ORIGINAL PAPER

Deep learning applications in pulmonary medical imaging: recent updates and insights on COVID-19

Hanan Farhat[1] · George E. Sakr[1] · Rima Kilany[1]

Abstract

Shortly after deep learning algorithms were applied to Image Analysis, and more importantly to medical imaging, their applications increased significantly to become a trend. Likewise, deep learning applications (DL) on pulmonary medical images emerged to achieve remarkable advances leading to promising clinical trials. Yet, coronavirus can be the real trigger to open the route for fast integration of DL in hospitals and medical centers. This paper reviews the development of deep learning applications in medical image analysis targeting pulmonary imaging and giving insights of contributions to COVID-19. It covers more than 160 contributions and surveys in this field, all issued between February 2017 and May 2020 inclusively, highlighting various deep learning tasks such as classification, segmentation, and detection, as well as different pulmonary pathologies like airway diseases, lung cancer, COVID-19 and other infections. It summarizes and discusses the current state-of-the-art approaches in this research domain, highlighting the challenges, especially with COVID-19 pandemic current situation.

Keywords Coronavirus Deep Learning · Pulmonary Imaging · Medical Image Analysis · Convolutional Neural Networks

1 Introduction

In 1995, Lo et al. [1] were the first to apply Convolutional Neural Networks to medical imaging, and just after, their applications widely became a research interest especially due to advances in GPU and availability of new large publicly available datasets and algorithms. An important survey by Litjens et al. [2] was published in 2017, summarizing approaches of deep learning in the medical imaging field. Specifically, deep learning applied to Lung images were the subject of 34 papers, where 35% of these targeted lung cancer diseased patients, and 63% used CT image modality to conform their tasks. This number of contributions to lungs was in the middle level between contributions to pathology, which were emerging highly, and contributions to bones and retinas

✉ Hanan Farhat
  hanan.farhat@net.usj.edu.lb

  George E. Sakr
  georges.sakr@usj.edu.lb

  Rima Kilany
  rima.kilany@usj.edu.lb

1 Saint Joseph University of Beirut, Mar Roukos, Beirut, Lebanon

which was had the fewest contributions. This directed us to a possible emerging of deep learning applications to Lungs very soon, especially that lung cancer was the highest cancer type leading to death in 2018 as 2 million cases were recorded according to WHO [3]. After a short time, and specifically in December 31st of 2019, a novel coronavirus (COVID-19) was eventually identified [3], and research directions eventually deviated toward this pandemic. Lung imaging became a faster key to the solution, and deep learning became the prosperous research field to invest in.

Since then, many surveys targeted pulmonary diseases detection/diagnosis, deep learning-based applications on pulmonary targets, or applications on certain pulmonary image modalities. This survey covers deep learning in pulmonary medical imaging including most approaches on all deep learning tasks and medical image modalities. It summarizes around 160 contributions in deep learning on lung medical image analysis, analyzing the research directions prior to and after February 1, 2017, including response to COVID-19 pandemic.

Before 2017, and referring to the same survey [2], CT scans came third in usage of all deep learning contributions to organs medical imaging (19.3%) after Microscopy (21%) and MRI (27%). Thus, dominance of CT applications over CXRs

was obvious in pulmonary medical imaging, but applications of deep learning-based approaches was promising either way. The analysis of this paper aims to include medical imaging directions and favors due to their importance in determining the deep learning algorithms results.

Selection of papers was done at three steps: once in middle of 2019 and an update in December 2019. Search was based on the terms "deep learning", "medical imaging", "chest radiographs", "chest CT", "pulmonary nodules", and "convolutional neural networks". It was performed on Google Scholar, PubMed, ArXiv and among most of the proceedings of MICCAI and IPMI conferences. Unrelated papers were excluded, such as ones targeting medical imaging of other organs, or ones targeting pulmonary diseases but not deep learning based. Chosen papers were also used to add relevant references. The first search resulted with 89 studies after excluding irrelevant papers. In December, 36 papers were added. After then, another heavy update was done in May 2020 due to the COVID-19 pandemic, adding 37 papers. This third search was done through Google Scholar website, using combinations of "COVID-19", "deep learning", "medical imaging", "CT", and "X-rays". It resulted with around 1984 search results, minimized to 140 upon first scan, then to 37 after second scan to include papers targeting only deep learning applications to medical imaging for COVID-19 diagnosis and excluding pre-prints.

The collected papers show how deep learning techniques are used to perform specific tasks on many type of diseases using different image modalities. Targeted **diseases** approximate percentages were as follows: 61% for lung cancer, 20% for image anatomy and quality, around 27% for infections, airway diseases and general thoracic diseases altogether. The rest targeted pulmonary embolism (PE), pneumothorax, pulmonary edema and interstitial lung diseases (ILD). At the level of **image modality**, around 46% of the contributions use chest computed tomography (CT), 38.5% use X-rays, and around 14% use both image modalities while 1.6% go for PET and MRI usage. Finally, from a **task** perspective and referring to Litjens et al. [2], previously, 41.2% of papers handled detection, 35.5% handled classification, and 6% handled image retrieval. Image enhancement, feature extraction, and segmentation are each handled by 2.9% of the 34 papers listed in his paper, while 8.8% target other tasks. Therefore, detection and classification were competitively researched in the domain of deep learning tasks applied on chest, with scope expansion to enhance the input features, segmented organs and patches, and scans as a whole. However, at the level of tasks performed by deep learning in the later 3 years, the ranking came approximately as follows: classification came first (33%) sharply followed by detection (31%), segmentation (23%), image enhancement (7%), feature extraction (4.7%) and finally registration (1.6%). When emphasizing on COVID-19, approximately classification and detection had

similar shares (42% of contributions for each) with segmentation having the rest. The percentage is expected to vary as soon as the many contributions that are still in process of publications get published.

The coronavirus epidemic can be considered the trigger to move the maths of deep learning fast into the clinics of medical doctors. Even though not yet implemented, it became the concern of researchers around the world, and the subject of their new experiments. Commercial applications took the chance to market their products, and artificial intelligence became the hand to support in any future pandemics. On medical imaging analysis level, it was the right time to put the present architectures and possible future improvements in the service of humans health sector. A lot of challenges still exist, starting from legal and ethical concerns, reaching the radiologists trust-gaining journey. And technically, big multi-centered datasets that are well annotated are needed along with technological resources to train the algorithms and come up with the best AI-COVID-19 assistant, that could possibly be the any-virus-assistant in the soon future.

The rest of this survey is organized as follows:

- Section 2: gives an overview of medical image modalities, deep learning and surveys on deep learning in medical imaging, in addition to available datasets for pulmonary medical images.
- Section 3: summarizes surveys on deep learning-based applications and approaches on pulmonary medical images.
- Section 4: defines COVID-19, describes related medical imaging concerns, summarizes reviews on deep learning applied to COVID-19 medical imaging analysis, and finally listing and describing contributions to this domain.
- Section 5: discusses the challenges in this explicated research domain and points out to its future directions.

2 Background overview

This section gives an overview of different medical image modalities. In addition, it highlights deep learning and provides surveys on its application in medical imaging. Finally, it provides available datasets for pulmonary medical images.

2.1 Medical image modalities

Medical imaging term refers to techniques used to reveal the internal organs or tissues of the body in order to diagnose or detect diseases presence and evolution. Many modalities of digital medical images exist, such as CT, magnetic resonance imaging (MRI), X-Ray, ultrasound (US), and positron emission tomography scans (PET). Some modalities are

organ-specific like retinal photography, and others examine multiple organs such as CT and MRI [4].

Goel et al. [5] defined common medical imaging modalities and compared some with respect to availability, cost, radiation effect, data acquisition, speed and resolution. X-rays are 2D images produced by electromagnetic waves penetrating the body and absorbed non-uniformly by tissues and bones. Similarly, CT scans or CAT scans, use electromagnetic waves to create detailed cross-sectional images, resulting in multiple two-dimensional images that can be used to create 3D representation of the target organ. CT scans are of better resolution than X-Rays, but are more expensive. In addition to electromagnetic waves, MRI uses radio waves to produce three-dimensional images that are of better resolution and has no ionizing effect in contrary to X-Rays and CT scans. However, they are last at the level of availability, while X-Ray, US and CT are widely available. US imaging, also named sonography, uses high-frequency sound waves that have no ionizing effect, with affordable fees. Another imaging modality is PET scans that is commonly used beside CT scans and requires injection of the patient with radiopharmaceuticals that define this imaging modality as nuclear.

Medical image modalities differ also in their best-use cases. PET is important for tissues and organs functionality monitoring and diagnosis [5], while for lung nodule detections, CT scans are the most sensitive modality [6] as they compete by their rapid acquisition, availability and cost effectiveness. On the other hand, Candemir and Antani [7] states that chest radiographs (aka CXRs) as the most conventional imaging modality for diagnosis of pulmonary disorders and cardio-thoracic. Candemir and Antani [7] add that X-rays are efficient for tuberculosis (TB) too, are widely available, emit radiations less than other modalities and are affordable for under-resourced regions of the world where infectious diseases spread quickly.

The next part provides surveys on applications of deep learning in medical images analysis in general.

2.2 Surveys on deep learning in medical images analysis

Deep learning (DL) is a genre of machine learning (ML) and is mainly an extension of previous artificial neural network (ANN) forms. DL is based on computational models that learn features from raw data at many levels of conciseness [8], bypassing manual feature extraction [9], imitating the human neurons structure and resulting with a model of high computational complexity [10]. In comparison with traditional ML techniques, Sahiner et al. [11] state two reasons behind the standing out of DL: the depth of the model and its composition

Deep learning targets a wide variety of tasks including classification, regression, clustering, image reconstruction,

artifact reduction, lesion detection, segmentation and others [10,11]. In order to perform these tasks, many deep learning paradigms were developed: convolutional neural networks (CNN), recurrent neural networks (RNN), reinforcement learning, general adversarial networks (GAN), auto-encoders(AE), and many others [12]. These paradigms are categorized into three: supervised learning that usually seeks a specified neural network output, unsupervised learning that involves inferring unlabeled datasets [12], and reinforcement learning which is a trade-off between exploitation and exploration [13], Which is based on the action-reward principle where the algorithm tries different actions and based on the rewards; it adjusts itself.

The most popular and commonly used deep learning paradigm is the convolutional neural network (CNN). The CNN is built up from several layers that function differently but complementary. The building blocks layers are: convolutional, activation, pooling, dropout regularization, and batch normalization layers [9]. A simple demonstration of CNN architecture is in Fig. 1. The total number of layers implicates the large number of design decisions such as the kernel size, the activation layer type, regularization level and type, loss function type, etc. [11].

Deep learning application areas have been expanding, reaching agriculture [14], mobile and wireless networking [15], Internet of Things [16], bio-informatics [17], health management systems [18] and many fields. Nevertheless, deep learning in medical imaging became widespread in 2012, and work on it developed quickly after then [19].

The availability, reliability and affordability of computer-assisted diagnosis for early cancer diagnosis can lessen the inequalities between populations at the level of mortality and save more lives according to Liu et al. [20]. And this definitely can be generalized to involve all fatal diseases. The most outstanding deep learning architecture for imaging is the convolutional neural network (CNN) [11] (Fig. 1), and, many models were developed later on based on it such as AlexNet [21], VGG [22], GANs [19], GoogLeNet [23], ResNet and others [9].

Many studies surveyed and reviewed deep learning on medical imaging from different perspectives. Contributions to this field can be chronologically summarized as done next.

The most popular paper in 2016 is that of Shin et al. [24][1] which exploited the deep convolutional neural networks for computer-aided detection at three levels; firstly different CNN architectures were compared, and secondly the impact of datasets characteristics was evaluated to finally verify when and where transfer learning is useful. Transfer learning is a deep learning technique that allows the use of networks on a dataset having been trained on another one. Thoraco-abdominal lymph node (LN) detection and ILD

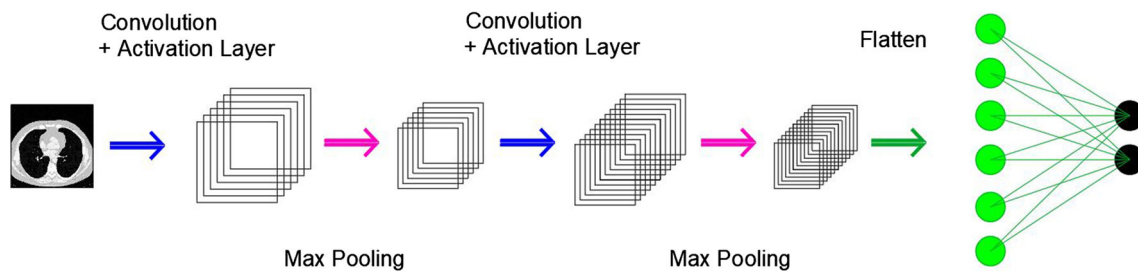[1] Cited 1672 times at the time of writing.

