European Journal of Radiology 123 (2020) 108774

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Artificial intelligence applications for thoracic imaging

Guillaume Chassagnon[a,b,d], Maria Vakalopoulou[b,d], Nikos Paragios[b,c,d], Marie-Pierre Revel[a,d,*]

[a] Radiology Department, Groupe Hospitalier Cochin Broca Hôtel-Dieu – Université Paris Descartes, 27 Rue du Faubourg Saint-Jacques, 75014, Paris, France
[b] Laboratoire Mathématiques et Informatique pour la Complexité et les Systèmes, Ecole CentraleSupelec, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France
[c] TheraPanacea, 27 Rue du Faubourg Saint-Jacques, 75014, Paris, France
[d] Center for Visual Computing, Ecole CentraleSupelec, 3 Rue Joliot Curie, 91190, Gif-sur-Yvette, France

ARTICLE INFO

Keywords:
Artificial intelligence
Deep learning
Machine learning
Thoracic imaging

ABSTRACT

Artificial intelligence is a hot topic in medical imaging. The development of deep learning methods and in particular the use of convolutional neural networks (CNNs), have led to substantial performance gain over the classic machine learning techniques. Multiple usages are currently being evaluated, especially for thoracic imaging, such as such as lung nodule evaluation, tuberculosis or pneumonia detection or quantification of diffuse lung diseases. Chest radiography is a near perfect domain for the development of deep learning algorithms for automatic interpretation, requiring large annotated datasets, in view of the high number of procedures and increasing data availability. Current algorithms are able to detect up to 14 common anomalies, when present as isolated findings. Chest computed tomography is another major field of application for artificial intelligence, especially in the perspective of large scale lung cancer screening. It is important for radiologists to apprehend, contribute actively and lead this new era of radiology powered by artificial intelligence. Such a perspective requires understanding new terms and concepts associated with machine learning. The objective of this paper is to provide useful definitions for understanding the methods used and their possibilities, and report current and future developments for thoracic imaging. Prospective validation of AI tools will be required before reaching routine clinical implementation.

1. Introduction

Artificial intelligence has become a hot topic in radiology these last years, with already 150 deep learning articles only focusing on medical imaging in 2018 [1]. Machine learning gives computers the ability to learn from data and reproduce human interpretations without being explicitly programmed. Computer vision, a scientific field of particular interest for radiologists, shares a number of objectives with machine learning. The goal is to make it possible for machines to analyze, process and understand digital to automate tasks that the human visual system can do.

Traditional machine learning techniques have been used from the 80 s. The availability of large datasets and the increasing computing capabilities driven from the development of graphic processing units (GPU) have been the foundational elements for the development of neural networks, and especially of convolutional neural networks (CNNs) [2]. Neural networks are designed to mimic the way the human brain processes information. They combine multiple formal neurons, each of them processing part of the information, with their intelligent combination leading to the final decision rule. Deep neural networks are characterized by a specific configuration where neurons are organized in multiples layers. Deep learning currently represents the state-of-the art in machine learning for a variety of applications. Classification, disease detection and segmentation have been primarily the tasks that benefited from the development of deep learning. Computational imaging encompasses other objectives as well, such as radiomics. Radiomics consist of extracting and analyzing imaging features (e.g. characteristics which are invisible to the human eye) such as shape, texture, or intensity of voxels [3]. The selected features can be used for classification purposes, such as the benign or malignant nature of a lung nodule but also for evaluating the prognosis or the probability of response to treatment of lung malignancies. Radiomic analysis relies on statistics, using various methods such as clustering or dimensionality reduction, random forest, linear regression and others. Computer aided diagnosis (CAD) refers to tools which have been developed to assist image reading. The main difference between classical machine learning and deep neural networks is that deep learning replaces the process of feature extraction and disease classification of the traditional CAD systems, but requires large datasets for training.

Thoracic imaging is an important domain for developing solutions

* Corresponding author at: Service de Radiologie A, Groupe Hospitalier Cochin Broca Hôtel-Dieu, AP-HP, 27 Rue du Faubourg Saint-Jacques, 75014, Paris, France.
E-mail address: marie-pierre.revel@aphp.fr (M.-P. Revel).

https://doi.org/10.1016/j.ejrad.2019.108774
Received 27 June 2019; Received in revised form 13 November 2019; Accepted 21 November 2019

based on artificial intelligence, for several reasons. If we consider chest radiography (CXR), it is one of the most frequently performed procedures in medicine. In all institutions, it represents a huge work load, and in some countries there are not enough radiologists to interpret CXR. Furthermore, detection errors are common due the low contrast between lesions and the surrounding lung and the superposition of bone structures and computed assisted tools can be helpful. CAD has been reported to improve reader accuracy for the detection of lung cancers previously missed on CXR [4]. Lung cancer screening with low dose CT has been demonstrated to reduce lung cancer-related mortality [5] and it is expected that after the United States, European countries could start large scale screening [6] which will require a lot of resources in view of the huge number of eligible patients. This is another reason why AI-based algorithms could assist thoracic radiologists.

In this review, we will present the current state of AI in thoracic imaging, an active field of research for the development of AI algorithms, starting with the applications for chest radiography and then discussing those for chest computed tomography.

2. Artificial intelligence applied to chest radiograph (CXR) reading

The World Health Organization estimates that two thirds of the global population lack access to imaging and radiology diagnostics [7]. Thoracic imaging techniques such as digital chest radiography have the major advantage to be easy to use and affordable, even in developing or underdeveloped areas. It consists of 2D images and several billions have already been stored on picture archiving and communication systems (PACS) and linked to radiological reports. However, there is a shortage of experts who can interpret chest radiographies, even when imaging equipment is available, which opens tremendous perspectives for the impact of artificial intelligence applied to thoracic imaging.

