Artificial Intelligence for Lung Cancer Research and Considerations

Junchao Sun

College of Bioengineering, Qilu University of Technology, Shandong, China.

Abstract

Lung cancer is a common tumor with high lethality, and with the rapid development of artificial intelligence (AI) technologies, great progress has been made in the detection, diagnosis, and efficacy prediction of lung cancer.This review is focused on AI research applications, among others, with a view to improving clinical awareness of AI research.

Keywords

Artificial Intelligence; Lung Cancer; Feature Extraction.

1. Introduction

Lung cancer is a common malignancy affecting approximately 1.8 million people per year worldwide [1], and the incidence has significantly increased over the past 50 years [2].

Lung cancer is clinically divided into two categories, small cell lung cancer (NSCLC) and non-small cell lung cancer (NSCLC), of which NSCLC accounts for 80%, and the most common histological subtypes of NSCLC are adenocarcinoma and squamous cell carcinoma (SqCC) [3]. According to data published by the American Cancer Society in January 2020, the incidence and mortality of lung cancer ranks first among malignant tumors, and early diagnosis and treatment can effectively extend the survival time of patients. CT is the most common method for lung cancer screening and clinical diagnosis, and has always played an important role in the early detection and follow-up treatment of lung cancer. With the progressive popularity of low-dose CT lung cancer screening technology, the number of lung nodules detected has been rising [4]. As the number of tests continues to rise, the workload of radiologists continues to increase. At the same time, the modern diagnosis and treatment of lung cancer (lung cancer) demand is constantly rising for many aspects, such as the requirements of doctors and the requirements of diagnosis.In recent years, with the continuous development of computer technology, artificial intelligence technology gradually walks into people's horizons and has received extensive attention.Using artificial intelligence technology can provide physicians with technical support for lung cancer diagnosis [6-8]. In this review, we give an introduction to AI technologies and AI research applications for lung cancer imaging are summarized below.

2. Introduction to artificial intelligence technology

Artificial intelligence (AI), a new technological science in the study, development of theories, methodological techniques for modeling, extending and expanding human intelligence, and application systems [8]. The most representative are artificial neural networks, support vector machines, and convolutional neural networks. AI technology is the current hot revolutionary technology, combined with medical imaging, trained by retrospective machine learning of large sample cases to reach the purpose of diagnosis and prediction of diseases [9]. AI in clinical practice is the diagnostic and predictive analysis of various medical images, including photographs, pathology photomicrographs, and radiological images of retinal and skin lesions, among others [10-13]. Can significantly improve the comprehensive performance of lung cancer medical imaging in terms of diagnostic speed, accuracy as well as reporting quality. Artificial intelligence includes machine learning; And deep learning, an advanced technology in machine learning, is now as the most commonly used method for artificial intelligence research in lung cancer (Figure 1). Deep learning contains multiple working layers, such as the visual layer, with the workflow going from the initial layer to the uppermost layer [14]. AI applying deep learning techniques can be a model is continually optimized to match practical needs, and the research path is widely used.

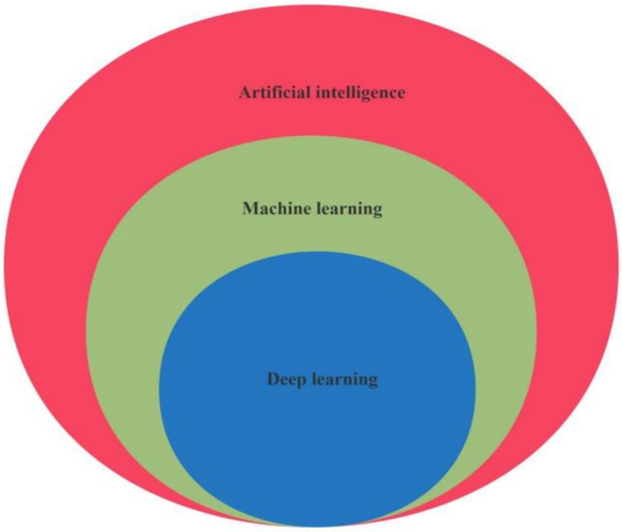

Fig. 1 A diagram of the categorical relationship between AI and machine learning, deep learning

Radiomics employs feature building models to diagnose and predict clinical diseases mainly based on high-throughput extraction of features from medical images [15-16]. The entire analysis pipeline includes image acquisition, lesion delineation as well as feature extraction, feature selection, and machine modeling. Deep learning models are an artificial intelligence method that continually learns to optimize, in contrast to omics models that are well established, model parameters generally do not produce modifications, and do not have the means to optimize autonomously. Therefore, radiomics is not a simple AI approach, but a comprehensive use of AI with partial techniques and statistical analysis, and its requirements for study subject sample quality as well as sample size are less demanding than those of machine learning and deep learning. According to recent literature observations [17-19], its research applications in imaging medicine and nuclear medicine are wider.