Fig. 1 Visualization of convolutional neural network architecture

classification were targeted in its experiments to conclude that 8-layered and 22-layered deep CNN architectures are useful when the training data is limited. However, optimal solutions for computer-aided detection problems should take into consideration the trade-off between using better learning models and using more training data. For instance, developing well annotated datasets is as necessary as developing new learning algorithms. A substitute for newer datasets is transfer learning from available natural-images datasets or exploring the handcrafted features complementary properties.

In the same year, Wang [25] published his *perspective on deep learning* which he started by emphasizing on medical imaging and explaining why using deep learning for image reconstruction is as evident for image analysis, expressing his enthusiasm that work on deep imaging will accelerate the reinvention of the future of health care and is not a bypassing wave of research.

A task-specific survey on deep learning done by Miranda et al. [26] reviews classifications techniques of medical image analysis, including but not restricted to convolutional neural networks, which aim to achieve high accuracy and identify which parts of the body are infected. In addition, the survey covered image modalities used, datasets and trade-offs for each technique, and the improvements allowing accuracy and sensitivity enhancement. The challenges were listed as follows: the continuous increase in diversity and number of images, mathematical formulations, and computing power. The survey ended up by an expectation of employing image classification techniques in computer-aided diagnosis.

Reaching out to 2017, the interest in this domain notably increased. *A survey on segmentation techniques for medical image processing* by Merjulah and Chandra [27] summarized the efficient methods, compared them and concluded that CNNs achieved the highest accuracy outperforming non-deep techniques. In addition, the survey noted that CNNs have the potential to perform detection and segmentation while applying classification, yet its success is dependent on the given problem and the suitable corresponding architecture.

In the same year, Erickson et al. [28] gave a valuable introduction about machine learning for medical imaging and its types but emphasized on supervised learning. Their paper included important definitions for used terminology, explained the stages of Machine Learning process, defined supervised machine learning types especially CNNs, listed open-source tools and libraries, and commented on them. The drawn-out conclusion is the necessity of understanding the learning process to avoid misusing it, and the benefits of CNNs over traditional ML methods are the no-need to compute features manually.

Shen et al. [29] published a paper in which they introduced the fundamentals of deep learning methodologies and evaluated their performance in many application areas such as computer-aided diagnosis and prognosis, tissue segmentation and others. The authors summarized the challenges defeating approaches by the need to use smaller patches as input, augment training data for better learning process, and use different forms of transfer learning. In addition, they concluded that PET data could be estimated given MRI data. Current research directions were discussed agreeing that deep learning advances are due to the development of GPU, available datasets and algorithms. As recommendations, it remains a challenge to interpret the learned model, and the development of algorithms should take into consideration the need to generalize over different imaging protocols, along with the need for architectures dependent on domain-specific information.

Ker et al. [4]'s survey was first released in December 2017 and updated in February 2018. It covered all previous work referencing most important books and reviews in deep learning for medical analysis and 200 most cited papers in the last 3 years. According to their survey and agreeing with Shen et al. [29], the majority of published algorithms employ CNN and the advances in this domain are due to GPU advances and the availability of larger datasets. However, they addressed two more challenges in training data: the case of data imbalance and the need to estimate how much labeled data is needed. Concerning the future expectations, this survey added the use of radio images to predict underlying molecular origins of tissues, combining content-based image retrieval to computer-aided diagnosis, generating better quality of MRI images, and classify lung cancer sub-types.

Among all, the most popular survey in 2017 is that of Litjens et al. [2]. It reviewed the major deep learning concepts suitable for medical image analysis and summarized more than 300 contributions to this domain. It surveyed the deep learning tasks (Segmentation, Classification, Detection, Registration, etc.) and application areas (Retinal, Breast, Cardiac, abdominal, etc.), ending with a critical discussion of the open challenges and future research directions. A drawn conclusion is that the most preferred trained CNNs are the end-to-end ones, on which transfer learning has high impact. Yet, the exact architecture, according to the survey, is not enough for a good solution. The model hyper-parameters have no rules to choose them accordingly; the input sizes should be relevant to the problem context (not to over-fit); the acquisition of relevant annotations for images should be noiseless and as fast as possible, and the features should be balanced not to accidentally exclude the clinical ones.

In 2018, many surveys were done in this field. Lundervold and Lundervold [9] aimed to introduce deep learning, describe its application on MRI processing and analysis, and provide a start-up bench for future contributors listing state-of-the-art open-source codes, datasets, educational references and possible problems. Challenges are categorized into: "Data," "Interpret-ability, trust and safety," and "Workflow integration and regularization." The authors agreed in their survey that the most used networks are standard deep neural networks and characterized them as "data-hungry" networks.

In the same year, Meyer et al. [30] emphasized on radiotherapy complex treatment facilitated by artificial intelligence. It presented the common network architectures emphasizing on CNNs and shed the light on published work in their specified field, classified into seven categories relevant to the workflow of patients. The authors hoped their work will inspire researchers to work on radiotherapy-specific applications. According to them, despite the advantages of deep neural networks, they remained empirical whether choosing the general architecture or deciding on the hyper-parameters of models. Even tricks that improve the performance lacked justifying theories. In addition, the authors assured that the need for building well-annotated datasets being as pivotal as new algorithms development. Finally, they concluded that this is just the start for radiotherapy and expected that work on it will rapidly evolve.

Moving forward to the beginning of 2019, research has expanded on medical image analysis using deep learning and that was clearly proved by the count of papers and surveys published this year. Latif et al. [31] published a review on medical imaging using machine learning and deep learning algorithms. They provided an outline for researchers including existing medical imaging techniques with their advantages and drawbacks. They also discussed multi-dimensional medical data and methods for analyzing distinctive diseases. In conclusion, disease patterns are better classified and categorized by deep learning algorithms permitting to extend their goals and predicting their performance when used for patients' treatment. Finally, researches are considering challenges and still flourishing continuously in health and other application fields.

Wang et al. [32] surveyed deep learning for image super-resolution. They targeted anomaly detection in different areas, aiming to review firstly: research deep-learning-based methods for anomaly detection, and secondly their applications in various domains along with corresponding assessments. The authors' categorized methods, drew an outline including variants and assumptions related to anomalous behavior differentiation for each method, listed their advantages and limitations, and discussed the real computational complexities when applied. Last but not least, the survey ended by a detailed discussion on each level of network design, learning strategies, evaluation metrics, unsupervised super-resolution and future directions for real-world scenarios (Dealing with images degradation, domain-specific applications, and multi-scale super-resolution).

Again in 2019, Altaf et al. [33] described the deep learning effect on medical-image-analysis as a paradigm shift. The reviewed literature is organized according to human anatomy surveying the recent developments in the targeted topic. The lack of well-annotated large-scale datasets is stated as the core challenge. However, this survey points out to the importance of collaborations between experts in medical imaging and experts in computer vision fields in order to significantly improve the application of machine learning and computer vision tasks in medical image analysis and health care.

An overview on algorithms and concepts that would enhance deep learning performance was given by Thaler and Menkovski [34]. The paper motivated the use of machine learning and pointed out to some shortcomes of it. It described, as well, the major building blocks of deep learning methods and the way to use them in solving problems. However, the main content of the paper is applications of deep learning technologies on health care problems of which the authors provide a background knowledge for. Different data modalities (not only images) are defined and structured within the paper, declaring the remarkable success of deep learning in many tasks and application areas even with the large number of challenges present, starting from the uninterpretable inner process reaching bias-free, large-scale, available and relevant training data choice.

Another survey is by Haskins et al. [12] targeting a specific medical image analysis task that is image registration. It outlines the evolution of deep learning and its limitations in this context. Besides, it states the future research directions as follows: deep adversarial image registration, reinforcement learning-based registration, and raw imaging domain registration.

A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis by Cheplygina et al. [35] traces back the origin of deep neural networks and introduces applications for different tasks and diseases summarizing the development of frameworks and various algorithm models on datasets.

As surveys and reviews on deep learning in medical image analysis are continuously released and updated, it is worth noting that research directions are continuing on organ/target specific levels. Lung was targeted in only 11.3% of studies issued before February 2017 [2]; however, the number of publications on deep learning-based applications targeting lung have increased notably as around 57% of the total papers mentioned in this paper were published in 2019. Similarly, around 48% of the publications targeted lung cancer conforming with the fact that lung cancer recorded the highest percentage of new cases in 2018 [36].

With no doubt, the CNN architecture was a great advance for medical image analysis. Until 2017, many proimising architectures proved their efficiency, starting with AlexNet in 2012 [21] reaching Inception-**V4 in 2017 [?], passing along with VGG [22], GoogLeNet [23], DenseNet [?] and ResNet [?]. By then, transfer learning has been studied extensively and it became worth noting that using natural images as training datasets was found to be useful even for medical image analysis. Moreover, surveys and approaches became more target-specific in means of organs, image modalities, or deep learning task (details in Sect. 3). And as deep networks became the promising future, the training data was a challenging barrier as public data is not enough in quantity and needs to be accompanied by extensive precise work of radiologists. However, some challenges like LUNA16 [37] and CheXpert [38] improved the chances for better trained models, but other sources of large datasets declared that their labeling of their data is not fully accurate such as [39], which recommends investing in the labeling process as well as the deep learning application. Hence, arouses the need for unsupervised approaches [35]. However, the publicly available datasets that were used by the aforementioned applications are presented next.

2.3 Datasets for pulmonary medical images

There are a plenty of medical images databases, that are either available publicly or upon conditional request. For example, The cancer imaging archive (TCIA) [40] consists of many collections that target cancer in different organs for varying imaging modalities. In addition, neuro-imaging datasets of the brain are available at The Open Access Series of Imaging Studies (OASIS) [41]. Besides, there exists datasets that target Alzheimer [42], retinas [43], knee MRIs [44], and sometimes many organs at once [45].

Emphasizing on pulmonary medical imaging, Qin et al. [46] manifests the top available CXR datasets that researchers rely on: the Indiana dataset [47], the KIT dataset [48], the MC dataset [49], the JSRT dataset [50,51], the Shenzhen dataset [49], and ChestXray14 dataset [39]. Zhang et al. [6] states that most significant and well-known databases for pulmonary nodules in CT are:

- Automatic Nodule Detection 2009 (ANODE09) [52]
- Lung Image Database Consortium and Image Database Resource Initiative (LIDCIDRI) [53]
- Lung Nodule Analysis 2016 (LUNA16) [37]

In addition, the following public datasets provide lung CT images:

- Deep Lesion [54] [55]
- COPDGene [56]

while CXRs are available through:

- CheXpert [38]
- MIMIC-CXR [2]

Some public datasets used by contributions discussed in Sect. 3.2 are defined in Table 1.

3 Deep learning and pulmonary medical imaging analysis

This section presents an overview of the literature of deep learning applications to the specific topic of medical imaging analysis. The first part of this section presents an overview of the surveys on deep learning applications to pulmonary medical image analysis, while the second part presents papers that prove the improvement brought by deep learning to medical image analysis. Papers are contributions of researchers to deep learning applications on pulmonary medical image analysis. They are clustered according to deep learning tasks: registration, image enhancement, detection, feature extraction, and classification. Contributions to each of them is categorized according to target organ, object, lesion, or disease.

3.1 Surveys on deep learning applications to pulmonary medical imaging analysis

Machine learning targets different chest parts and objectives. Examples on contributions areas are rib detection and suppression, fissure extraction, airway segmentation, and nodule

---

[2] To request credentials for access: [57].