The first application of artificial intelligence is workflow optimization, by detecting CXR with possible abnormalities that should be read first among all CXR of the work list. Using density and texture-based features, Kao et al. developed a CAD system to automatically determine abnormal chest examinations in the work list of radiologists interpreting chest examinations. The turnaround time for reporting abnormal CXR was reduced by 44 % [8]. CAD can be used for specific detection tasks on chest radiograph, such as detection of tuberculosis, pneumonia or lung nodule, and even more advances tasks such as multiple disease detection are being developed as well [9].

2.1. Lung nodule diagnosis

Among specific tasks, a major application of CAD is the diagnosis of lung nodules on chest radiography. This includes CAD for detection (CADe), and CAD for characterization (CADx) used to evaluate the nodule probability of malignancy or a combination of them. Whereas radiomics is often used for CADx, either using deep learning or classic machine learning techniques, the current tendency for developing CADe tools is to use deep learning.

Traditional pulmonary nodule CAD systems include image preprocessing, nodule detection using various algorithms, extraction of features and classification of the candidate lesions as nodules and non-nodules. The number of selected features (intensity, shape, texture, size) and the machine learning algorithm used for classification (support vector machine, Fisher linear discriminant and others) depend on the CAD system. The objective is to have adequate sensitivity with a low number of false positives detections. The development of convolution neural networks has opened new perspectives, but require large annotated chest radiograph datasets, in order to avoid under and overfitting (defined in upcoming section). Transfer learning could overcome this requirement. It consists on training algorithm non-medical, everyday images on a large data set and initializing the network with its parameters on the smaller medical image dataset. Bush et al. pre trained a CNN model on a subset of the ImageNet dataset which contains millions of labeled real-word images and retrained it to classify chest radiographs as positive or negative for the presence of lung nodules with a sensitivity of 92 % and a specificity of 86 % [10].

More recently Nam et al. developed a deep learning-based detection algorithm for malignant pulmonary nodules on chest radiographs and compared its performance with that of physicians, with half of them being radiologists. They used a dataset of 43 292 chest radiographs with a normal to diseased ratio of 3.67. Using an external validation dataset, they found AUC of the developed algorithm was higher than that of 17 of the 18 physicians. All physicians showed improved nodule detection when using the algorithm as second reader [11].

2.2. Tuberculosis diagnosis

Automated detection of tuberculosis on chest radiographs is another important field of research. Tuberculosis is an important cause of death worldwide, with a high prevalence in underdeveloped areas where radiologists are lacking. Several approaches have been used to detect tuberculosis manifestations in CXRs. Traditional machine learning approaches mainly used textural features, with or without applying bone suppression as pre-treatment of CXR images. Rohmah et al. used statistical features in the image histogram to identify TB positive radiographs and reached an accuracy of 95.7 % [12]. Others used a combination of textural, focal, and shape abnormality analysis [13]. Hwang et al. introduced a deep learning for automated detection of active pulmonary tuberculosis on chest radiographs. Their solution outperformed physicians including thoracic radiologists [14]. Lakhani et al. retrained two CNNs (AlexNet and GoogLeNet) pre-trained on non-medical images on a dataset of 1007 CXRs, with an equivalent number of positive and negative tuberculosis cases. The AUC was 0.99 for the best performing classifier combining the 2 pre-trained CNNs which were activated on areas of the lung where the disease was present, in the upper lobes [15]. However, as acknowledged by the authors, the model was trained for a specific task, which was differentiating normal versus abnormal CXR regarding tuberculosis suspicion. The model might falsely consider as tuberculosis positive chest radiographs with pathologic findings having a similar radiographic appearance, such as lung cancers and bacterial pneumonia. This limits the use of the algorithm to areas of high tuberculosis prevalence and few mimickers, such as lung cancer also affecting the upper lung zones.

2.3. Detection of pneumonia

In addition to pulmonary nodules and tuberculosis there are acute conditions that can be detected using such computer-aided solutions, like pneumonia. Rajpurkar et al. trained Chexnet, a deep learning algorithm for pneumonia detection of and compared its performance to that of 4 radiologists, using F1 score metric. Their model performed better than the averaged radiologists even though no better than the best radiologist [16].

2.4. Detection of common chest radiograph anomalies

Beyond lung nodule detection or other specific detection tasks, detection of multiple abnormalities is more challenging but in phase with the clinical practice, since frequently there are multiple abnormalities in the chest radiographs. Automated chest radiography reading based on deep learning is currently an intense field of research. As previously mentioned, deep learning-based algorithms need to be trained on large datasets. The amount and the diversity of data are of major importance. In a perfectly balanced scenario it could produce optimal results but often two other phenomena are observed, namely overfitting and underfitting. Underfitting occurs when the model fails to does not perform well on both training and validation datasets. Overfitting refers to model achieving excellent performance on the training data, but low

