ARTIST
JEFF ZHANG

CRITICAL REVIEW

Artificial intelligence in diagnosing lung cancer: Applications and future developments

doi:10.35493/medu.39.18

JUSTIN PHUNG
Bachelor of Health Sciences (Honours), Class of 2024, McMaster University
Correspondence: phungj3@mcmaster.ca

ABSTRACT

Computer-aided diagnosis (CAD) systems lie at the intersection of medicine and computer science. Over the last couple of decades, consistent research and technological advances have resulted in a steady improvement of CAD systems that are capable of assisting in the detection and diagnosis of various diseases. However, several limitations prevent CAD from being implemented in clinical practice. The primary purpose of this review is to provide a general overview of CAD systems in the context of lung cancer, as well as assess the critical challenges that CAD must overcome. Such challenges include data privacy and sharing laws, radiologist workflow integration, and the lack of a standardized performance evaluation. Thus, coordination between radiologists, researchers, and medical institutions will play a pivotal role in shaping the future development of CAD systems in healthcare.

CONTEXT

The applications of artificial intelligence (AI) in the field of medicine have made much headway in recent years, especially with regards to detection and diagnosis of diseases using computer-aided diagnosis (CAD) systems. AI utilizes computers to simulate human intelligent processes, including learning, reasoning, and thinking.[1] In particular, convolutional neural networks (CNNs), a subset of AI, can automatically extract image features after training on labelled samples.[1,2] Given the availability of large datasets and increased computing power, CNN-based CAD systems have produced promising results for many tasks, including image classification, correct image detection, and segmentation, proving themselves capable of replacing current CAD systems based on manual input.[3]

Although there are variations in how CAD systems function, most undergo five general steps: acquisition, preprocessing (increasing the precision and accuracy of algorithms), segmentation (separating the study region from other organs and tissues in radiographic images), nodule detection (marking the location of pulmonary nodules in the image), and elimination of false positives.[4] CAD systems are further divided into computer-aided detection (CADe) and computer-aided diagnosis (CADx).[3] This distinction is critical as their functions differ: CADe systems detect potential lesions and reveal abnormalities in medical images, whereas CADx systems primarily serve to characterize, classify, and distinguish lesions.[5]

Although the purpose of CAD is to assist radiologists in the detection of lung nodules, CAD has not been implemented in routine clinical practice, with research efforts still focused primarily on improving performance.[4] Currently, computerized tomography (CT) scans are the most common imaging modality for radiomics analyses, especially for diagnosing lung cancer, due to their high spatial resolution, cost-effectiveness, wide availability, and noninvasiveness.[5,6] CNNs have been successfully developed to detect pulmonary nodules and lesion segmentation through training on publicly available databases and testing on various datasets such as the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).[7,8]

EVALUATING EFFECTIVENESS OF CAD

CAD can potentially reduce radiologists' workloads and enhance their performances. Reading time is significantly shorter with CAD used as a concurrent reader compared to when the reading is done after interpretation by a radiologist.[4] Moreover, CAD has been shown to improve the performance of experienced and inexperienced radiologists.[9] Kligerman et al. determined that CAD can improve a radiologist's ability to accurately detect lung nodules that were initially missed, while Sahiner et al. further demonstrated that it can help detect small size nodules under five millimetres, which are easily overlooked by inspection alone.[10,11] As such, CAD consistently improves performance among thoracic radiologists and can play a critical role in the detection of lung nodules.

Despite the clear impact of CAD on lung nodule detection, its performance varies significantly depending on the study. Generally, a research study's CAD system can be evaluated through several metrics including accuracy, precision, sensitivity, specificity, true positive rate, and false positive rate.[2] Although the establishment of databases like the LIDC-IDRI has provided a large repository of training images that have steadily improved sensitivity and specificity, CAD systems remain inconsistent depending on the method and training dataset.[12] For example, one model trained on the LIDC-IDRI dataset was a multi-view CNN, a 3D model that encodes for richer spatial information. Its error rate was 5.41%, and the sensitivity and specificity rates were 90.49% and 99.91%, respectively.[12] Conversely, another study conducted by Nishio et al. utilized the deep CNN method on a clinical dataset; their highest reported accuracy score was 68.0%.[13] Overall, the CAD effectiveness varies depending on the method and training dataset used.