Table 1 Public datasets used by contributions to deep learning applications in pulmonary medical imaging analysis; ImO: image modality, CXR: chest X-ray, CT: computed tomography

Public dataset	ImO	Description
LIDC [53]	CT	Lung image database consortium image collection (LIDC-IDRI); consists of 1018 cases of diagnostic and lung cancer screening thoracic CT scans with marked-up annotated lesions
DIR-Lab [58]	CT	Consists of 10 breath-hold CT image pairs that are randomly selected from the COPDgene
NLST [59]	CT	The National Lung Screening Trial; consists of 75000 CT exams images and aimed to determine whether screening for lung cancer with low-dose helical CT reduces mortality from lung cancer in high-risk individuals relative to screening with chest radiography
COPDGene [56]	CT	Chronic Obstructive Pulmonary Disease Gene; consists of 400 CT scans and aims to test why some smokers develop COPD while others do not
Computed Tomography Emphysema Database [60]	CT	Comprises 115 high-resolution CT (HRCT) slices as well as 168 square patches manually annotated in a subset of the slices targeting emphysema
Deep Lesion [54,55]	CT	Contains over 32,000 annotated lesions identified on CT images, which represent 4,400 unique patient
ANODE09 [61]	CT	Consists of 55 CT scans for evaluation of nodule detection algorithms
ChestXray14 [39]	CXR	Consists of 112120 frontal view radiographs targeting 14 thoracic diseases
JSRT [50]	CXR	Japanese Society of Radiological Technology; consists of 247 PA chest radiographs with/without lung nodules
Shenzhen and Montgomery [49]	CXR	Shenzhen; contains 662 frontal CXRs with/without manifestations of TB including pediatric X-rays (AP). Montogomrg; contains 138 frontal CXRs with/without manifestations of TB
RIH-CXR [62]	CXR	Rhode Island Hospital chest radiographs consists of 17,202 frontal view chest radiographs labeled as normal or abnormal
LOCT and chest X-rays [63]	CXR	Labeled Optical Coherenece Tomography and CXRs dataset; consists of thousands of validated OCT and CXRs images that are split into training and testing sets. OCT Images are labeled as (disease)–(randomized patient ID)–(image number by this patient) and split into four directories: CNV, DME, DRUSEN, and NORMAL
MIMIC-CXR [64]	CXR	Consists of 371,920 CXRs associated with 227,943 imaging studies labeled with 14 diseases
CheXpert [38]	CXR	Contains 224,316 chest radiographs of 65,240 patients, with 14 observations in radiology reports
TCIA [65]	Multiple	The Cancer Imaging Archive; is a service that provides "Collections" of medical images of cancer for public download. The collections are organized according to patients related by a common disease (e.g. lung cancer), image modality (MRI, CT, etc) or research focus

detection, characterization and classification [66]. Similarly, deep learning methods arouse for plenty of them.

According to [8], applications in chest imaging using Radio-graphs are mainly lung nodule detection, TB diagnosis, and multiple abnormal pattern (MAP) detection (patterns such as pneumonia, pleural effusions, etc). In addition, using chest CT, deep learning can be applied on nodule detection/screening, ILD, chronic obstructive pulmonary disease (COPD), and image normalization.

This claim complies with the fact that disregarding the rarely used image modalities for pulmonary diagnosis such as US [67] and PET, the emphasis is on CT scans and CXRs in around 130 papers on deep learning application to chest, published since 2017.

Several surveys and reviews handled the deep learning challenge, but the interpretation of what is happening in the inner processing stages stayed vague. Nevertheless, since 2017, contributions to medical targeting or only pointing to chest diagnosis were general, task-specific, image-modality-specific, target-specific, treatment-method-specific, or any combination of the aforementioned specificities.

Table 2 Surveys and reviews mentioning or targeting pulmonary diagnosis, G: general, TaS: task specific, IMS: image modality specific, TgS: target specific, TtS: treatment specific, y: checked

Title	Year	Citations	G	TaS	IMS	TgS	TtS
A survey on deep learning in medical image analysis	2017	2631	Y				
Deep learning in medical	2017	870	Y				
Machine learning for medical imaging	2017	244	Y				
Deep learning applications in medical image analysis	2017	185	Y				
Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology	2017	121	Y				
Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning	2017	58	Y				
Abnormality detection and localization in chest X-rays using deep convolutional neural networks	2017	41		Y	Y	Y	
Segmentation technique for medical image processing: a survey	2017	6		Y			
Computer-aided detection systems to improve lung cancer early diagnosis: state-of-the-art and challenges	2017	5		Y		Y	
An overview of deep learning in medical imaging focusing on MRI	2018	73			Y		
Survey on deep learning for radiotherapy	2018	58					Y
Computer-aided detection in chest radiography based on artificial intelligence: a survey	2018	35			Y		
Automatic nodule detection for lung cancer in CT images: a review	2018	9		Y	Y	Y	
Comparing deep learning models for population screening using chest radiography	2018	9			Y		
Deep learning in medical imaging and radiation therapy	2019	72					Y
Artificial intelligence in cancer imaging: clinical challenges and applications	2019	66	Y				
Deep learning in medical image registration: a survey	2019	16		Y			
Going deep in medical image analysis: concepts, methods, challenges and future directions	2019	7	Y				
Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: a systematic review	2019	5		Y		Y	
A review on lung boundary detection in chest X-rays	2019	4		Y	Y	Y	
Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET	2019	0			Y		
Evolving the pulmonary nodules diagnosis from classical approaches to deep learning aided decision support: three decades development course and future prospect	2019	0				Y	
A comprehensive review on multi-organs tumor detection based on machine learning	2020	1		Y		Y	
Image segmentation using deep learning: a survey	2020	0		Y			

Out of 24 surveys and reviews in Table 2, 10 were emphasizing on chest images and/or problems. Eight studies were published in 2019: [7,11,12,20,33,68–70]. Yet, it is expected that more studies will be issued in the next year and thus exceed the earlier rates.

In 2018, targets were chest radiography [71,72], MRI [9], radiotherapy [30], and pulmonary cancerous nodules [6]. A year earlier, lung cancer early diagnosis [73], mediastinal lymph node metastasis of non-small cell lung cancer (NSCLC) classification [74] and chest imaging in general [66] were targeted.

The rest of contributions targeted either medical images in general including chest imaging, or a deep learning task that involves application on different organs including chest such as registration, or a disease that infects lungs and others like cancer.

Wang et al. [74] compared one deep learning method (used AlexNet CNN architecture) and four classical machine learning methods for classifying mediastinal lymph node metastasis of NSCLC. Image patches from two modalities were used: PET and CT all at once, which may have limited the performance of CNN. The study favored diagnostic fea-

tures over texture features in working with lymph nodes due to their small size. The CNN's performance did not significantly vary from best methods, without even using important diagnostic features, directing research toward incorporating diagnostic features into a newly designed dual-modality PET/CT images.

Van Ginneken [66] summarized 50 years of computer analysis in chest Imaging. It handles rib detection and suppression in chest radiographs. And in CTs, it handles fissure extraction, airway segmentation, nodule detection, classification and characterization. The authors concluded that convnets perform as feature extractors and classifiers at once. They can be used for producing filtered images, and their generalization can introduce them into many applications rapidly. In addition, deep learning continuously allows the integration of text and image analysis to improve the performance.

Even though Kim et al. [75] compared deep and shallow learning methods for classifying the regional pattern of diffuse lung diseases, some limitations existed, such as the popular training-data-size dilemma. The accuracy of deep learning exceeded that of shallow in all inter-scanner variations, attempting to consider the whole lung quantification. Yet, it did not address the misclassification caused by airways, lung boundaries and vessels. Volumetric CT scans were recommended to be used in the deep learning process.

Qin et al. [71] surveyed computer-aided AI-based detection in chest radiography. They referenced important datasets and image pre-processing techniques and reviewed specific disease detection like pulmonary nodules and TB, in addition to multiple disease detection. Out of many, deep learning methods proved to be the most accurate in classification. Moreover, they can predict many suspected disease types' presence simultaneously, with limitations caused by the imbalance or insufficiency of datasets. Features' extraction is also time-consuming, which implied the need to research using other domain datasets to optimize the initial hyper-parameters decisions. Besides, multiple-disease detection was recommended to be given attention, as it is clinically vital to realize their co-presence.

Traverso et al. [73] reviewed the computer-aided detection systems that improve the early diagnosis of lung cancer. Based on their results, the best computer-aided system involves the use of CNN for false positive reduction and candidate detection. Even though the combination of deep learning and other methods appeared to perform better, the sensitivity saturated already starting from 2 false-positives per scan.

Sivaramakrishnan et al. [72] compared deep learning models for population screening using frontal chest radiography. The results demonstrated that pre-trained CNNs are promising for feature extraction in medical images, especially for TB. Moreover, it emphasized the need for large datasets to enhance the performance and increase the accuracy. The comparison between pre-trained and customized deep learning models favored pre-trained ones, as favored features from shallow layers over those from deep layers of the pre-trained CNNs.

Zhang et al. [6] reviewed automatic nodule detection for lung cancer in CT scans. They mentioned the techniques better used for lung nodule detection, yet pointed general and specific challenges for this task, as nodules vary in type, size, texture, location and their respective clinical records. The review detailed data acquisition, pre-processing, lung segmentation, nodule detection and false-positive reduction. High sensitiveness was achieved by several works, however, at high false-positive rates. The use of LIDC-IDRI dataset was frequent among the papers reviewed, and the traditional methods results were satisfying, but AI-based methods (such as deep learning) have shown better performance and set the expectations higher. The major advantage of CNN, according to them, was declared to be its ability to learn from different sources of data, and determining by itself the unknown features required for the learning process. CNNs underperformed SVM classifiers, but were still promising to breakthrough. The future challenges mentioned were many, such as the need to consider the different types, sizes, locations, and textures of pulmonary nodules. Besides, building a set of features to reduce the false-positives' rates was pointed as a challenge. The authors agreed with others that the need for large public annotated datasets is vital for training the models, and the cooperation between academic institutions and medical organizations will help optimize the efforts to achieve better results.

Labaki and Han [76] questioned if deep learning will make chest imaging smarter. They noted the impact of implementing approaches on larger scale. In addition, it was impossible to examine all slices of each CT scan; thus, images composed of four cuts were used instead. Larger number of scans improved performance of the model, and thus, more is needed for training, taking into consideration the effect of varying imaging protocols. Furthermore, clinical outcomes accompany potential imaging features such as airway disease. That is why engaging clinical data into the predictive process was recommended, as well as prioritizing the compatibility of models with standard workflows.

Benzaquen et al. [77] discussed lung Cancer Screening (LCS) and suggested three methods to improve it. First, selection criteria should be refined (risk factor assessment). Second, computer-aided diagnosis should be used to interpret chest CTs. And finally, biological blood signatures should be used for early diagnosis of cancer. The second method is our concern, where concisely deep learning was applied for imaging interpretation, but the involvement of three methods is recommended to optimize the performance, especially

that CNNs are still black-box-like where their inner world is unrevealed yet.

Liu et al. [20] summarizes three decades of pulmonary nodules' diagnosis and concludes with a future prospect. It starts with very early approaches from 1980s and ends with deep learning various methods like two-/three-dimensional, multi-view, multi-scale, multi-stream, multi-tasking deep convolutional neural networks, deep belief networks (DBN), auto-encoder (AE) networks, and ensemble methods. According to the paper, tremendous challenges still exist even though the improvements in field of pulmonary nodules diagnosis are noticeable. The challenges were identified on data scarcity, diagnostic accuracy and training efficacy, which are multi-faceted barriers that need more than a single solution.

Pehrson et al. [70] also reviewed pulmonary nodule detection. However, they target deep learning exactly rather than computer-aided methods in general. Specifically, they focus on papers using lung image database consortium image collection (LIDC-IDRI) for training and testing models concerned of detecting lung nodules in thoracic CT scans. The majority of feature-based algorithms included in the review achieved more than 90% accuracy, while deep learning methods achieved accuracies in the range 82.2–97.6%. The authors concluded that even high accuracies do not prove the preference of machine learning method applied over others, especially that different hyper-parameters and heterogeneous selected scans are usually used. A limitation is also the lack of labeled training data, but LIDC-IDRI is considered as a step forward, which spots the light on significance of the images' relevant annotations' acquisition rather than their availability. Feature-based ML algorithms perform better than DL; yet, the importance of DL algorithms is amplified when features are not identified as they are able to identify features themselves. Advances in GPUs initially created for massive gaming have created an opportunity to benefit from the computing capabilities in a more sincere manner: DL experimenting. Many algorithms introduced pre-processing techniques, like transfer learning and defining bounding boxes prior to prediction. Finally, some contributions were picked by the authors to rely on for future work.