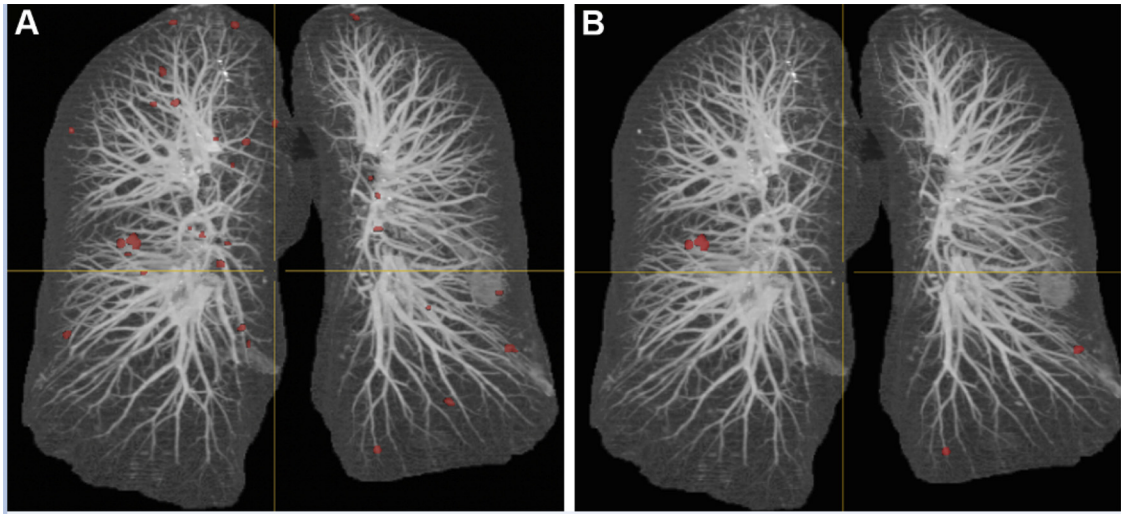

Fig. 1. CADe for lung nodule detection. Using commercially available software (Thoracic VCAR, GE healthcare, Buc, France), the number of candidate lesion is very high (32 red spots) when the sensitivity is adjusted at its highest level, to detect lung nodules from 2 mm (A). Using a different setting, to only detect nodules of at least 4 mm, the number of candidate lesions goes down to 5 (B). When adjusted to 6 mm, no more candidate lesions are detected (not shown).

performance on new, unseen images which is often called poor generalization. Generalization is a critical point for deep learning. If neural networks are trained on standardized images, acquired with the same protocol from only one vendor the risk is that the developed algorithm will not perform well for new unseen cases acquired with different parameters. This is also the reason why the data must be split into training, validation and test datasets with performances ideally validated on an external independent cohort. Once the training and testing datasets have been defined, the next required step towards leveraging artificial intelligence - in particular for methods such as deep learning - refers to the annotation of data. For image analysis there are several ways to generate annotated datasets depending on the level of detail that these annotations will provide and depending on the task that they address.

For classification tasks, labeling is a way of annotation where images are globally labeled with one or several classes (labels) that is often called weak annotations. The exact localization of the anomaly is not necessarily provided on the image. Several large databases of annotated chest radiographies are publicly available for developing research projects. One of the largest databases is chestX -ray8, already mentioned, built from the clinical PACS of the hospitals affiliated to the National Institute of Health. This database includes 112,120 frontal views of 30,805 patients and initially the image labels of 8 diseases, then extended to 14 diseases (chestX –ray14), including atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis, effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass, and hernia.

Rajpurkar et al. compared the performance of Chexnet, trained on chestx-ray14 dataset to that of 9 radiologists on a validation set of 420 images containing examples of the pathology labels. The radiologists achieved statistically significantly higher AUC performance on cardiomegaly, emphysema, and hiatal hernia, whereas for other pathologies, AUCs reached with the algorithm were either significantly higher (atelectasis) or with no statistically significant difference (other 10 pathologies) [17]. Hwang et al. developed a deep learning-based algorithm able to distinguish normal and abnormal chest radiograph results, including malignant neoplasm, active tuberculosis, pneumonia, and pneumothorax. The algorithm was trained on a dataset of 54 221 normal chest radiographs and 35 613 with abnormal findings. External validation using486 normal and 529 abnormal chest radiographs was performed. With a median 0.979 AUC, the algorithm demonstrated significantly higher performance than non-radiology physicians, board-certified radiologists, and thoracic radiologists. Human readers' performance increased when assisted by AI algorithms[18].

These results open new perspectives, but it is noteworthy that the algorithm was tested for anomalies present as isolated pathologic findings. The performance was not evaluated for combined abnormalities, which is critical for a successful implementation in clinical practice.

3. Artificial intelligence applied to chest computed tomography (CT) reading

3.1. Lung nodule

3.1.1. Lung nodule detection

The application of medical image analysis to thoracic CT is not a novel research area. CAD has been used for automated lung nodule detection on CT. Early approaches at the beginning of the 2000's were based on traditional machine learning approaches, such as Support Vector Machines (SVMs). Commercially available computer-aided detection packages were proposed by companies like Siemens (Lung Vcare) General electric (CT ALA for advanced lung analysis), R2 Technology (Image checker) and others.

Even though none of the two large randomized lung cancer screening studies, NLST (National lung cancer screening trial) [19]and NELSON [20] used CAD for lung nodule detection, an ancillary study from the NELSON group, published in 2012 [21] compared CAD and double reading by radiologists, in a cohort of 400 CT scans randomly selected from the NELSON database. The lung CAD algorithm used in this study was commercial software from Siemens, available since 2006 (LungCAD VB10A). Ground truth was established by a consensus reading from expert chest radiologists. The sensitivity for lung nodule detection was 78.1 % for double reading and 96.7 % for CAD, at an average cost of 3.7 false positive detections per examination. However, there were only 5 subsolid nodules (either non-solid or part-solid) in the 400 selected CT scans, and 2 of them were not detected by CAD. Using another commercial CAD, only 50 % of subsolid nodules were detected at best with the highest sensitivity setting, at the average cost of 17 CAD marks per CT [22]. Visual confirmation remains necessary for reducing false positives when using a CAD for the detection of subsolid nodules [23].