3. Artificial intelligence research applications in lung cancer

Lung cancer has been a research hotspot in clinical medicine and life sciences because of its high incidence combined with high lethality. Accompanied by the dramatic development of AI technologies, lung cancer has largely focused on the detection, diagnosis as well as prediction of lung lesions in etiological analysis, various diagnoses (e.g. live testing, histopathology, genetic classification, etc.), prognostic evaluation, and treatment options preferred [20-26]. Using artificial intelligence technology can significantly improve the detection efficiency of lung cancer and scientifically manage patients [20-21], multicenter studies by sim et al [21] showed that combining deep convolutional neural network can improve the reading sensitivity of chest X-ray images from 65.1% to 70.3% and false positives of single images from 0.20 to 0.18 in patients, respectively. Liu et al [22] study used convolutional neural network to establish deep learning model, and the fully automated detection of CT scan images of 12754 patients, whether solid nodules, or subsolid nodules in different size locations, the sensitivity of the fully automated detection model was higher than two radiologists, the total positive rate of the model was 70.1%, higher than two radiologists (39.4% and 56.6%), High false positives (48.4%) existed for the model. The study by Hawkins et al23 explored the ability of radiomics analysis of baseline low-dose CT images to detect and predict subsequent carcinogenesis of pulmonary nodules, investigated the use of data from the US national lung cancer screening trial to detect lung nodules, acquired 23 stable features by random forest classifiers, and built a lung nodule model to predict 1 - and 2-year carcinogenesis, The accuracy could reach 80% and 79% (AUCs of 0.83 and 0.75, respectively).

Traditional features to differentiate the pathological types and cover of gene mutations in lung cancer although progress has been made, it still faces great challenges [24-25].With the help of research methods of artificial intelligence such as radiomics and deep learning, image information features that cannot be obtained by the naked eye can be obtained, with excellent performance in discriminating lung cancer pathological types and gene mutations [26].

Wang et al [24] retrospectively analyzed the predictive ability of imaging omics technique in differentiating benign and malignant lung cancer nodules, the study acquired CT images of 593 patients with lung tumors in the international open database lung imaging database consortium and image data resource program, quantitatively extracted 150 imaging omics features, after data processing and modeled with support vector machine, Finally, the sensitivity, specificity and accuracy were 82.5%, 89.5%, 86.0% and 74.6%, 78.9%, 76.1% in the experimental and validation cohorts, respectively. Using radiomics features to review 58 invasive lung adenocarcinomas and 28 preinvasive lesions, Chen et al27 developed an ANN model by calculating its mean CT, standard deviation of CT values, quality, kurtosis, and entropy values, and the AUC reached 0.981 for discriminating between invasive and preinvasive lesions. Gao et al28 showed that the radiomics texture feature identification of ground glass nodular Aspen invasive lung adenocarcinoma on CT plain scan thin section images had a high value (AUC) of 0.890) compared with enhanced scan (AUC of 0.868), and the difference was not statistically significant (P = 0.1897). ROIs et al29 identified cases positive and negative for the growth factor receptor (EGFR) by using radiomics, and after combining the radiomics signature (AUC of 0.69) with clinical features associated with EGFR (AUC of 0.70), the combined model AUC reached 0.75, and patients with EGFR positive and KRAS positive tumors could be identified using radiomics (AUC of 0.80).

Artificial intelligence techniques have been more studied and applied not only in the detection and diagnosis of lung cancer, but also in the prediction of prognosis survival, radiotherapy and response to immunotherapy in lung cancer patients [30]. Song et al31 retrospectively analyzed the CT radiomics features of lung cancer (1 032) for EGFR mutations in patients with non-small cell lung cancer (NSCLC) stage w disease and modeled the prediction of progression free survival in this group of patients treated with EGFR TKIs, showing that, compared with the clinicopathological characterization model, The combined model of combined radiomics yielded significant clinical benefit (P < 0.0001). Using radiomics features, Huang et al32 retrospectively evaluated disease-free survival (DFS) in 282 patients with early-stage (I a to Hb) NSCLC and revealed that radiomics features were significantly associated with DFS and were significant predictors independent of clinicopathological risk factors, whereas the c-index of the nomogram for predicting DFS was 0.72 (95% CI 0.71-0.73), Above the clinicopathological risk factor C index of 0.69 (95% CI 0.68 to 0.70) (P < 0.000 1).Pyka et al33 retrospectively analyzed data of stereotactic body radiotherapy from 45 NSCLC patients from whom PET/CT image data were extracted to study the correlation between stereotactic body radiotherapy and local recurrence and survival, and showed entropy values and correlations with local recurrence in both subgroups of NSCLC patients at T1 (tumor diameter W3 cm) and T2 (tumor diameter ≥ 3 cm), The AUC was 0.801 and 0.776, respectively, and the entropy value was found to be an independent predictor of disease-related survival in multiple regression analysis (HR = 7.48, P = 0.016).Recently, Daniel Xie, a scientist at Google health research unit (GHR) in the USA, and simultaneously, developed a new deep learning model. Trained with 42290 CT scan images, the system predicted the prevalence of lung nodules even without human involvement. Showed that the AI system had an accuracy of 94% and could detect extremely small pulmonary nodules in 6716 tested cases. In the case of unlimited anterior CT scan images, the system performed better than 6 expert radiologists; But these findings need large-scale clinical validation, but the model will advance improved scientific management and prognosis of lung cancer patients in the future.In January 2021, the US FDNA analysis technology company trained artificial intelligence to identify rare genetic syndromes using ten thousand images of real patient faces.