Gooen et al. [78] compared deep learning approaches in pneumothorax detection and localization using CXRs. Three methods are compared: CNN, FCN and MIL. CNN has achieved best performance in terms of area under curve (AUC) while fully convolutional networks(FCN) and multiple-instance learning(MIL) outperformed it in terms of localization confidence. Authors recommend elaborating more techniques, possibly combining the three approaches by either merging their architectures or cascading them.

Candemir and Antani [7] handle the lung boundary detection approaches using CXRs issued between 2006 and 2017, using both frontal X-rays and lateral ones. They highlight radio-graphic measures that were extracted from lung boundaries, and their uses in cardiopulmonary abnormality detection, concluding the challenges facing researches. The review references publicly available CXRs datasets and classifies deep learning-based methods as best-performing over other methods, even if time-consuming and computationally costly. Moreover, the authors realize the fact that all research on this topic target adults while pediatric CXRs are usually more noisy and challenging but disregarded. Thus, datasets and studies should be developed and performed on pediatrics.

As a conclusion, reviews and surveys have agreed on the promising future of deep learning methods, especially CNNs, as well as on the tremendous challenges striking them. The next part details approaches since 2017 and classifies them according to their tasks: Registration, Image Enhancement, Segmentation, Detection, Feature Extraction, and Classification highlighting pulmonary diseases.

3.2 Contributions to deep learning applications in pulmonary medical imaging analysis

The machine learning methodology of analyzing a medical image involves many tasks, that most of them may be applied using deep learning. Yet, researchers do not mind imposing non-deep-learning methods into the analysis process to improve the performance of deep learning proposed approaches. Referring to the work of Latif et al. [31], the ML workflow starts by feeding the medical image into the algorithm, segmenting, extracting features, selecting features and discarding noises, classifying, and finally detecting the targets and deciding on the diagnosis. According to the same reference, deep learning algorithms can categorize, classify and enumerate patterns of diseases from images upon processing. This allows raising the expectations of predictions based on the image processing output. Figure 2 visualizes two deep learning tasks applications on chest: Bone suppression and segmentation, in addition to two targeted chest diseases: pneumonia and TB.

Target pulmonary applications are mainly divided into many parts: general thoracic diseases, lung cancer, ILD, infections, pulmonary edema, PE, airways diseases, and pneumothorax.

General thoracic diseases include tasks that target multiple pathologies, image quality, lungs anatomy, and diseases occurrences.

Lung cancer here refers to tumors and nodules. Lung nodules are frequently detected on chest imaging performed to screen for lung cancer or metastasis from other malignancies or to evaluate respiratory symptoms. The risk of lung cancer in these nodules depends on its size, morphology, evolution over time and patient risk factors.

Diffuse parenchymal lung diseases are characterized by bilateral and multilobar involvement of the lungs. They are
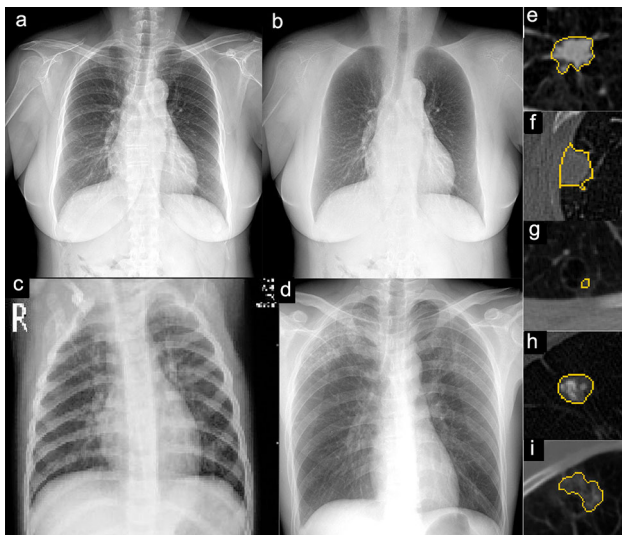
Fig. 2 Visualization of some deep learning tasks applications on chest and some target chest diseases. Bone suppression as application of image enhancement [79]: **a** standard chest radiograph, **b** soft tissue image. Example on pneumonia [80]: **c** sample chest image. Example on TB [81]: **d** posteroanterior chest radiograph showing upper lobe opacities with pathologic analysis–proven active TB. Segmentation [82]: **e–h** lung nodules segmented using U-Net

frequently caused by ILD and occasionally by infections or systemic diseases.

Infections covered are two types: TB and pneumonia. Chest radiology is frequently ordered to diagnose pneumonia which is an infection of the lung parenchyma that can be caused by bacteria, viruses or fungal infection, while TB is an infection by mycobacterium TB that is endemic in certain populations and have characteristic findings on chest radiology.

Pulmonary edema refers to the accumulation of fluid in the lungs and is frequently caused by heart failure (cardiogenic), other causes include inflammation (non-cardiogenic). Pulmonary edema has features on CXR and CT scan of bilateral alveolar filling in addition to vascular engorgement and pleural effusions.

PE is a disease whereby a clot or thrombus occlude one or multiple branches in the pulmonary arteries. It is usually detected by CT scan with contrast administration to detect a filling defect in the pulmonary vasculature.

Airway anatomy and segmentation is important to localize lesions in the lung and guide procedures such as bronchoscopy and lung biopsy. In addition, the airways are affected by multiple diseases including asthma, COPD, bronchiectasis and cystic fibrosis.

Finally, pneumothorax is an accumulation of air between the visceral and parietal pleura covering the lungs, diagnosed by CXR or CT scan. Recently, US is used for diagnosis.

Definitions of tasks and approaches to each with corresponding target disease are detailed in the next parts.

Table 3 Contributions to registration in pulmonary medical imaging analysis

Authors	Year	Citations	Modality
de Vos et al. [83]	2019	55	CXR
Hering et al. [84]	2019	0	CT

3.2.1 Registration

Images registration transforms different image datasets into one system that has matched imaging content [12]. Previously, it was manually done because it requires clinical expertise. However, deep learning has changed the landscape of image registration research.

An approach to apply deep learning on medical image registration was proposed by de Vos et al. [83] where stacked layers of trained ConvNets are used to exploit the image similarity analogous to conventional intensity-based image registration and thus allow the architecture to predict the registration of unseen images. The approach was comparable to conventional image registration methods but faster by several orders of magnitude. Another registration contribution is by Hering et al. [84] which proposed a using the whole image rather than patches, depending mainly on two blocks: a convolutional neural network in addition to the loss function (U-net architecture), and the embedding into a multilevel approach from coarse to fine. The approach allows predicting a 3D deformation field. Both contributions are classified under image anatomy and quality, one targeting X-rays and another targeting CT scans (Table 3).

3.2.2 Image enhancement

Image enhancement is a pre-processing technique. It improves the visual representation of the image and thus enhances its analysis [85]. It could be in the form of denoising, bone suppression, tissue-bone separation, reduction of spatial resolution loss, or the reconstruction of the image itself. The contributions are summed up in Table 4.

3.2.2.1. Denoising

Dealing directly with denoising, Umehara et al. [86] defined image super-resolution as producing a high-resolution image from low resolution one. They compared the image quality of the Super-resolution convolutional neural network (SRCNN) and conventional image interpolation methods, nearest neighbor, bi-linear and bi-cubic interpolations. SRCNN scheme significantly outperformed conventional interpolation algorithms for enhancing image resolution at the quantitative level (peak signal-to-noise ratio (PSNR) and structural similarity (SSIM)) and visual level (sharper edges with no obvious artifacts) yielding to the improvement of image quality of magnified chest radiographs.

Table 4 Contributions to image enhancement in pulmonary medical imaging analysis

Authors	Year	Citations	Image modality	Target
Lee et al. [94]	2019	1	CXR	Dual-energy CXR generation with single X-ray exposure
Lee et al. [93]	2019	2	CT	Image resolution enhancement
Lee et al. [91]	2019	9	CT	Spatial resolution loss reduction
Umehara et al. [92]	2018	19	CT	Image conversion among different reconstruction kernels
Umehara et al. [86]	2017	9	CXR	Denoising
Vidya et al. [89]	2019	0	CXR	Denoising
Ahn et al. [88]	2019	0	CT	Denoising ultra-low-dose CT scans
Tang et al. [87]	2019	0	MRI	Denoising
Yang et al. [79]	2017	51	CXR	Bone suppression
Zarshenas et al. [90]	2019	2	PET	Bone and soft-tissue suppression

Secondly, Tang et al. [87] targeted denoising also, but for PET scans using artificial neural networks. A three-layer ANN architecture was adopted, with 128 hidden layers. Datasets were customized; 1 for training and 9 for evaluation. The model proved its efficiency in noise reduction of low-count (1/10-count) chest PET images as it recorded an average of 40% decrease at the level of 1/10-count images. Besides, 40% signal-to-noise ratio increase was recorded for ANN processed images for all patients. The model was noted to be promising as it is not time-consuming nor computationally costly.

Moreover, Ahn et al. [88] presented a denoising deep learning approach for ultra-low-dose chest CT. A modified U-net model was used, with kernel size 4x4 and five layers. It was trained by anonymized regular-dose chest CT scans, which were also used to produce low-dose CT scans and low-dose noise by a simulator. The model then predicted the denoised image by subtracting the predicted noise image from the ultra-low-dose CT image. The bronchial wall, lung fissure, and soft tissue were measured visually. Besides, the standard deviation of soft tissue was calculated.

Noise can be caused by different scanners types and manufacturers. Vidya et al. [89] aimed to improve the learning process applied to medical images, specifically chest radiographs, reducing these effects. Global normalization and a local enhancement filter (for finer structures and opacities) are applied on three public and one private data sources. The model used for experimenting is DenseNet and recorded 0.043 mean enhancement (increase) with standard deviation decrease by 0.013. This proved the efficiency of the proposed transformations.

3.2.2.2. Bone suppression

Bone suppression and separating it from soft tissues are also denoising from a clinical point of view in some cases. For such reason, Zarshenas et al. [90] aimed to develop a model that separates ribs and clavicles from soft tissue to better visualize chest radiographs. The proposed method included CNN of two scopes: anatomy-specific and orientation-specific. The anatomy-specific CNN was designed previously but redefined by the authors, to separate bones from soft tissues. Besides, they presented different orientation-specific CNNs—with the corresponding—frequencies that were trained. Several additions to the end of the process resulted in a higher similarity in comparison with gold-standard bone separation techniques.

Another approach was held by Yang et al. [79] who used convolutional networks for bone suppression to refine the predicted bone gradients progressively. The architecture based on many cascaded convolutional networks worked to predict bone gradients at varying resolutions and scales. Finally, the gradients were fused to produce an estimation to be subtracted from the original CXR where bone components were to be eliminated.

3.2.2.3. Image reconstruction and others

Reconstruction of images from sparsely sampled CT scans was also a concern at the level of image quality.

Spatial resolution loss of predicted images was aimed to be reduced by Lee et al. [91], who proposed fully convolutional networks replacing pooling layers with wavelet transform to predict high-quality images. The hybrid reconstruction technique reduced the blurring effect of deep learning and the streak artifacts resulting from sparse sampling conditions. This approach was applied for sparsely sampled CT scans such that images were restored with quality similar to that of fully sampled images.

Umehara et al. [86] continued their work after training SRCNN on JSRT to train on CT scans from The Cancer Imaging Archive in 2018 [92]. The results showed a highly restored reconstructed image comparable to the reference image and magnified twice. Thus, they suggested that SRCNN may become a potential solution for generating high-resolution CT images from standard CT images.

Lee et al. [93] aimed to develop and validate a CNN that converts CT images reconstructed with one kernel to images

with different reconstruction kernels, and it showed adequate performance with high accuracy and speed, indicating its usefulness for clinical application.

Reconstruction of images can be also for preventive reasons, like reducing the need for extra exposure of patients to radiation. This was the motivation of Lee et al. [94], besides improving the diagnostic accuracy. The objective was achieved by developing a methodology to synthesize dual-energy chest radiographs from given single-energy chest ones. The proposed method was a modified U-net, in addition to an anti-correlated relationship (ACR) of dual-energy chest radiographs. The model was trained, tested and evaluated by calculating the modulation transfer function and coefficient of variation. The structural similarity approach (SSIM) comparing predicted and correct DECR was over 0.85 which is one of the highest found in the literature. Moreover, the produced images' quality measured better than that of U-net.