For solid nodules, easier to detect, sensitivity should be adjusted to only detect lung nodules of at least 6 mm, according to lung-RADS and Fleischner guidelines [24]. This is a way to limit the false positive detections, and the number of candidate lesions to evaluate (Fig. 1).

Using CAD for the detection of lung nodules in patients with extra-thoracic malignancies improved the detection of lung nodules, at the cost of an 11 % increase of reading time [25].

3.1.2. Lung nodule volumetry

If not used for lung nodule detection, CAD has been used for calculating the volume of screen-detected nodules in the NELSON study and estimating the volumetry-based doubling time. This strategy was the basis of lung nodule management in the NELSON study, nodules of less than 50 mm^3 were considered as negative screen together with lung nodules between 50 and 500 mm^3 for which the volumetry-based doubling time, calculated at 3 months was more than 400 days [26]. This strategy was proven to notably reduce the false positive rate. The ratio of positive screens (true and false positives) was 6.6 %, in the NELSON study, compared to 24.1 % in NSLT, where nodule diameters were manually measured and any nodule of at least 4 mm was considered as a positive screen. The limits of diameter manual measurements are well known, the intra and inter reader repeatability are 1.4 and 1.7 mm respectively [27], which does not allow to reliably detect malignant growth at 3 months for nodules of less than 10 mm. Conversely, software-based volumetric measurements are highly repeatable [28] and doubling times of more than 500 days for solid nodules have a 98 % negative predictive value for the diagnosis of malignancy [29]. This is the reason why the European position statement recommends volume measurement and volume-doubling time estimation for the management of detected solid nodules [6].

Volumetry software are less reliable for subsolid nodules, even though doubling times of solid and nonsolid component of part-solid nodules can be separately estimated (Fig. 2).

3.1.3. Approach by radiomics

Besides volumetry-based doubling time estimation, another approach for lung nodule characterization is to use radiomics to analyze imaging features derived from medical images. Radiomics can be used to characterize tumor aggressiveness, viability, response to chemotherapy and/or radiation [30]. Therefore, a radiomic approach can help to reveal unique information about tumor biological behavior. It can be used for prognosis estimation in confirmed lung cancers [31] or to estimate the risk of distant metastasis [32]. Radiomics has also been used to predict histology and mutational profile of lung tumors [33]. Using principal component analysis (PCA) on stable, reproducible features, the authors obtained a radiomic signature able to successfully discriminate between EGFR + and EGFR- cases, with an AUC of 0.69.

The problem of radiomics is the robustness and generalizability of the learned signatures. Indeed, radiomics analysis performed on images acquired under specific, homogeneous imaging conditions, are not representative of clinical routine [34]. Another important condition in the selection of features is their reproducibility.

3.1.4. Deep learning

The use of CNN for CT images is more complex than for 2D chest radiograph images, due to the 3D nature of images, the high number of slices and smaller size of datasets, requiring data augmentation techniques. To overcome these problems, some studies use 2D CNNs applied to each slice, whereas others choose to adopt a patch-based approach or reduce the image size at the cost - for both cases - of a loss of information. Despite these technical difficulties, results are promising and CNNs generally allow obtaining better results than traditional machine learning methods. Using deep learning, Zhao et al. obtained an AUC value of 0.758 for predicting EGFR mutation[35].