4. Limitations and challenges of AI research in lung cancer imaging

Lung cancer imaging artificial intelligence research has achieved certain results, but it is still in the primary stage of development.Therefore, the limitations of its research applications and the challenges to be faced need to be recognized: ① sample size problem [34-35], most of the studies are based on a single center and have a small number of cases, which may lead to biased study results.The relative paucity of cases also resulted in a large proportion of studies adopting an imageomics approach rather than a more intelligent deep learning or machine learning approach.Therefore, it is necessary to establish a database with large sample size in multicenter, based on which the lung cancer research with relevant artificial intelligence can obtain more objective and scientific results. ② Delineation or labeling problems, for technical reasons, automatic delineation or labeling cannot be achieved in many cases for lung cancer lesions, whereas manual manual delineation, semi-automatic delineation, or labeling causes reproducibility problems.Since there is currently no large sample to support deep learning training for automatic delineation or labeling models, automated delineation or labeling software has not been widely applied. With the development of computer technology and the construction of multicenter large sample database, the automatic delineation or marking technique will be widely verified. ③ The problems of standardization and standardization of models, because they are at the primary stage of development, various models, software are specific and specialized, lack standards, is harder to standardize, unified, which will lead to problems with the reproducibility and comparability of research results. Therefore, relevant research teams need to collaborate as early as possible to establish a unified standard feature calculation technique and method that facilitates the research of radiomics features and high-quality model construction by deep learning, machine learning, etc. in order to further achieve efficient, precise diagnosis and prediction of target diseases. AI is being driven globally into clinical applications, but AI's work path, with its enormous potential, can carry enormous risks while helping physicians and entrepreneurs gain untapped medical knowledge. Hospitals with large technology companies are advancing AI applications to the clinic, but evidence that AI helps physicians detect missed diseases is now sparse and machine learning models are opaque. And it remains to be confirmed whether AI will show as much accuracy in a laboratory for a diverse set of patients of different race and community. When AI is at its best side and at its worst, this is the opportunism of launching over hype and revealing innovations that are the real solution.

5. Summary

Now, we have reason to believe that the careful application of AI in appropriate settings can make medical treatment more efficient and even life saving.But the challenge of AI is that it is critical to do so without violating privacy as well as data rights, not letting offsets in the data perpetuate, and not letting technology interfere with medicine.The commonly used artificial intelligence research methods for lung cancer imaging are deep learning developed on machine learning techniques, and the imaging methods associated with it.AI is highly valuable for tumor detection, efficient diagnosis, and efficacy prediction, but it is still in its early stage of development, and there are limitations such as small sample size, manual or semi-automatic predominance of lesion delineation and labeling, and lack of standards and specifications for feature or model calculation methods, which should be explored in lung cancer imaging.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017[J]. CA Cancer J Clin, 2017, 67: 7-30.

[2] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012[J]. CA Cancer J Clin, 2015, 65: 87-108.

[3] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015 [J]. CA Cancer J Clin, 2016, 66: 115-132.

[4] KAKINUMA R,OHMATSU H,KANWKO M,et al. Progression of focal pure ground-glass opacity detected by low-dose helical computed tomography screening for lung cancer [J]. J Comput Assist Tomogr, 2004,28(1):17-23.

[5] Roque Ramos L, DiMaio CJ, Sachar DB, et al. Autoimmune pancre atitis and inflammatory bowel disease: case series and review of the literature[J]. Dig Liver Dis, 2016, 48: 893-898.

[6] Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges[J]. Magn Reson Imaging, 2012, 30: 1234-1248.

[7] Lambin P, Rios -Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012, 48: 441-446.

[8] PARK S H, HAN K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction [J]. Radiology, 2018 , 286 (3): 800-809.

[9] GULSHAN V, PENG L, CORAM M, et al. Development and vali- dation of a deep learning algorithm for detection of diabetic retinop athy in retinal fundus photographs[J]. JAMA,2016,316 (22):2402-2410.