Much of the existing literature is focused on CADe systems, involving the detection of nodules or lesions. Zhang et al. noted in their appraisal that there have been many reviews about CADe systems, but few regarding CADx systems.[5] CADe and CADx are often regarded separately, with most research focused on individually optimizing either system.[14]

However, a study by Ozdemir et al. suggests that coupling CADe and CADx systems can result in improved performance.[14] Using an open-source dataset, CADe and CADx systems were developed simultaneously, allowing for an optimized system which can reduce the false positive rate and serve as an end-to-end automated diagnostic tool for lung cancer.[14]

ETHICAL AND PRACTICAL LIMITATIONS

Several limitations currently plague the field of CAD and hinder its development. Data scarcity is one such challenge, as properly and reliably training a CAD system is resource-intensive. Labeling tools must be made available to radiologists to create high-quality datasets, and although databases are expanding, a higher volume of data is required to improve functionality.[15] There are several work-arounds, such as transfer learning —a machine learning method that involves pre-training on a large dataset, followed by fine-tuning on another dataset— that has been demonstrated to show improved accuracy despite the lack of labeled images.[13] Supervised learning could also be utilized on a small portion of the dataset to train networks, followed by unsupervised learning that classifies the remaining unlabeled data.[15] However, the primary concern pertains to legal and ethical issues surrounding data privacy.

Numerous data confidentiality laws govern the use of patient images in academic settings, and medical institutions may face fines if personal health information is compromised.[3] Despite the need for extensive high-quality images to train CNNs, healthcare organizations may be deterred from contributing to a shared learning dataset as the possibility of mishandling medical images and facing litigation is not worth the risk. Nevertheless, data security in healthcare continues to improve, and new privacy-preserving models such as de-identification and anonymization are crucial to safeguarding and managing patient records.[16] Once the security of privacy-enhancing technologies is strengthened, the risks associated with uploading and sharing patient data can be drastically reduced, and de-identified patient data can be used for research with the approval of a research ethics board. Thus, a vast number of medical facilities may be incentivized to collaborate and contribute to a shared training dataset for CAD systems.

Another challenge that must be addressed is the diagnostic accuracy of CAD. The performance of CAD systems is steadily improving, but still varies with the study or method. The lack of a standardized performance assessment for CAD systems adds to the problem. Although at least one measure of accuracy is recommended by the Standards for Reporting of Diagnostic Accuracy Studies, the specific measures are not clearly stated and measures reported in published articles remain inconsistent depending on the application of the CAD system (e.g. segmentation or classification).[17] By establishing a well-defined standard such that the most important metrics are consistently reported in the literature, comparing different CAD systems becomes easier and more reliable, which would provide a direction for future researchers to build on past studies. In particular, Gao et al. highlighted eight common evaluation metrics that can be used to form a standard, including accuracy, precision, sensitivity, specificity, and true and false positive rates.[2] CAD systems that perform well against other systems can then be replicated; for example, the optimal deep CNN method that yields high results across several metrics can be further investigated by researchers on alternate datasets.[18] Thus, developing a standardized evaluation of various CAD systems is necessary to unify research efforts.

Finally, coordination between radiologists and researchers is needed to integrate CAD into clinical practice. Despite its potential to improve radiologist performance, CAD systems remain absent in clinical settings as they fail to integrate with radiologists' workflow.[19] As a result, CAD systems are likely to be regarded as low priority compared to their clinical tasks.[2,19] Nevertheless, combining the knowledge of radiologists and computer analysis can enhance CAD system performance.[5] Radiologists have a strong understanding of the diagnostic process and can identify the strengths and areas of improvement