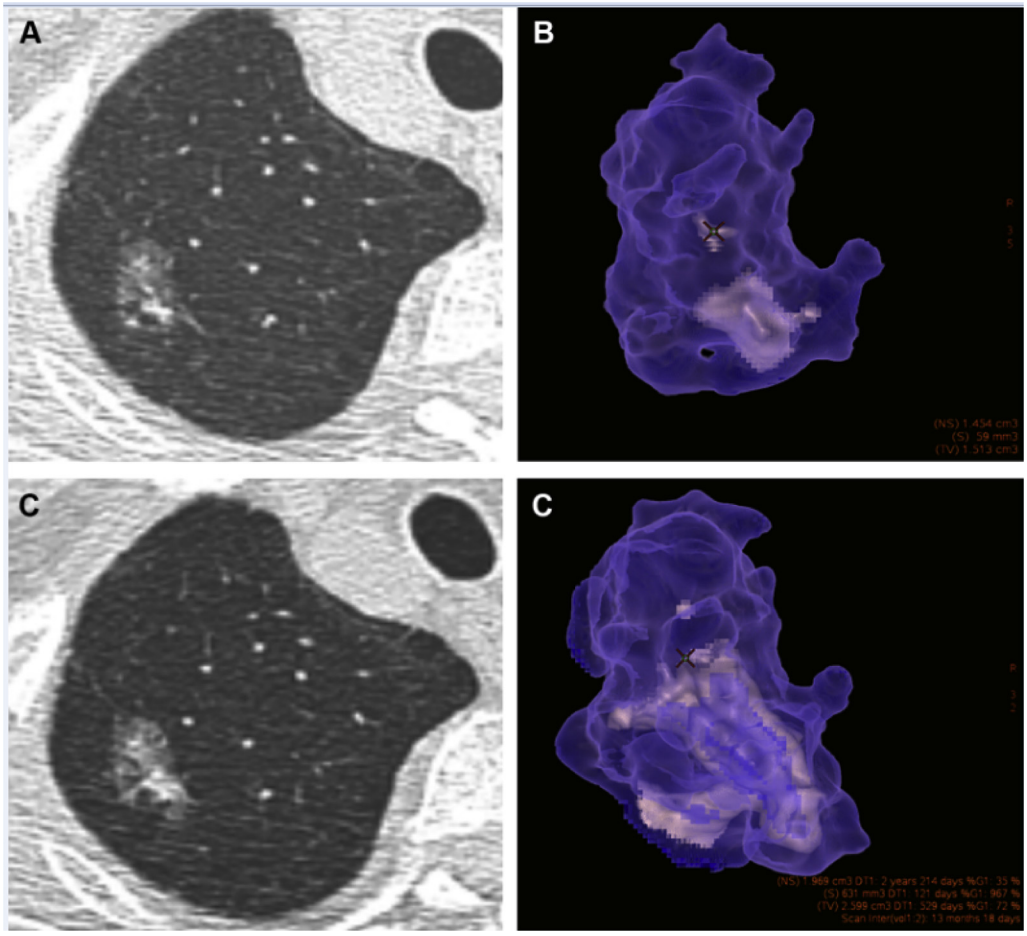

Fig. 2. Volumetry-based doubling time measurement of a part-solid nodule. Baseline CT demonstrates a part-solid nodule of the right upper lobe (A), with small solid component of 59 mm^3 (B). Follow-up CT performed 13 months later shows an increase of the solid portion (C). The whole nodule doubling time is 529 days (Thoracic VCAR software), relatively indolent, but the solid component doubling time is only 121 days, typically in the malignant range and reflecting aggressiveness, whereas the nonsolid component doubling time is almost 3 years (D).

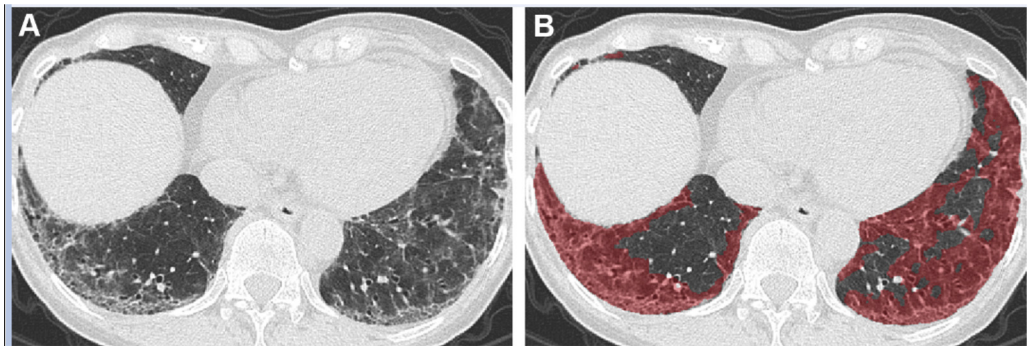

Fig. 3. Semantic segmentation of lung fibrosis, based on deep learning. Unenhanced CT scan axial transverse image through the lung bases, demonstrating ground glass reticulations and bronchiolectasis with subpleural predominance, typical for non specific interstitial pneumonia (NSIP) (A). Deep learning-based automated segmentation of fibrotic areas (AtlasNet), allowing calculating the volume of diseased lung on CT (B).

For the 2017 Kaggle Data Science Bowl (KDSB17), whose objective was to predict the cancer risk at 1 year, based on lung cancer screening CT examinations, all frontrunner teams used deep learning.

Ardila et al. trained a deep learning algorithm on a NLST dataset from 14,851 patients, 578 of whom having developed lung cancer within the next year [36]. They then tested the model on a first test dataset of 6716 cases, achieving an AUC of 94,4 %. Comparison to 6 radiologists was performed for a subset of 507 patients, and the model' performance was equivalent or higher to all of them when a single CT was analyzed, whereas performances were equivalent when the model and the radiologists made a decision including patients' previous CT scans.

3.2. Diffuse lung diseases

The use of CNN for thoracic CT is not restricted to nodule evaluation but can also be applied to diagnose and stage COPD and predict acute respiratory distress (ARD) and mortality in smokers [37]. Training a CNN on the CT scans of 7,983 COPDGene participants, AUC for the detection of COPD was 0.856 in a non overlapping cohort of 1000 another COPDGene participants. AUCs for ARD events were 0.64 and 0.55 in COPDGene and ECLIPSE participants, respectively.

CNNs can also be used for the detection and quantification of infiltrative lung diseases (ILD) or for automated classification of fibrotic lung diseases. Indeed, even though classification criteria have been established by consensus of 4 expert societies, inter radiologist agreement is only moderate at best, even among experts, and there is a shortage of experts [38].

Walsh et al. trained a CNN algorithm for automated classification of fibrotic lung disease on a database of 1157 high-resolution CT scans from two institutions showing evidence of diffuse fibrotic lung disease. When comparing the model performance with that of 91 radiologists, the model accuracy was 73·3 % compared to a median radiologist accuracy of 70·7 % [39].

The majority of previous work on ILD pattern detection was based on 2D image classification using a patch-based approach. This approach consists in dividing the lung into numerous small patches of the same size (e.g., 32 × 32 pixels) and to classify them into one of the ILD pattern classes. Different classifiers can be used, such as SVM, Boltzmann machines, convolutional neural networks (CNNs) local binary patterns and multiple instance learning [40]. These classifiers are trained on datasets including thousands of annotated patches, representatives of each class to identify, normal ground glass, honeycombing, emphysema. Caliper software was developed using the patch-based approach, for the quantification of disease extent and change in idiopathic pulmonary fibrosis [41,42]. The main advantages of this approach are the possibility to separately quantify each anomaly, and the need for only week annotation (e.g. categorization), which is less time consuming than semantic segmentation which requires precisely contouring disease extent on CT images.