The listed contributions are clear evidence of image enhancement role in improving the accuracy of deep learning processes by refining the input images, reconstructing them, and possibly synthesizing new images from them.

3.2.3 Segmentation

Segmentation, which is a pre-processing technique, aims to extract Regions of Interest (ROI) from medical images aiming to optimize the image analysis process. It is actually the process of dividing images into meaningful parts, which refer in the case of medical images to organs, tissues or other biological structures [27]. The division process locates exact boundaries of targeted objects, but cover the entire image when collected. Contributions to this task are given in Table 5. It can be of two types: organ segmentation or lesion segmentation. Categorization of contributions is under these two types and targets which are multi-organ, cardio-thoracic, lungs and bones, lung parenchyma, pulmonary nodules, lung cancer tissue, lung tumors, pulmonary vessels, and airways diseases.

3.2.3.2. Organ or substructure segmentation

In contributions to segmentation in pulmonary medical imaging analysis, authors targeted sometimes organs in general, including lungs or parts of it.

Multi-organ, cardio-thoracic

Zhang et al. [95] proposed dense image-to-image (DI2I) network for multi-organ segmentation, that was trained on digitally reconstructed radiographs rendered from CT volumes, followed by a task-driven GAN that consists of a modified cycle-GAN substructure for pixel-to-pixel translation between DRRs and X-ray images in addition to a module leveraging the pre-trained DI2I for consistency. The

TD-GAN aimed to achieve style transfer from unseen real X-ray images.

Two approaches emphasized on CXR and segmented organs from chest area. Dai et al. [96] proposed SCAN framework that consists of a segmentation network that plays the role of generator in GAN, and a critic network. The critic takes either the ground truth mask or the predicted mask and outputs the probability estimate whether the input is the ground truth or predicted mask.

Moreover, Dong et al. [97] proposed an unsupervised adaptation framework on adversarial networks, that learns domain-invariant feature representations from available open sources and produces accurate chest organ segmentation for unlabeled datasets. A discriminator is added to distinguish segmentation predictions from ground truth masks.

Lungs and bones

In fact, Gordienko et al. was more precise segmenting only lungs and bones from CXRs. His two papers: Gordienko et al. [98] and Gordienko et al.[99], studied the impact of pre-processing techniques on dimensionality reduction and performance. Gordienko et al. [99]'s comparison showed that bone shadow exclusion demonstrates the best accuracy and loss results in comparison to other pre-processed datasets after lung segmentation. However, Gordienko et al. [98] compared an original dataset to datasets of different combinations of lung segmentation, bone shadow exclusion and outliers filtering. The pre-processed dataset obtained after lung segmentation, bone shadow exclusion, and filtering out the outliers by t-SNE demonstrated the highest training rate and best accuracy in comparison to the other pre-processed datasets. This emphasizes the importance of lung segmentation in addition to other image enhancements prior to training.

Lung parenchyma

Closer approaches to the main course of pulmonary medical imaging analysis segment lung parenchyma, which is the portion of the lung involved in gas transfer, and other times segment pulmonary lobes and fissures.

Hooda et al. [100] proposed a segmentation method based on deep convolutional network targeting lungs, to indicate precise regions of interest in CXRs. Proposed models were based on standard FCN-4 architecture and applied dropout layers for comparison. The proposed model achieved satisfactory performance; 98.75% accuracy and 96.10% overlap.

Besides, Huynh and Anh [101] proposed a deep learning lung segmentation method emphasizing on large CXR images. The architecture consisted of convolutional, maxpooling, flattening, and fully connected layers. Experimenting was performed on images from Hoan My Hospital (15 images for training, 50 for testing) and 93% accuracy was

Table 5 Contributions to segmentation in pulmonary medical imaging analysis

Authors	Year	Citations	Image modality	Target
Zhang et al. [95]	2018	36	X-ray	Multi-organ
Dai et al. [96]	2018	38	CXR	Cardio-thoracic
Dong et al. [97]	2018	22	CXR	Cardio-thoracic
Gordienko et al. [99]	2018	23	CXR	Lungs and bones
Gordienko et al. [98]	2018	15	CXR	Lung and bones
Hooda et al. [100]	2019	1	CXR	Lung parenchyma
Huynh and Anh [101]	2019	1	CXR	Lung parenchyma
Kitahara et al. [102]	2019	1	CXR	Lung parenchyma
Furutani et al. [103]	2019	0	CXR	Lung parenchyma
Skourt et al. [104]	2018	22	CT	Lung parenchyma
Gerard and Reinhardt [105]	2019	1	CT	Lung parenchyma
Wang et al. [106]	2019	0	CT	Lung parenchyma
Wang et al. [82]	2017	100	CT	Pulmonary nodules
Mobiny and Van Nguyen [115]	2018	38	CT	Pulmonary nodules
Jin et al. [114]	2018	37	CT	Pulmonary nodules
Wang et al. [113]	2018	5	CT	Pulmonary nodules
Hu et al. [116]	2019	4	CT	Pulmonary nodules
Tang et al. [118]	2019	1	CT	Pulmonary nodules
Wang et al. [117]	2019	0	CT	Pulmonary nodules
Moriya et al. [120]	2019	0	CT	Lung cancer tissue
Jiang et al. [121]	2018	24	Multiple	Lung tumors
Jue et al. [122]	2019	0	CT	Lung tumors
Astaraki et al. [123]	2019	0	CT	Lung tumors
Cui et al. [107]	2019	1	CT	Pulmonary vessels
Yun et al. [108]	2019	6	CT	Airways diseases
Nadeem et al. [109]	2019	0	CT	Airways diseases
Qin et al. [110]	2019	0	CT	Airways diseases
Zhao et al. [111]	2019	0	CT	Airways diseases
Wang et al. [112]	2019	0	CT	Airways diseases

achieved versus 76% for conventional CNN. Similarly, the proposed model outperformed the conventional CNN at the level of precision, recall, F1 and Dice score.

Kitahara et al. [102] targeted lung segmentation for CXRs using Fully Convolutional Neural Networks. They aimed to reduce the mis-recognition of lungs, and used pre-processing techniques to achieve their goal. A customized dataset of inhale and exhale radiographs was used in order to validate the efficiency of model, where the change in area of lungs throughout frames was used to assess COPB presence (reduced change rate is a sign of disease), and experiments recorded 94% accuracy.

Again, Furutani et al. [103] aimed to segment lungs from CXRs proposing a model based on U-net deep architecture. The Dice-coefficient achieved was 0.91 on average.

Skourt et al. [104] proposed a lung CT image segmentation using the U-net architecture, consisting of a contracting path to extract high-level information and a symmetric expanding path to recover the needed information. Results showed an accurate segmentation with 0.9502 Dice-coefficient index, and the capability of applying it to a wide area of different segmentation tasks in medical imaging.

For the same segmentation target, Gerard and Reinhardt [105] aimed to segment pulmonary lobes and fissures from chest CT scans and thus proposed a deep learning framework made up of a novel pipeline of 3D CNN series to serve the purpose. It was experimented on COPDGene subset and achieved 0.993 Dice coefficient and 0.138 mm median average symmetric surface distance, showing the robustness of the model to different image quality, inspiration levels and pathologies.

Finally, Wang et al. [106] proposed a pulmonary lobes' segmentation model by first applying automated lung segmentation and then volumetric CNN (V-net) to CT scans. Additional feature maps are generated by coordination-

guided CNNs to reduce misclassification. The model has achieved 0.947 dice coefficient index.

Pulmonary vessels

As for pulmonary vessels, Cui et al. [107] proposed a framework for automated segmentation based on a 2.5D convolution network, where slice radius is introduced to convolve adjacent information, and multi-planar fusion is owed to optimize the presentation of intra/inter-slice features. After then, segmentation results are refined using component information of pulmonary vessel tree.

Airways diseases

Reaching for airways segmentation, approaches were held on CT scans only.

Yun et al. [108] proposed a 2.5D model that starts by extracting airway-candidate patches that are then classified by 2.5D CNN resulting with a likelihood map used for segmentation of airways finally. On the other side, Nadeem et al. [109] targeted airway segmentation from CT scans using deep learning in addition to conventional methods. The experiments showed significant advances in terms of accuracy at branch-level compared to unedited results from conventional industry method. Besides, the segmentation's leakages were significantly reduced. The proposed model involved a 3D U-Net that computes a likelihood map of the airway lumen space at total lung capacity from chest CTs, where the map is then fed into an augmentation conventional process to remove leakages consequently.

Another approach was by Qin et al. [110] who proposed a voxel-connectivity aware approach for accurate Airway segmentation that transforms conventional binary segmentation to 26 tasks of connectivity prediction, learning both airway structure and relationship between neighboring voxels, feeding the lung distance map and voxel coordinates into AirwayNet as additional semantic information.

Zhao et al. [111] proposed a two-stage 2d+3D neural network and a linear programming based tracking algorithm for airway segmentation, followed by a bronchus classification algorithm based on the segmentation results.

Last but not least, Wang et al. [112] introduced a 3D slice-by-slice convolutional model in a U-net architecture, and a novel loss function called radial distance loss.

3.2.3.2. Lesion Segmentation

Deeper approaches of segmentation target tumors and/or pulmonary nodules/tissues as an important step prior to analyzing them.

Pulmonary nodules

Firstly, Wang et al. [82] pointed out to analysis and measurement tools efficient for segmentation. They are Dice similarity coefficient DSC, which measures the overlapping between two segmentation results, the ASD which measures the average boundary distance between surfaces of two segmentation results, and the positive predictive value PPV beside the conventional specificity and sensitivity. The authors of the same paper proposed a data-driven model that segments lung nodules from heterogeneous CT images, named the Central Focused Convolutional Neural Networks CF-CNN. The proposed model aimed to capture a diverse set of features that are nodule-sensitive from 3D and 2D CT images and classify the voxels taking into consideration the neighboring voxels' effect. These key insights were targeted by a central pooling layer that retains much information on voxel patch center, followed by a patch learning strategy. Weighed sampling facilitated the training of the model that shows a performance superior to conventional models, achieving 82.15% and 80.02% dice scores for two performed experiments.

Another approach was done by Wang et al. [113] who proposed a deep region-based network (RCNN) for detection of pulmonary nodules in 3D CT images generating simultaneously a segmentation mask for each instance, in addition to a deep active self-paced learning (DASL) strategy for reducing annotation effort and making use of un-annotated samples(weekly supervised).

Jin et al. [114] coupled a 3D CGAN with a novel multi-mask loss to generate CT-realistic high-quality lung nodules conditioned on a VOI with an erased central region.

CapsNet was proposed by Mobiny and Van Nguyen [115] as alternative to CNN, besides a modified routing mechanism speeds up the process 3 times.

Based on probabilistic U-net, Hu et al. [116] proposed a model for segmentation that outputs two kinds of quantifiable uncertainty: aleatoric and epistemic.

Wang et al. [117] proposed a mixed supervised dual-network (MSDN) that consists of a network for detection and another one for segmentation, and used "Squeeze and Excitation" for transferring information from auxiliary detection to help segmentation.

NoduleNet was proposed by Tang et al. [118], which is an end-to-end 3d DCNN incorporating two design tricks: decouples feature maps for nodule detection and false-positive reduction and segmentation refinement subnet for increasing nodule segmentation precision.

At a larger scale, models genesis developed by Zhou et al. [119] aimed to generate powerful application-specific target models through transfer learning. Models genesis is a collection of generic source models built directly from unlabeled 3D image 3D data with a unified self-supervised method. It detects pulmonary nodules, but segments them first.

Lung cancer tissues

Moriya et al. [120] proposed a 2-phase deep unsupervised generative segmentation model; reconstructs image patches after inferred by categorical latent variables from unlabeled images and estimates, after training, the probability of belonging to each category to obtain segmented image.

Lung tumors

Tumors segmentation was held by several approaches too.

Jiang et al. [121] proposed an adversial domain adaptation-based deep learning approach for tumor segmentation: a tumor-aware unsupervised cross-domain adaptation(CT to MRI), then a semi-supervised tumor segmentation using U-net trained with synthesized and limited number of original MRIs. Introduced also a tumor-aware loss for unsupervised cross-domain adaptation.