1. Cong L, Feng W, Yao Z, Zhou X, Xiao W. Deep learning model as a new trend in computer-aided diagnosis of tumor pathology for lung cancer. *J Cancer*. 2020;11(12):3615-22. Available from: doi:10.7150/jca.43268.
2. Gao J, Jiang Q, Zhou B, Chen D. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. *Math Biosci Eng*. 2019;16(6):6536-61. Available from: doi:10.3934/mbe.2019326.
3. Lee SM, Seo JB, Yun J, Cho YH, Vogel-Claussen J, Mark L, et al. Deep learning applications in chest radiography and computed tomography. *J Thorac Imaging*. 2019;34(2):75-85. Available from: doi:10.1097/RTI.0000000000000387.
4. Al Mohammad B, Brennan PC, Mello-Thoms C. A review of lung cancer screening and the role of computer-aided detection. *Clin Radiol*. 2017;72(7):433-42. Available from: doi:10.1016/j.crad.2017.01.002.
5. Zhang G, Yang Z, Gong L, Jiang S, Wang L, Cao X, et al. An appraisal of nodule diagnosis for lung cancer in CT images. *J Med Syst*. 2019;43(7):181. Available from: doi:10.1007/s10916-019-1327-0.
6. Chen B, Zhang R, Gan Y, Yang L, Li W. Development and clinical application of radiomics in lung cancer. *Radiat Oncol*. 2017;12(1):154. Available from: doi:10.1186/s13014-017-0885-x.
7. Benzaquen J, Boutros J, Marquette C, Delingette H, Hofman P. Lung cancer screening, towards a multidimensional approach: Why and how? *Cancers (Basel)*. 2019;11(2):212. Available from: doi:10.3390/cancers11020212.
8. Armato SG 3rd, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med Phys*. 2011;38(2):915-31. Available from: doi:10.1118/1.3528204.
9. Marten K, Seyfarth T, Auer F, Wiener E, Grillhösl A, Obenauer S, et al. Computer-assisted detection of pulmonary nodules: Performance evaluation of an expert knowledge-based detection system in consensus reading with experienced and inexperienced chest radiologists. *Eur Radiol*. 2004;14(10):1930-38. Available from: doi:10.1007/s00330-004-2389-y.
10. Kligerman S, Cai L, White CS. The effect of computer-aided detection on radiologist performance in the detection of lung cancers previously missed on a chest radiograph. *J Thorac Imaging*. 2013;28(4):244-52. Available from: doi:10.1097/RTI.0b013e31826c29ec.
11. Sahiner B, Chan HP, Hadjiiski LM, Cascade PN, Kazerooni EA, Chughtai AR, et al. Effect of CAD on radiologists' detection of lung nodules on thoracic CT scans: Analysis of an observer performance study by nodule size. *Acad Radiol*. 2009;16(12):1518-30. Available from: doi:10.1016/j.acra.2009.08.006.
12. Kang G, Liu K, Hou B, Zhang N. 3D multi-view convolutional neural networks for lung nodule classification. *PLoS One*. 2017;12(11):e0188290. Available from: doi:10.1371/journal.pone.0188290.
13. Nishio M, Sugiyama O, Yakami M, Ueno S, Kubo T, Kuroda T, et al. Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. *PLoS One*. 2018;13(7):e0200721. Available from: doi:10.1371/journal.pone.0200721.
14. Ozdemir O, Russell RL, Berlin AA. A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. *IEEE Trans Med Imaging*. 2020;39(5):1419-29. Available from: doi:10.1109/TMI.2019.2947595.
15. Liu B, Chi W, Li X, Li P, Liang W, Liu H, et al. Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: Three decades' development course and future prospect. *J Cancer Res Clin Oncol*. 2020;146(1):153-85. Available from: doi:10.1007/s00432-019-03098-5.
16. Abouelmehdi K, Beni-Hssane A, Khaloufi H, Saadi M. Big data security and privacy in healthcare: A review. *Procedia Comput Sci*. 2017;113(1):73-80. Available from: doi:10.1016/j.procs.2017.08.292.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351(1):h5527. Available from: doi:10.1136/bmj.h5527.
18. Savitha G, Jidesh P. Advances in Intelligent Systems and Computing. Singapore: *Springer*; 2019. 11-23 p.
19. Ather S, Kadir T, Gleeson F. Artificial intelligence and radiomics in pulmonary nodule management: Current status and future applications. *Clin Radiol*. 2020;75(1):13-9. Available from: doi:10.1016/j.crad.2019.04.017.
20. Yanase J, Triantaphyllou E. The seven key challenges for the future of computer-aided diagnosis in medicine. *Int J Med Inform*. 2019;129(1):413-22. Available from: doi:10.1016/j.ijmedinf.2019.06.017.

of current CAD systems. The implementation of a feedback system would offer insight and guidance to academic researchers developing practical CAD systems.[20]

CONCLUSION

Artificial intelligence has the potential to improve patient outcomes by detecting, classifying, and diagnosing pulmonary nodules. Although substantial progress in CAD research has been made, there are still significant barriers to widespread clinical implementation. Current research suggests that developing CADe and CADx simultaneously can optimize performance. High-quality databases are required to train CAD systems, while privacy and security are crucial to ethically and legally share patient data for research. Furthermore, the lack of a standardized performance assessment tool has persistently made comparison between published literature difficult. Successful collaboration between professionals in the field of medicine and computer science is necessary to improve the effectiveness of CAD systems and, by extension, patient care.

REVIEWED BY: DR. DAVID KOFF

Dr. David Koff is a Professor of radiology and Chair of the Department of Radiology at McMaster University. He leads research projects on radiation risk, the validation of technology, and applications of artificial intelligence to medical imaging. He is currently the Chair of Canada Safe Imaging (CSI), an initiative he launched to promote radiation safety in Canada.

EDITED BY: NICK TELLER & TAAHA HASSAN