However patch-based methods do not integrate spatial information such as the subpleural location and basal predominance. In the central lung portion, some bronchi may be misclassified as honeycombing. Furthermore, the results might be disappointing when the model is applied to the whole CT image, as problematic patches including more than one pattern might have been excluded from the training datasets, similarly to frontier patches at the very lung periphery, close to the chest wall. Another approach is the segmentation of the whole fibrotic extent without quantifying each component [43]. This requires contouring the abnormal fibrotic areas on every abnormal slices, which is time consuming but allow then applying the model to the whole lung volume (Fig. 3).

4. Conclusion

Machine learning has already been part of the radiologists' daily life for several years. Multiple applications are currently being developed with deep learning-based approaches and will require prospective clinical evaluation. These developments shouldn't be considered as a threat but more as an opportunity. Radiologists can benefit from workflow optimization, and gain performance for detection, characterization and quantification tasks, especially in the field of thoracic imaging. Human validation remains necessary as full automation of imaging tasks cannot be considered at this time.

Declaration of Competing Interest

None of the authors report conflict of interest

References

[1] L. Saba, M. Biswas, V. Kuppili, E. Cuadrado Godia, H.S. Suri, D.R. Edla, T. Omerzu, J.R. Laird, N.N. Khanna, S. Mavrogeni, A. Protogerou, P.P. Sfikakis, V. Viswanathan, G.D. Kitas, A. Nicolaides, A. Gupta, J.S. Suri, The present and future of deep learning in radiology, Eur. J. Radiol. 114 (2019) 14–24, https://doi.org/10.1016/j.ejrad.2019.02.038.

[2] S. Soffer, A. Ben-Cohen, O. Shimon, M.M. Amitai, H. Greenspan, E. Klang, Convolutional neural networks for radiologic images: a radiologist's guide, Radiology 290 (2019) 590–606, https://doi.org/10.1148/radiol.2018180547.

[3] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: Images Are More than Pictures, They Are Data, Radiology 278 (2016) 563–577, https://doi.org/10.1148/radiol.2015151169.

[4] S. Kligerman, L. Cai, C.S. White, The effect of computer-aided detection on radiologist performance in the detection of lung cancers previously missed on a chest radiograph, J. Thorac. Imaging 28 (2013) 244–252, https://doi.org/10.1097/RTI.0b013e31826c29ec.

[5] National Lung Screening Trial Research Team, D.R. Aberle, A.M. Adams, C.D. Berg, W.C. Black, J.D. Clapp, R.M. Fagerstrom, I.F. Gareen, C. Gatsonis, P.M. Marcus, J.D. Sicks, Reduced lung-cancer mortality with low-dose computed tomographic screening, N. Engl. J. Med. 365 (2011) 395–409, https://doi.org/10.1056/NEJMoa1102873.

[6] M. Oudkerk, A. Devaraj, R. Vliegenthart, T. Henzler, H. Prosch, C.P. Heussel, G. Bastarrika, N. Sverzellati, M. Mascalchi, S. Delorme, D.R. Baldwin, M.E. Callister, N. Becker, M.A. Heuvelmans, W. Rzyman, M.V. Infante, U. Pastorino, J.H. Pedersen, E. Paci, S.W. Duffy, H. de Koning, J.K. Field, European position statement on lung cancer screening, Lancet Oncol. 18 (2017) e754–e766, https://doi.org/10.1016/S1470-2045(17)30861-6.

[7] D.J. Mollura, E.M. Azene, A. Starikovsky, A. Thelwell, S. Iosifescu, C. Kimble, A. Polin, B.S. Garra, K.K. DeStigter, B. Short, B. Johnson, C. Welch, I. Walker,

D.M. White, M.S. Javadi, M.P. Lungren, A. Zaheer, B.B. Goldberg, J.S. Lewin, White Paper Report of the RAD-AID Conference on International Radiology for Developing Countries: identifying challenges, opportunities, and strategies for imaging services in the developing world, J. Am. Coll. Radiol. JACR 7 (2010) 495–500, https://doi.org/10.1016/j.jacr.2010.01.018.

[8] E.-F. Kao, G.-C. Liu, L.-Y. Lee, H.-Y. Tsai, T.-S. Jaw, Computer-aided detection system for chest radiography: reducing report turnaround times of examinations with abnormalities, Acta Radiol. 56 (2015) (1987) 696–701, https://doi.org/10.1177/0284185114538017.

[9] C. Qin, D. Yao, Y. Shi, Z. Song, Computer-aided detection in chest radiography based on artificial intelligence: a survey, Biomed. Eng. Online 17 (2018) 113, https://doi.org/10.1186/s12938-018-0544-y.

[10] I. Bush, Lung Nodule Detection and Classification Vol. 20 (2016), pp. 196–209.

[11] J.G. Nam, S. Park, E.J. Hwang, J.H. Lee, K.-N. Jin, K.Y. Lim, T.H. Vu, J.H. Sohn, S. Hwang, J.M. Goo, C.M. Park, Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs, Radiology 290 (2019) 218–228, https://doi.org/10.1148/radiol.2018180237.

[12] R.N. Rohmah, A. Susanto, I. Soesanti, Lung tuberculosis identification based on statistical feature of thoracic X-ray, 2013 Int. Conf. QiR, IEEE (2013) 19–26.