[10] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. Nature,2017,542(7639): 115-118.

[11] EHTESHAMI B, VETA M, JOHANNES D P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer[J]. JAMA, 2017, 318(22):2199-2210.

[12] RECHT M P, DEWEY M, DREYER K, et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommen dations [J/OL]. Eur Radiol, 2020. [2020-02-17]. https:// www. ncbi. nlm. nih. gov/ pubmed/ term =Integrating +artificial +intellig ence +into +the +clinical +practice +of +radiology %3A +challenges +and +recommendations. DOI: 10.1007/ s00330-020-06672-5.

[13] CHEN J H, ASCH S M. Machine learning and prediction in medicine-beyond the peak of inflated expectations[J]. N Engl J Med,2017,376(26):2507-2509.

[14] GILLIES R J, KINAHAN P E, HRICAK H. Radiomics: images are more than pictures, they are data[J]. Radiology, 2016, 278(2):563-577.

[15] WILSON R, DEVARA J A. Radiomics of pulmonary nodules and lung cancer[J]. Transl Lung Cancer Res, 2017,6(1):86- 91.

[16] THAWANI R, MCLANE M, BEIG N, et al. Radiomics and radiogenomics in lung cancer: a review for the clinician [J]. Lung Cancer, 2018,115: 34-41.

[17] LAMBIN P, LEI T H, DEIST T M, et al. Radiomics: the bridge between medical imaging and personalized medicine [J]. Nat Rev Clin Oncol, 2017,14(12) :749-762.

[18] WANG B, TANG Y, CHEN Y, et al. Joint use of the radiomics method and frozen sections should be considered in the prediction of the final classification of peripheral lung adenocarcinoma manifesting as ground-glass nodules [J]. Lung Cancer, 2020, 139: 103-110.

[19] HASSANI C, VARGHESE B A, NIEVA J, et al. Radiomics in pulmonary lesion imaging [J]. AJR Am J Roentgenol, 2019, 212(3):497-504.

[20] SIM Y, CHUNG M J, KOTTER E, et al. Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs [J]. Radiology, 2020, 294:199-209.

[21] LIU K, LI Q, MA J, et al. Evaluating a fully automated pulmonary nodule detection approach and its impact on radiologist performance [J]. Radiology: Artificial Intelligence, 2019,1(3): e180084.

[22] HAWKINS S, WANG H, LIU Y, et al. Predicting malignant nodules from screening CTs[J]. J Thorac Oncol,2016,11(12):2120-2128.

[23] WANG J, LIU X, DONG D, et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method [J]. Conf Proc IEEE Eng Med Biol Soc, 2016,2016:1272-1275.

[24] IKEDA K, AWAI K, MORI T, et al. Differential diagnosis of ground-glass opacity nodules: CT number analysis by three dimensional computerized quantification[J]. Chest, 2007, 132(3):984-990.

[25] COUDRAY N, OCAMPO P S, SAKELLAROPOULOS T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning[J]. Nat Med, 2018, 24(10):1559-1567.

[26] CHAE H D, PARK C M, PARK S J, et al. Computerized texture analysis of persistent part-solid ground-glass nodules: differenti-ation of preinvasive lesions from invasive pulmonary adenocarcinomas [J]. Radiology, 2014,273(1):285-293.

[27] GAO C , XIANG P, YE J, et al. Can texture features improve the differentiation of infiltrative lung adenocarcinoma appearing as ground glass nodules in contrast-enhanced CT? [J]. Eur J Radiol, 2019, 117:126-131.

[28] RIOS V E, PARMAR C, LIU Y, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer [J]. Cancer Res, 2017,77(14):3922-3930.

[29] BANNA G L, OLIVIER T, RUNDO F, et al. The promise of digital biopsy for the prediction of tumor molecular features and clinical outcomes associated with immunotherapy [J]. Front Med, 2019, 6:172.

[30] SONG J, SHI J, DONG D, et al. A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy [J]. Clin Cancer Res, 2018,24 (15):3583-3592.

[31] HUANG Y, LIU Z, HE L, et al. Radiomics signature:a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer [J]. Radiology, 2016, 281(3): 947-957.

[32] PYKA T, BUNDSCHUH R A, ANDRATSCHKE N, et al. Textural features in pre-treatment [F18] -FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy [J]. Radiat Oncol, 2015,10: 100.

[33] ZHANG Y, OIKONOMOU A, WONG A, et al. Radiomics-based prognosis analysis for non-small cell lung cancer[J]. Sci Rep, 2017, 7: 46349.

[34] FIGUEROA R L, ZENG T Q, KANDULA S, et al. Predicting sample size required for classification performance [J]. BMC Med Inform Decis Mak, 2012, 12:8.