Besides, Jue et al. [122] proposed a cross-modality educed deep learning segmentation that combines CT and pseudo-MR produced from CT by aligning their features to obtain segmentation on CT. The proposition was implemented using U-net and Dense-FCN.

Finally, Astaraki et al. [123] proposed normal appearance auto-encoder that automatically replaces lung nodules/masses with healthy appearing tissue, incorporates the output along with the original latter to a segmentation network, and trains the normal-appearance auto-encoder using semi-automated in-painting network.

As a result, segmentation, involving 2D and 3D scans, certainly enhances the performance of classifiers, whether at the scale of organ (lung) or deeper as for nodules. U-Net convolutional neural networks were notably used as basis for proposed architectures. Besides, the combination of segmentation with other pre-processing techniques appeared to come up with better accuracies.

3.2.4 Detection

Detection is a key part of the diagnosis, typically consisting of localization and identification of specific lesions in the image [2]. Computer-aided detection is usually referred to as CADe and sometimes as localization. It could be detection of a single disease such as TB or multi-disease detection such as anomalies detection [46]. Brief list of detection contributions can be found in Table 6.

General thoracic diseases

Starting by general thoracic disease detection, many approaches are to be considered.

Rajpurkar et al. [124] developed what they called CheXNeXt neural network, which is based on 121-DenseNet architecture to discover 14 pathologies from CXR. They trained it on ChestXray-14 dataset in two steps. Besides, the model is initialized with parameters from a network pre-trained on ImageNet dataset.

For detecting multiple abnormalities, Singh et al. [125] assessed the accuracy of Qura AI (commercially available DL algorithm) using Chest Xray 8 dataset and proposed a standard of reference to select images and direct radiologists. The algorithm was found to be of high accuracy, and might serve as second reader to improve radiologists performance.

Cai et al. [126] proposed an attention mining AM strategy to improve CNN's sensitivity or saliency to disease patterns. Moreover, the ResNet CNN model was modified to include multi-scale aggregation (MSA) to improve the localization of small-scale disease findings.

For anomaly detection on CXRs, CXNet-m1 was proposed by Xu et al. [127] based on deep learning. They aimed to overcome the conventional limitations of the existing deep learning techniques such as over-fitting and low transfer efficiency by a shorter, thinner and more powerful design than fine-tuned CNN. The CXNet-m1 achieved 67.7% accuracy, 73.6% precision, 73.8% recall, 73.7% F-measure, and 65.8% AUC. All values were the best among the experimented networks except for precision which came next after Inception-ResNet. This approach assured the importance of the proper design rather than the fine-tuning, but still agreed that the more training, the better learning process would be.

For automatic triaging of adult chest radiographs, Annarumma et al. [128] developed and tested an artificial intelligence based on deep convolutional neural networks. An NLP system was used to analyze free-text reports corresponding to images of the adopted training dataset, and the model achieved a sensitivity of 71% and specificity of 95% beside 73% and 94% for positive predictive value (PPV) and negative predictive value (NVP), respectively, assuring the clinically acceptable performance of the developed system.

On the another hand, Gerard et al. [129] targeted pulmonary fissure detection in CT images using deep learning. It is based on a novel coarse-to-fine cascade of ConvNets, named FissureNet, and a novel 3D segmentation architecture named Seg3DNet. Fissure detection was evaluated with two rule-based methods (Hessian and DoS) and two learning-based methods (FissureNet and U-Net). Upon experiment, learning-based methods outperformed rule-based methods, and FissureNet outperformed U-Net. FissureNet achieved high sensitivity for fissure detection and few false-positives were produced. It was also proved to be robust against variations in image modalities, scanning protocols and inspiration levels. The overall AUC achieved for FissureNet was 0.98 beating that of U-Net and Hessian.

Infections: pneumonia and tuberculosis

Table 6 Contributions to detection in pulmonary medical imaging analysis

Authors	Year	Citations	Image modality	Target
Rajpurkar et al. [124]	2018	107	CXR	General thoracic diseases
Annarumma et al. [128]	2019	28	CXR	General thoracic diseases
Singh et al. [125]	2018	17	CXR	General thoracic diseases
Cai et al. [126]	2018	9	CXR	General thoracic diseases
Xu et al. [127]	2019	7	CXR	General thoracic diseases
Gerard et al. [129]	2019	15	CT	General thoracic diseases
Rajpurkar et al. [130]	2017	450	CXR	Infections, pneumonia
Ayan and Unver [131]	2019	0	CXR	Infections, pneumonia
Heo et al. [132]	2019	6	CXR	Infections, tuberculosis
Ho et al. [133]	2019	0	CXR	Infections, tuberculosis
Lin et al. [134]	2019	0	CT	Pulmonary embolism
Taylor et al. [135]	2018	16	CXR	Pneumothorax
Park et al. [136]	2019	2	CXR	Pneumothorax
Nam et al. [137]	2019	53	CXR	Lung cancer
Zhao et al. [138]	2019	6	CXR	Lung cancer
Pesce et al. [145]	2017	2	CXR	Lung cancer
Chang and Moturu [149]	2019	0	CXR	Lung cancer
Setio et al. [148]	2017	493	CT	Lung cancer
Dou et al. [147]	2017	207	CT	Lung cancer
Ardila et al. [153]	2019	99	CT	Lung cancer
Huang et al. [150]	2017	66	CT	Lung cancer
Dou et al. [143]	2017	60	CT	Lung cancer
Jiang et al. [141]	2018	58	CT	Lung cancer
Hamidian et al. [139]	2017	43	CT	Lung cancer
Masood et al. [142]	2018	41	CT	Lung cancer
Yan et al. [158]	2018	36	CT	Lung cancer
Khosravan and Bagci [156]	2018	28	CT	Lung cancer
Kuan et al. [144]	2017	23	CT	Lung cancer
Zhu et al. [157]	2018	19	CT	Lung cancer
Gu et al. [151]	2018	12	CT	Lung cancer
Wang et al. [155]	2018	10	CT	Lung cancer
Bhatia et al. [159]	2019	7	CT	Lung cancer
Winkels and Cohen [160]	2019	5	CT	Lung cancer
Cha et al. [140]	2019	3	CT	Lung cancer
Wang et al. [154]	2019	2	CT	Lung cancer
Tang et al. [118]	2019	1	CT	Lung cancer
Gong et al. [152]	2019	0	CT	Lung cancer
Astaraki et al. [123]	2019	0	CT	Lung cancer
Zhang et al. [161]	2019	0	CT	Lung cancer
Gonzalez et al. [162]	2018	60	CT	Airways diseases

Reaching out for infectious diseases, pneumonia and TB were a target for localization task using deep learning-based models. Rajpurkar et al. [130] targeted pneumonia detection on CXRs and developed for that an algorithm of 121 convolutional layers. An extension was applied to detect 14 other diseases. Its F1 score value exceeded the average performance of 4 radiologists (0.435 vs. 0.387 at 95% bootstrap confidence interval), but only frontal radiographs were used with no access to patients history.

Again, for better detection of pneumonia from CXRs, Ayan and Unver [131] compared two deep learning models; VGG-16 and Xception. Both models were fine-tuned and involved transfer learning to enhance their performance. Many parameters were used to compare VGG-16

and Xception, and each outperformed the other at some. For example, VGG-16 outperformed Xception in accuracy as they recorded 87% and 82%, respectively, yet Xception was more successful at sensitivity. Each network proved to have its own capabilities even when tested on same datasets.

Whereas for TB detection, Heo et al. [132] used deep learning on chest radiographs of annual workers' health examination data using different feature extractors. Then, they compared the performances of convolutional neural networks (CNNs) based on images only (I-CNN) to CNNs including demographic variables (D-CNN). CNNs using demographic variables recorded higher AUC values (0.957 vs. 0.9714) and greater sensitivity values but less attenuating (0.815 vs. 0.775), validating that machine learning facilitates detection of TB in CXRs and demographic values improve the results.

Besides, Ho et al. [133] compared the performance of three deep learning models (ResNet-152, Inception-ResNet and DenseNet-121) for automated detection of pulmonary TB and evaluated their efficiency on chest radiography diagnosis as early detection can reduce high mortality rates due to this disease. One training dataset (ChestXray14) and two external datasets (Montgomery and Shenzhen) were used for experimenting detection of diseased cases. Pre-processing techniques were applied (augmentation and tSNE visualization) and increased average AUC of DCNNs.

Pulmonary embolism

At the level of PE, Lin et al. [134] proposed an end-to-end network that consists of: a 3D candidate proposal network for detecting cubes containing suspected PE, a 3D spatial transformation sub-net for generating fixed-sized vessel-aligned image representation for candidates, and a 2D classification network which takes the three cross sections of the transformed cubes as input and eliminates false-positives.

Pneumothorax

Moving to pneumothorax, Taylor et al. [135] developed various automated image classifiers that detect clinically significant (moderate and large) pneumothorax and trained them on a customized dataset, willing to avoid life-threatening delay of radiologists reviews in urgent cases such as overnights. Another contribution was that of Park et al. [136] who proposed a 26 layer CNN that detects pneumothorax.

Lung cancer

Lung cancer detection targets nodules and tumors of different sizes and types. It is actually the most targeted for deep learning-based detection task.

Nam et al. [137] proposed a deep CNN (DLAD) with 25 layers and 8 residual connections. It used batch normalization technique to speed up the training and used pixel intensity of chest radiography as input to output the location and malignant nodule presence.

Moreover, Zhao et al. [138] targeted detection of EGFR-mutations in pulmonary adenocarcinoma and developed a 3D-deep learning-based methodology to serve its purpose using CXRs. The proposed model, named 3D DenseNets, learns strong representations with supervised end-to-end training, and is finely tuned with another nodules subset. Augmentation is applied to avoid over-fitting, and experiments recorded 75.8% and 75% AUC for holdout and public test sets, respectively, for detection of the EGFR mutations and thus is a promising model. Besides, deep learned features were found to be related to conventional radiomics, but more robust, compact and expressive.

For automatic detection of lung nodules in chest CT, Hamidian et al. [139] trained a 3D CNN. It involved two stages: screening and discrimination, that aim to generate candidate regions of interest at the first place, and then, a more specialized CNN classifies them as nodule or background. This screening architecture reduced the size of initial search space and thus lead to a 800-fold speed-up compared to using the brute-force method of sliding the 3D CNN across the volume to obtain the classification scores for the whole CT exam. The approach was multi-scale to detect nodules with varying sizes at similar sensitivities compared to one another, and the recorded results were 80% and 95% sensitivities at 22.5 and 563 false positive rates, respectively.

Cha et al. [140] proposed a deep convolutional neural network-based model for detecting operable lung cancer with chest radiographs. It resulted with an overall sensitivity of 76.8% with a 0.3 false-positive per image and AUC of 0.732. The sensitivity of DLM performance was superior to average of 6 human readers and demonstrated its high diagnostic capability.

Adding to that, Jiang et al. [141] proposed an effective lung nodule detection scheme based on multi-group patches cut out from the lung images, pre-processed by the Frangi filter. The results demonstrated that the multi-group patch-based learning system is efficient to improve the performance of lung nodule detection and greatly reduces the false-positives in the case of huge data by achieving 80.06% sensitivity with 4.7 false-positives per scan and 94% sensitivity with 15.1 false-positives per scan.

Masood et al. [142] proposed an IoT-enabled computer-assisted decision support system for detection of pulmonary cancer and classification of its stages by using a novel deep learning-based model and metastasis information obtained from medical body area network (MBAN). DFCNet is proposed and based on the deep fully convolutional neural network (FCNN) which is used for classification of each

detected pulmonary nodule into four lung cancer stages. The proposed architecture achieved 84.58% accuracy for DFC-Net in comparison with 77.6% for conventional CNN and a potential for generalization to detect other cancer types.

Also for pulmonary nodule detection, Dou et al. [143] proposed 3D ConNets along with online sample filtering and hybrid-loss residual learning. The framework consisted of two stages: candidate screening where 3D FCNN is established and trained with the online sample filtering, and then, false-positive reduction using hybrid-loss residual network that control the information of nodules (location and size) to guide the nodule recognition process. The overall sensitivity of the framework at the last stage along with FCN, OSF, ResNet and HL was 90.5% at 1 false-positive per scan.