[13] L. Hogeweg, C.I. Sánchez, P. Maduskar, R. Philipsen, A. Story, R. Dawson, G. Theron, K. Dheda, L. Peters-Bax, B. van Ginneken, Automatic detection of tuberculosis in chest radiographs using a combination of textural, focal, and shape abnormality analysis, IEEE Trans. Med. Imaging 34 (2015) 2429–2442, https://doi.org/10.1109/TMI.2015.2405761.

[14] E.J. Hwang, S. Park, K.-N. Jin, J.I. Kim, S.Y. Choi, J.H. Lee, J.M. Goo, J. Aum, J.-J. Yim, C.M. Park, DLAD Development and Evaluation Group, Development and Validation of a Deep Learning-Based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs, Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. (2018), https://doi.org/10.1093/cid/ciy967.

[15] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, Radiology 284 (2017) 574–582, https://doi.org/10.1148/radiol.2017162326.

[16] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning, ArXiv Prepr (2017) ArXiv171105225.

[17] P. Rajpurkar, J. Irvin, R.L. Ball, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C.P. Langlotz, B.N. Patel, K.W. Yeom, K. Shpanskaya, F.G. Blankenberg, J. Seekins, T.J. Amrhein, D.A. Mong, S.S. Halabi, E.J. Zucker, A.Y. Ng, M.P. Lungren, Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists, PLoS Med. 15 (2018) e1002686, , https://doi.org/10.1371/journal.pmed.1002686.

[18] E.J. Hwang, S. Park, K.-N. Jin, J.I. Kim, S.Y. Choi, J.H. Lee, J.M. Goo, J. Aum, J.-J. Yim, J.G. Cohen, G.R. Ferretti, C.M. Park, DLAD Development and Evaluation Group, Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs, JAMA Netw. Open. 2 (2019) e191095, , https://doi.org/10.1001/jamanetworkopen.2019.1095.

[19] National Lung Screening Trial Research Team, D.R. Aberle, C.D. Berg, W.C. Black, T.R. Church, R.M. Fagerstrom, B. Galen, I.F. Gareen, C. Gatsonis, J. Goldin, J.K. Gohagan, B. Hillman, C. Jaffe, B.S. Kramer, D. Lynch, P.M. Marcus, M. Schnall, D.C. Sullivan, D. Sullivan, C.J. Zylak, The National Lung Screening Trial: overview and study design, Radiology 258 (2011) 243–253, https://doi.org/10.1148/radiol.10091808.

[20] Y. Ru Zhao, X. Xie, H.J. de Koning, W.P. Mali, R. Vliegenthart, M. Oudkerk, NELSON lung cancer screening study, Cancer Imaging Off. Publ. Int. Cancer Imaging Soc. 11 (2011) S79–84, https://doi.org/10.1102/1470-7330.2011.9020 Spec No A.

[21] Y. Zhao, G.H. de Bock, R. Vliegenthart, R.J. van Klaveren, Y. Wang, L. Bogoni, P.A. de Jong, W.P. Mali, P.M.A. van Ooijen, M. Oudkerk, Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume, Eur. Radiol. 22 (2012) 2076–2084, https://doi.org/10.1007/s00330-012-2437-y.

[22] J. Benzakoun, S. Bommart, J. Coste, G. Chassagnon, M. Lederlin, S. Boussouar, M.-P. Revel, Computer-aided diagnosis (CAD) of subsolid nodules: Evaluation of a commercial CAD system, Eur. J. Radiol. 85 (2016) 1728–1734, https://doi.org/10.1016/j.ejrad.2016.07.011.

[23] M. Silva, C.M. Schaefer-Prokop, C. Jacobs, G. Capretti, F. Ciompi, B. van Ginneken, U. Pastorino, N. Sverzellati, Detection of Subsolid Nodules in Lung Cancer Screening: Complementary Sensitivity of Visual Reading and Computer-Aided Diagnosis, Invest. Radiol. 53 (2018) 441–449, https://doi.org/10.1097/RLI.0000000000000464.

[24] H. MacMahon, D.P. Naidich, J.M. Goo, K.S. Lee, A.N.C. Leung, J.R. Mayo, A.C. Mehta, Y. Ohno, C.A. Powell, M. Prokop, G.D. Rubin, C.M. Schaefer-Prokop, W.D. Travis, P.E. Van Schil, A.A. Bankier, Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017, Radiology 284 (2017) 228–243, https://doi.org/10.1148/radiol.2017161659.

[25] L. Vassallo, A. Traverso, M. Agnello, C. Bracco, D. Campanella, G. Chiara, M.E. Fantacci, E. Lopez Torres, A. Manca, M. Saletta, V. Giannini, S. Mazzetti, M. Stasi, P. Cerello, D. Regge, A cloud-based computer-aided detection system improves identification of lung nodules on computed tomography scans of patients with extra-thoracic malignancies, Eur. Radiol. 29 (2019) 144–152, https://doi.org/10.1007/s00330-018-5528-6.

[26] D.M. Xu, H. Gietema, H. de Koning, R. Vernhout, K. Nackaerts, M. Prokop, C. Weenink, J.-W. Lammers, H. Groen, M. Oudkerk, R. van Klaveren, Nodule management protocol of the NELSON randomised lung cancer screening trial, Lung Cancer Amst. Neth. 54 (2006) 177–184, https://doi.org/10.1016/j.lungcan.2006.08.006.

[27] M.-P. Revel, A. Bissery, M. Bienvenu, L. Aycard, C. Lefort, G. Frija, Are two-dimensional CT measurements of small noncalcified pulmonary nodules reliable? Radiology. 231 (2004) 453–458, https://doi.org/10.1148/radiol.2312030167.

