Artificial Intelligence-Enhanced Pneumoconiosis Screening and Staging from Chest Radiographs

PI: Prof. Chen Hao
Assistant Professor, Department of Computer Science, HKUST

Objectives:
The research project aims at developing an AI-based system for automated screening and staging of pneumoconiosis.

Background:
Pneumoconiosis is a lung disease caused by inhalation of dust particles. Artificial intelligence–assisted diagnosis shows promise in accelerating pneumoconiosis screening and staging, reducing misdiagnosis rates, alleviating the burden on radiologists, and supporting the prevention and treatment of occupational diseases. AI-powered automated screening and staging systems can serve as effective complementary diagnostic tools in Hong Kong, the Chinese Mainland, and other regions.

Methodology:
The AI system first screened and classified chest X-rays collected from an open-access pneumoconiosis database. For suspected positive cases, the system further assessed disease severity and identified other possible complications. In parallel, the system performed anatomical segmentation and modelling of the chest images and analysed pneumoconiosis-related radiographic findings. In addition, the system quantitatively evaluated disease severity in each lung region.

Impact:
The system can be widely used in hospitals for complementary pneumoconiosis diagnosis and quantitative classification in the future, thereby accelerating the efficiency of diagnosis and assessment, and effectively promoting the prevention and treatment of pneumoconiosis.

Results and Conclusion:
1. Approximately one million chest X-ray images from 5 public datasets (MIMIC-CXR, VinDr-CXR, NIH CXR, PadChest and CheXpert) were collected for large-scale pretraining.
2. A total of 2 336 specific pneumoconiosis and 2 162 silicosis data were obtained from the Hospital Authority Data Collaboration Laboratory (HADCL), ensuring that the model was tested on a local demographic trend.
3. The pneumoconiosis screening model achieved 97.03% accuracy and 97.43% area under the curve (AUC), exceeding the proposal's target of 95% accuracy.
4. The lung segmentation model achieved a Dice score of 95.64%, meeting the 95% target, and was able to divide the lungs into upper, middle, and lower regions for anatomical analysis.
5. The lesion localization and segmentation model achieved a Dice value of 90.19%, exceeding the target of 85%, and was able to identify lesion regions.
6. The pneumoconiosis staging model integrated outputs from the lung segmentation and lesion localization models and calculated lesions ratio within specific lung regions. It achieved 91.13% classification accuracy, exceeding the target of 90%.
7. The multi-disease classification model achieved 92.73% average accuracy (pleural thickening 97.77%, pneumonia 96.42%, pneumothorax 96.02%, consolidation 95.57%,

cardiomegaly 90.51%, edema 89.21%, atelectasis 88.78% and pleural effusion 87.56%), exceeding the target of 90%.

8. Continued accumulation of clinical information would further improve system accuracy and performance.
9. The integrated workflow was expected to allow retrieving clinical data to the PACS interface, prioritizing critical cases for immediate review, validation of any positive findings by physicians, digital report signing, and transferring results to the electronic health system (eHealth).