Kuan et al. [144] presented a framework for computer-aided lung cancer diagnosis, that ranked 41st out of 1972 teams in the Kaggle Data Science Bowl 2017. They aimed to detect the nodules in 3D CAT scans and then classify them as malignant or not, to finally assign a cancer probability based on the results. The log-loss valued 0.52712 where only 4 features were used in the competition (number of nodules, mean, std, and sum of softmax output). For additional features, results for the Malignancy detector recorded 0.719 sensitivity, 0.653 specificity, 0.558 F1 score, and 0.484 log-loss. The combination of detector with the nodule classifier achieved better results except for sensitivity value.

In 2017, Pesce et al. [145] contribution was published and then got updated in 2019 [146]. It basically proposed two architectures for lung nodules detection from chest radiographs using visual attention networks; the first is CNN with attention feedback CONAF and the other is recurrent with annotation feedback RAMAF, accompanied with an NLP system for tagging images automatically to be validated. CONAF achieved for Localization the highest sensitivity and average overlap when compared lesions to normal only (0.74 and 0.45 respectively) and same when compared lesions to all others (0.65 and 0.43 respectively). Yet, for precision, CONAF recorded 0.21 for lesions vs. normal only and 0.15 for lesion vs. all others.

For the sake of false positive reduction in automated pulmonary nodule detection, Dou et al. [147] proposed a novel method employing three-dimensional convolutional networks. The proposed method used volumetric CT scans, that adds to the 3D architecture to encode more spatial information and extract more Representative features than 2D architectures using 2D training samples. The methodology embedded a multilevel strategy to meet the challenges caused by the variations and hard mimics of pulmonary nodules. An advantage of the framework is its generalizability as it can be extended to other 3D detection tasks.

Using multi-view convolutional networks (ConvNets), Setio et al. [148] proposed a CAD system for pulmonary nodules. Discriminative features were automatically learned from the training data. The input of network is nodule candidates obtained by combining three candidate detectors specifically designed for solid, sub-solid, and large nodules. For each candidate, a set of 2-D patches from differently oriented planes should be extracted. The proposed architecture comprises multiple streams of 2-D ConvNets, where the outputs are then combined using a dedicated fusion method to yield the final classification. In order to avoid over-fitting, data augmentation and dropout were applied. Experimenting the proposed framework resulted with a 90.1% and 85.4% sensitivity values at 4 and 1 false-positives per scan, respectively.

On the other hand, Chang and Moturu [149] targeted detecting early stage lung cancer using synthetically generated X-rays throughout a model based on CNN. Pre-processing techniques were applied in order to generate the X-Rays from CTs (due to the lack of real X-rays according to authors), followed by a random generation and placement of nodules to optimize the training process. The model achieved 97.45% validation accuracy for 1–3 cm diameter sizes of nodules and 1000HU[3] for radio-density. On the other hand, the least recorded accuracy was 70.55% at 3–150 HU radio-density for 0.3–3 cm diameter of nodule size. The main barriers are declared as optimizing the hyper-parameters, guaranteeing enough variability of training data, tweaking the size of patches and sliding window, and fine-tuning several other parameters.

Huang et al. [150] targeted lung nodule detection mainly using CT scans. The proposed system was based on 3D CNNs, leveraging perturbing anatomical structures and data-driven machine-learned features. The system first generates candidate nodules and estimates their local orientation. Candidates are then fed into a trained 3D CNN to predict whether they are nodules or not, achieving 90% sensitivity at 5 false-positives rate. Authors concluded the efficiency of involving a priori information and the preference of 3D CNN over 2D CNN for volumetric medical image analysis.

Continuing with lung nodule detection, Gu et al. [151] also used deep convolutional neural networks. The proposed model used a multi-scale prediction methodology designed for chest CTs, providing three schemes for selection according to need. As 3D CNNs can utilize richer spatial 3D contextual information than 2D CNNs; the proposed schemes including multi-scale cubes can be an outstanding solution for extremely small nodules' detection. The sensitivities were recorded at 1 and 4 false-positives rates and were 87.94% and 92.93%, respectively, implying the feasibility of extending the system to other medical fields.

Gong et al. [152] showed that deep learning observers strongly correlate to human observer performance, as it was proposed and tested on localization of lung nodules in chest

---

[3] Hounsfield units used to measure radio-density.

CT scans. A local customized dataset was used consisting of different variables instantiating varying experimental conditions (nodules sizes, nodule types, radiation dose levels, etc.). The correlation was measured by Pearson's coefficient and recorded 0.988 with 95% confidence interval.

Willing to improve lung screening using CT scans, Ardila et al. [153] proposed a model based on deep learning which performed equally to six radiologists when prior CT was provided. However, without prior CT, it outperformed the radiologists reducing false-positive rates by 11% and false-negatives by 5%.

Wang et al. [154] targeted lung cancer detection using deep CNN in addition to a random forest classifier to detail the diagnosis. A 3D attention-based deep CNN is proposed by the authors, using CT images, to detect lung cancer without prior suspicions of interest regions. Demographic clinical features were used for the classifier. Accuracy recorded for attention network and AUC recorded for clinical demographic features each apart were 68.7% and 63.5%, respectively. When combining both, AUC records 78.7%.

While Wang et al. [155] proposed a pulmonary detection framework consisting of feature pyramid network, conditional 3D non-maximum suppression, and an attention 3D CNN, Khosravan and Bagci [156] used a single feed-forward pass of a single network for detection, designed as a 3D CNN with dense connections and trained in an end-to-end manner.

Zhu et al. [157] proposed a deep 3D ConvNet framework augmented with expectation-maximization (EM), to mine weakly supervised labels in EMRs for pulmonary nodule detection.

To incorporate 3D context information efficiently, Yan et al. [158] developed a 3D context-enhanced region-based CNN by aggregating feature maps of 2D images.

Astaraki et al. [123] proposed normal appearance auto-encoder (NAA) that automatically replaces lung nodules/masses with healthy appearing tissue, incorporating the output along with the original latter to a segmentation network, and training the NAA using semi-automated in-painting network.

NoduleNet was developed by Tang et al. [118], which is an end-to-end 3d DCNN, incorporating two design tricks: decouples feature maps for nodule detection and false-positive reduction, and segmentation refinement subnet for increasing nodule segmentation precision.

Bhatia et al. [159] delineated a pipeline of pre-processing techniques highlighting lung regions and extracting features using U-net and ResNet models. Multiple classifiers are used after then to predict the probability of the CT scan being cancerous, trained on LIDC-IDRI dataset.

Winkels and Cohen [160] proposed 3D roto-translation group convolutions instead of standard translational convolutions in application to pulmonary nodules. Baseline network used consisted of 6 convolutional layers, batch normalization and ReLu nonlinearities, 3D max pooling and fully connected layer.

Last but not least, Zhang et al. [161] proposed a model that first segments lung parenchyma by region growing method, then comes the model PndDBN-5 which consists of three Restricted Boltzmann Machines (RBM).

Airways diseases

To detect and stage COPD and consecutively predict acute respiratory disease (ARD) using chest CT scans, Gonzalez et al. [162] studied the capability of CNN involving logistic and cox regression to assess COPD and mortality respectively. They have realized that CNNs provide a flexible and fast method that may allow assessing population-wide diseases, in addition to proving its efficiency for the previously mentioned objective.

Plenty of work has been done on detection in pulmonary medical imaging analysis, and proposed models proved to be competitively efficient at both speed and performance. Yet, many approaches involved external factors which improved their performance, such as: NLP to benefit from textual data, IoT to bring live data into the process, demographic features to add input information, pre-filtration of input patches, etc. In addition, most of the work on detection targeted lung cancer, specifically using CT modality, and a fair number of approaches targeted infections and general thoracic diseases, for which CXRs were used. Moreover, CXRs usage dominated pneumothorax while CT dominated PE and airways diseases. Relating to the architectures, AleXnet, ResNet, DenseNet, Inception and region-growing networks were used themselves or after editing, and in some approaches, segmentation (such as U-net-based models) was aggregated to the detector to support the process.

3.2.5 Features extraction

Features extraction is the characteristic discriminating deep learning from traditional machine learning methods, as it is done automatically in DL while done exhaustively manually in ML. Features are important for the learning process and could be acquired from training medical datasets, natural images datasets or by transfer learning from pre-trained medical image analysis networks. For a list of the contributions to this task, refer to Table 7. Contributions to feature extraction are clustered according to target: general thoracic diseases, infections, and pulmonary nodules.

General thoracic diseases

Nemoto et al. [163] aimed to generate features from normal volume patches only, using a deep convolutional auto-encoder (D-CAE) network. The D-CAE is trained by CT

Table 7 Contributions to features extraction in pulmonary medical imaging analysis

Authors	Year	Citations	Image modality	Target
Chen et al. [164]	2019	1	CXR	General thoracic diseases
Nemoto et al. [163]	2019	0	CT	General thoracic diseases
Lopes and Valiati [165]	2017	33	CXR	Infections, tuberculosis
Gozes and Greenspan [166]	2019	0	CXR	Infections, tuberculosis
Liang and Zheng [167]	2019	2	CXR	Infections, pneumonia
Chen et al. [168]	2017	45	CT	Pulmonary nodules

scans that include no lesions, which thus calculates the low-dimensional features and reproduction error from 2.5D volume patches. Combining both previous features and deep-CAE generated features increased the AUC to 98.9%.

Moreover, Chen et al. [164] proposed a dual asymmetric DCNN model, that is a complementary combination of ResNet and DenseNet such that feature extraction is imposed at two levels: feature and decision, then combining the two loss functions from both networks. The model functions as a multi-label thoracic disease classifier and was proved to be effective with respect to state-of-the-art baselines throughout experimenting on ChestX-ray14 dataset.

Infections: pneumonia and tuberculosis

Targeting infections, Lopes and Valiati [165] proposed three models for feature extraction applied on TB to refute that fine-tuned CNNs always surpass pre-trained ones. The first model used different CNN architectures (examples is VGG-19) to extract features each at once from resized images, which were then fed into SVM classifier. The second model took the same three CNN architectures and allowed them to extract features from certain regions of interest ROI, then combined them creating a global descriptor used to train a SVM. The final model was made up of best SVMs trained in models 1 and 2 to create ensembles of classifiers. Based on the results obtained in this paper, pre-trained networks were validated for their usefulness and power. Referring to the results, it is recommended that model 2 be applied on high-resolution datasets to extract valuable global descriptor even though it will need too much time. Different results may occur if other CNN architectures, classifiers and methods for visual dictionary generation were used instead of GoogLenet, ResNet and VGG, Support Vector Machine (SVM), and K-clustering method.

On the same side, Gozes and Greenspan [166] targeted TB willing to study impact of feature learning (pre-training deep model) specifically on CXR using DenseNet-121 CNN. The application incorporated meta data and trained the model on ChestX-ray14 dataset which includes 14 thoracic pathologies. The feature learning allows for better transfer learning on small-scale datasets for Tb (On Shenzen dataset: AUC recorded was 96.5%).

Also about infections, Liang and Zheng [167] aimed to develop a deep learning-based model that overcomes lack of spatial information in conventional deep CNN feature extraction, and thus improve accuracy of classifiers. In order to detect pneumonia in children CXRs, they proposed a framework that combined dilated convolution and residual thought to avoid over-fitting, depth-model degradation problems and spatial information loss. Results recorded were 96.7% recall rate and 92.7% f1-score. The model is considered reliable for classification of children pneumonia in CXRs.

Pulmonary nodules

Emphasizing on pulmonary nodules, Chen et al. [168] exploited three different multi-task learning schemes (MTL) to take advantage of heterogeneous computational features derived from deep learning models of convolutional neural network (CNN) and stacked de-noising auto-encoder (SDAE), in addition to hand-crafted Haar-like and HoG features. These extracted features aim to ease the description of 9 semantic features for lung nodules in CT images. As each semantic feature is considered an individual task, heterogeneous computational features are tasks selected by MTL schemes and mapped toward radiologists' ratings along with cross-validation evaluation schemes on the nodules that were randomly selected from LIDC dataset. Results predicted that MTL schemes ratings were closer to radiologists' than single-task methods and were considered robust. Besides, results of co-training CNN regression were more accurate than single-task regression, but did not surpass the multi-task regression. On the other hand, it was concluded that deeper CNN is not always equal to better regression performance, but bigger training datasets are expected to yield better regression results with CNN. Besides, the combination of all heterogeneous features can effectively boost the results.