[28] M.-P. Revel, C. Lefort, A. Bissery, M. Bienvenu, L. Aycard, G. Chatellier, G. Frija, Pulmonary nodules: preliminary experience with three-dimensional evaluation, Radiology 231 (2004) 459–466, https://doi.org/10.1148/radiol.2312030241.

[29] M.-P. Revel, A. Merlin, S. Peyrard, R. Triki, S. Couchon, G. Chatellier, G. Frija, Software volumetric evaluation of doubling times for differentiating benign versus malignant pulmonary nodules, AJR Am. J. Roentgenol. 187 (2006) 135–142, https://doi.org/10.2214/AJR.05.1228.

[30] T.P. Coroller, V. Agrawal, E. Huynh, V. Narayan, S.W. Lee, R.H. Mak, H.J.W.L. Aerts, Radiomic-Based Pathological Response Prediction from Primary Tumors and Lymph Nodes in NSCLC, J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer. 12 (2017) 467–476, https://doi.org/10.1016/j.jtho.2016.11.2226.

[31] D.V. Fried, S.L. Tucker, S. Zhou, Z. Liao, O. Mawlawi, G. Ibbott, L.E. Court, Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer, Int. J. Radiat. Oncol. Biol. Phys. 90 (2014) 834–842, https://doi.org/10.1016/j.ijrobp.2014.07.020.

[32] T.P. Coroller, P. Grossmann, Y. Hou, E. Rios Velazquez, R.T.H. Leijenaar, G. Hermann, P. Lambin, B. Haibe-Kains, R.H. Mak, H.J.W.L. Aerts, CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma, Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol. 114 (2015) 345–350, https://doi.org/10.1016/j.radonc.2015.02.015.

[33] E. Rios Velazquez, C. Parmar, Y. Liu, T.P. Coroller, G. Cruz, O. Stringfield, Z. Ye, M. Makrigiorgos, F. Fennessy, R.H. Mak, R. Gillies, J. Quackenbush, H.J.W.L. Aerts, Somatic mutations drive distinct imaging phenotypes in lung Cancer, Cancer Res. 77 (2017) 3922–3930, https://doi.org/10.1158/0008-5472.CAN-17-0122.

[34] K. Robinson, H. Li, L. Lan, D. Schacht, M. Giger, Radiomics robustness assessment and classification evaluation: A two-stage method demonstrated on multivendor FFDM, Med. Phys. 46 (2019) 2145–2156, https://doi.org/10.1002/mp.13455.

[35] W. Zhao, J. Yang, B. Ni, D. Bi, Y. Sun, M. Xu, X. Zhu, C. Li, L. Jin, P. Gao, P. Wang, Y. Hua, M. Li, Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning, Cancer Med. (2019), https://doi.org/10.1002/cam4.2233.

[36] D. Ardila, A.P. Kiraly, S. Bharadwaj, B. Choi, J.J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D.P. Naidich, S. Shetty, End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography, Nat. Med. (2019), https://doi.org/10.1038/s41591-019-0447-x.

[37] G. González, S.Y. Ash, G. Vegas-Sánchez-Ferrero, J. Onieva Onieva, F.N. Rahaghi, J.C. Ross, A. Díaz, R. San José Estépar, G.R. Washko, COPDGene and ECLIPSE Investigators, Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography, Am. J. Respir. Crit. Care Med. 197 (2018) 193–203, https://doi.org/10.1164/rccm.201705-0860OC.

[38] S.L.F. Walsh, L. Calandriello, N. Sverzellati, A.U. Wells, D.M. Hansell, UIP Observer Consort, Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT, Thorax 71 (2016) 45–51, https://doi.org/10.1136/thoraxjnl-2015-207252.

[39] S.L.F. Walsh, L. Calandriello, M. Silva, N. Sverzellati, Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study, Lancet Respir. Med. 6 (2018) 837–845, https://doi.org/10.1016/S2213-2600(18)30286-8.

[40] G.B. Kim, K.-H. Jung, Y. Lee, H.-J. Kim, N. Kim, S. Jun, J.B. Seo, D.A. Lynch, Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease, J. Digit. Imaging 31 (2018) 415–424, https://doi.org/10.1007/s10278-017-0028-9.

[41] J. Jacob, B.J. Bartholmai, S. Rajagopalan, M. Kokosi, A. Nair, R. Karwoski, S.M. Raghunath, S.L.F. Walsh, A.U. Wells, D.M. Hansell, Automated quantitative computed tomography versus visual computed tomography scoring in idiopathic pulmonary fibrosis: validation against pulmonary function, J. Thorac. Imaging 31 (2016) 304–311, https://doi.org/10.1097/RTI.0000000000000220.

[42] J. Jacob, B.J. Bartholmai, S. Rajagopalan, M. Kokosi, R. Egashira, A.L. Brun, A. Nair, S.L.F. Walsh, R. Karwoski, A.U. Wells, Serial automated quantitative CT analysis in idiopathic pulmonary fibrosis: functional correlations and comparison with changes in visual CT scores, Eur. Radiol. 28 (2018) 1318–1327, https://doi.org/10.1007/s00330-017-5053-z.

[43] M. Vakalopoulou, G. Chassagnon, N. Bus, R. Marini, E.I. Zacharaki, M.-P. Revel, N. Paragios, Atlasnet: multi-atlas non-linear deep networks for medical image segmentation, Int. Conf. Med. Image Comput. Comput.-Assist. Interv. Springer, 2018, pp. 658–666.