly relies on the sample size and dataset quality, which should accurately represent the actual clinical scenario. Our dataset contained 33,493 cases (of which 10,977 were positive for pneumoconiosis) from 21 hospitals nationwide in 16 provinces of China. To the best of our knowledge, this is the largest dataset reported in the field of pneumoconiosis CAD detection. In this study, the radiographs included were also obtained from various clinical scenarios and patients, containing normal and abnormal cases ranging from early pneumoconiosis to severe cases. Therefore, the dataset used to train the DL algorithm was diverse and comprehensive, which is highly conducive for the effective adoption of the model to various clinical scenarios. Third, we invited 10 occupational–disease–diagnosis experts and two professional radiologists nationwide who are experts in pneumoconiosis to annotate the radiographs evaluated in the study. Therefore, we could confirm the accuracy and authority of the annotation results. Additionally, we applied the convergent consistency weighted voting algorithm and excluded fourth-class images to ensure the accuracy of the results during the annotation process. Finally, given the above-mentioned efforts, our work surpasses previous studies by demonstrating a very high level of performance and a satisfactory low incidence of FPs when applied to a low-prevalence disease such as pneumoconiosis, with results superior to nearly all models reported for pneumoconiosis detection.

Despite the striking results, this study has some limitations that need to be addressed. Since the model sensitivity is not 100%, some cases of pneumoconiosis may not be detected by the algorithm. Additionally, pneumoconiosis diagnoses by radiologists based on radiographs are not always correct. Long-term follow-up or further testing is necessary to confirm the diagnosis.

Overall, our CAD scheme achieved a promising diagnostic performance, with an accuracy of 92.0%, a sensitivity of 99%, and an AUC value of up to 0.96, which has been installed in more than 10 hospitals in China. The CAD scheme is a potential tool for the early detection of pneumoconiosis, as performance will be further improved through model refinement and steadily larger sample sizes. Furthermore, based on our database, we will continue developing DL algorithm systems for pneumoconiosis category diagnosis, which will be useful tools in clinical practice.

[#]Correspondence should be addressed to LI Shu Qiang, Associate Senior Doctor, E-mail: shuqiangli@263.net, Tel: 86-10-82266276.

Biographical note of the first author: LI Xiao, female, born in 1989, MD, Doctor-in-charge, majoring in diagnosis and treatment of occupational disease.

Received: October 28, 2020;

Accepted: August 24, 2021

## REFERENCES

1. Mandrioli D, Schlünssen V, Ádámd B, et al. WHO/ILO work-related burden of disease and injury: protocol for systematic reviews of occupational exposure to dusts and/or fibres and of the effect of occupational exposure to dusts and/or fibres on pneumoconiosis. Environ Int, 2018; 119, 174−85.

2. International Labour Organization (ILO). Guidelines for the use of the ILO international classification of radiographs of pneumoconioses. ILO. 2011, 3-6.
3. Ather S, Kadir T, Gleeson F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. Clin Radiol, 2020; 75, 13−9.
4. Rajpurkar P, Irvin J, Ball RL, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med, 2018; 15, e1002686.
5. Singh R, Kalra MK, Nitiwarangkul C, et al. Deep learning in chest radiography: detection of findings and presence of change. PLoS One, 2018; 13, e0204155.
6. Turner AF, Kruger RP, Thompson WB. Automated computer screening of chest radiographs for pneumoconiosis. Invest Radiol, 1976; 11, 258.
7. Yu PC, Xu H, Zhu Y, et al. An automatic computer-aided detection scheme for pneumoconiosis on digital chest radiographs. J Digit Imaging, 2011; 24, 382−93.
8. Zhu L, Zheng R, Jin H, et al. Automatic detection and recognition of silicosis in chest radiograph. BioMed Mater Eng, 2014; 24, 3389−95.
9. Okumura E, Kawashita I, Ishida T. Development of CAD based on ANN analysis of power spectra for pneumoconiosis in chest radiographs: effect of three new enhancement methods. Radiol Phys Technol, 2014; 7, 217−27.
10. Okumura E, Kawashita I, Ishida T. Computerized classification of pneumoconiosis on digital chest radiography artificial neural network with three stages. J Digit Imaging, 2017; 30, 413−26.