Feature extraction is indirectly addressed in a lot of approaches tackling deep learning. They all disregard manual extraction of features, but prefer to maximize the number of features involved for better classification (also referred to as discrimination), noting that the more training is applied

to the architecture, better chances of surpassing fine-tuned CNNs are exposed. Children were targeted in one approach, and age prediction was suggested in another which introduce deep learning objectives. More about classification is present in the next part.

3.2.6 Classification

Classification is an important task in medical image analysis, that comes just after feature extraction and representation. It aims to map the input variables (images or records) into output variables that represent a specific class such as "diseased" or "healthy" [26]. Table 8 lists contributions to this task. They are clustered according to target: general thoracic diseases, interstitial lung diseases, infections, pulmonary edema, airway diseases and lung cancer.

General thoracic diseases

Contributions which handled general thoracic diseases classification are discussed first. Abiyev and Maaitah [169] demonstrated the feasibility of classifying chest pathologies in CXRs using CNN, and compares CNN, BPNN and CpNN networks.

DLAD was proposed by Hwang et al. [170], which is a classification algorithm with dense blocks comprising 5 classifiers, 1 for each disease and the fifth for normal/abnormal classification. Two loss types were used for training: classification and localization.

CXRs classification into anteroposterior or posteroanterior views was held by Kim et al. [171] who developed a ResNet-18 DCNN. It was trained on NIH ChestXray14 database that consists of adults and pediatrics CXRs. Another similar network was developed and trained only on pediatric CXRs. Recorded AUC values and accuracies were 99.7% and 98%, respectively, for pediatrics trained network, and 100% and 99.6%, respectively, for fully trained network (adults and pediatrics). Similarly, sensitivity and specificity of fully trained network outperformed those of pediatrics-trained network. However, the reduction is slight with respect to 95% training data reduction.

Tang et al. [172] targeted identification of abnormal CXR using the proposed model based on generative adversarial one-class identifier. The model is mainly trained to identify normal CXR by reconstructing it. If the input image is abnormal, the reconstructed image will be poor and thus be identified. The model has achieved 84.1% AUC with architecture composed of three DCNNs; auto-encoder U-Net, discriminator and decoder.

Besides, Pan et al. [62] compared two deep learning classifiers: DenseNet and MobileNet-V2, that were trained by both Rhode Island Hospital chest radiograph (RIH-CXR) and Health ChestXray14 (NIH-CXR) datasets for nor-

mal/abnormal and 14 thoracic diseases classification. When tested for normal/abnormal classification, DenseNet and MobileNet-V2 recorded 90% and 89.3% AUROC, respectively, when trained by NIH-CXR, versus 96% and 95.1% on RIH-CXR. On the other hand, MobileNet-V2 recorded AUROC with 1% average distance within that of DenseNet. As a result, MobileNet-V2 and DenseNet performance was comparable, and decreases slightly when tested by external dataset, which should be taken into consideration when applying to other institutions' datasets.

Evaluating the effect of augmentation on classification, Ogawa et al. [173] used augmented training datasets for abnormal chest radiographs detection based on deep convolutional neural network. The augmentation was followed by a binary classification of the images and the accuracy measured was higher than that with non-augmented datasets.

The ability of deep learning classification methods to handle label noise was studied by Calli et al. [174]. The test was applied on chest radiographs and exactly on ChestX-ray14 publicly available dataset. The experiments revealed that deep learning methods are robust against label noise but are not completely insensitive to them. Results show that 16% and 32% training label noise cause only 1.5% and 4.6% respective drops in accuracy.

Another approach to general classification was age prediction. Karargyris et al. [175] aimed to predict the patient's age from his CXR and compare it to his actual age to improve counseling particularly when there is a notable difference. CNN was trained in regression on a large publicly available dataset, and heat maps were explored to realize the significance of areas near spine, shoulders, mediastinum and clavicles for age prediction.

Wong et al. [176] aimed to classify the disease-free CXRs without risking to discharge sick patients. The proposed architecture is based on Inception-ResNet-v2 model that is trained using ImageNet to provide image features, then by CXRs labeled by radiologists. The precision is optimized to 100% and the recall to 50% in order to classify a good number of normal X-rays (but not all) making clinicians work easier.

Similarly, classification of CXR was the target of Ma et al. [177] who proposed a novel scheme of cross-attention networks (CAN) for automated thoracic disease classification, where features were got from pumping images into two networks that have different initializations, which will then go into ReLU layers. The feature maps then would be the input to a transition layer to transform two groups of features into the same shape, then cross-attention feature maps are produced using Hadamard product.

Purkayastha et al. [178] documented the implementation of deep learning in LibreHealth Radiology, a version of a modern electronic health record system (LiberHealth EHR) that is dedicated to radiology and imaging professionals. The

Table 8 Contributions to classification in pulmonary medical imaging analysis

Authors	Year	Citations	Image modality	Target
Abiyev and Maaitah [169]	2018	26	CXR	General thoracic diseases
Hwang et al. [170]	2019	18	CXR	General thoracic diseases
Tang et al. [172]	2019	5	CXR	General thoracic diseases
Pan et al. [62]	2019	5	CXR	General thoracic diseases
Ogawa et al. [173]	2019	2	CXR	General thoracic diseases
Kim et al. [171]	2019	2	CXR	General thoracic diseases
Calli et al. [174]	2019	1	CXR	General thoracic diseases
Purkayastha et al. [178]	2019	1	CXR	General thoracic diseases
Karargyris et al. [175]	2019	0	CXR	General thoracic diseases
Wong et al. [176]	2019	0	CXR	General thoracic diseases
Ma et al. [177]	2019	0	CXR	General thoracic diseases
Tang et al. [179]	2019	1	CT	General thoracic diseases
Gao et al. [180]	2018	118	CT	Interstitial lung diseases
Kim et al. [181]	2018	21	CT	Interstitial lung diseases
Lakhani and Sundaram [81]	2017	416	CXR	Infections, tuberculosis
Raju et al. [182]	2019	0	CXR	Infections, tuberculosis
Zech et al. [183]	2018	101	CXR	Infections, pneumonia
Stephen et al. [80]	2019	4	CXR	Infections, pneumonia
Wang et al. [184]	2019	0	CXR	Pulmonary edema
Zucker et al. [185]	2019	1	CXR	Airway diseases
Zhao et al. [111]	2019	0	CT	Airway diseases
Wang et al. [186]	2019	2	CXR	Pneumothorax
Pesce et al. [145]	2017	2	CXR	Lung Cancer
Takemiya et al. [187]	2019	1	CXR	Lung cancer
Shen et al. [192]	2017	172	CT	Lung cancer
Ciompi et al. [194]	2017	129	CT	Lung cancer
Liao et al. [196]	2019	76	CT	Lung cancer
Song et al. [203]	2017	69	CT	Lung cancer
Hussein et al. [197]	2017	46	CT	Lung cancer
Masood et al. [142]	2018	41	CT	Lung cancer
Hussein et al. [188]	2017	39	CT	Lung cancer
Wang et al. [195]	2017	37	CT	Lung cancer
Shen et al. [190]	2019	22	CT	Lung cancer
da Silva et al. [191]	2018	25	CT	Lung cancer
Kuan et al. [144]	2017	23	CT	Lung cancer
Xu et al. [199]	2019	19	CT	Lung cancer
Liu et al. [193]	2017	13	CT	Lung cancer
Khosravan et al. [198]	2019	9	CT	Lung cancer
Xie et al. [202]	2019	4	CT	Lung cancer
Luckehe and von Voigt [189]	2018	3	CT	Lung cancer
Srivastava and Purwar [201]	2019	1	CT	Lung cancer
Byun et al. [200]	2019	0	CT	Lung cancer
Zhou et al. [119]	2019	4	Multiple	Lung cancer

approach integrated CheXNet into LibreHealth Radiology allowing physicians to upload chest radiographs to diagnose in means of 30 seconds 14 thoracic diseases, reducing the usual delay and surpassing shortage of trained radiologists. A web service is provided to allow clients with poor computational resources to make use of the system, achieving 86% accuracy.

Finally, an approach used CT scans for classification. Tang et al. [179] targeted classification of four lung diseases: pneumonia, nodule, pulmonary edema and atelectasis using case-level weak supervision. A local dataset was prepared and labeled based on radiologists reports after being analyzed by rule-based models. Ten CT slices for each patient held the same label yet possibly did not all show the diseases. Performance of deep classifier (ResNet-50 with fourfold cross-validation) was recorded on slice-level (standalone slice) and on patient-level (mean probability of five slices chosen where they have the highest probability). For slice-level, AUC records were 71% for nodule, 79% for atelectasis, 96% edema, and 90% for pneumonia, whereas on patient-level, AUC recorded 74% for nodule, 83% for atelectasis, 97% for edema, and 91% for pneumonia. In addition, a heat map is generated to approximate the disease detector.

Interstitial lung diseases

ILD were targeted by two approaches. Gao et al. [180] proposed a network consisting of five convolutional layers, three FC layers and a softmax classification of 6 classes. Kim et al. [181] also employed convolutional neural network containing 6 layers (4 convolutional and 2 fully connected) for classification of diffuse lung disease regional patterns, and compared it with a shallow learning method (Support Vector Machine). The deep learning method significantly outperformed shallow ones, as the classification accuracy of CNN recorded 95.12%. Clinical information and additional training data are expected to enhance the performance.

Infections: pneumonia and tuberculosis

Classification of TB-diseased or pnuemonic images from healthy ones was also approached from deep learning point of view. Lakhani and Sundaram [81] compared two different deep convolutional neural networks: AlexNet and GoogLeNet. Both models were used to classify images as diseased with pulmonary TB or as healthy, and used for that both trained and untrained networks of ImageNet. Augmentation was applied as well, in addition to multiple pre-processing techniques. Ensembles were performed on the best-performing algorithms. In cases were classifiers resulted in contradictory classifications, an independent board-certified cardio-thoracic radiologist interpreted the images. An ensemble of the two DCCNs performed the best as it recorded the highest AUC (0.99). Besides, the pre-trained models surpassed untrained ones. Moreover, augmentation increased accuracies and radiologists assessment in cases of disagreement further improved the results.

Similarly, Raju et al. [182] targeted TB detection using a complete CNN approach, applied on CXRs. Data were intensified at the edges and then cropped by identifying the background. Resizing and normalization of pixel values were then applied before implementing two suggested methods. The first method proposed was deep residual network, and the second was Oxfordnet. For method 1, sensitivity was 82.08% and specificity was 93.80%, while for method 2, sensitivity was 84.91% and specificity was 93.02%. Additional training of the models would increase their robustness and applying more pre-processing methods can improve the performance.

Moving to pneumonia, Zech et al. [183] aimed to test the performance of deep learning models with variable generalization. To meet their objective, they compared the datasets impact on CNN classifiers used for pneumonia detection in chest radiographs. As a result, five natural comparison models were considered, each built up from different combinations of training and validation datasets. The experiment revealed that CNNs were best performing when internal datasets were involved (3 out of 5 natural comparisons), which may confound disease predictions. The best values achieved were 73.2% for accuracy, 93.4% for AUC, 95% for sensitivity and 70.9% for specificity.

Moreover, Stephen et al. [80] proposed a deep learning-based model for classification of pneumonia using CXRs from [63] dataset. The proposed model did not include transfer learning. Instead, it was built to extract features from the input X-ray and classify it. Several augmentation techniques were applied to enhance the validation accuracy allowing it to achieve remarkable results. The accuracies recorded best at 200 dataset size, where training and validation accuracies recorded 95.31% and 93.73% respectively.

Pulmonary edema

For the classification of pulmonary Edema severity using CXRs, Wang et al. [184] compared a number of deep learning based models; DenseNet, ResNet-50, Inception-V3, InceptionResNet-V2, NASNetMobile, DenseNet w/lung ROI, DenseNet w/Semi-Supervised I, DenseNet w/Semi-Supervised II). A large-scale dataset was used (MIMIC-CXR) and highest AUC recorded for multi-class severity classification was for DenseNet w/ Semi-Supervised II (81.3%), as well as for no-pneumonia and mild pneumonia (85.3% and 74.7% respectively). DenseNet w/ Semi-Supervised I achieved best AUC for severe pneumonia (88.9%) while NASNetMobile achieved best AUC record for moderate pneumonia. As a result, semi-supervised learning via self-training with pseudo labeling is promising with

respect to dealing with large-scale unlabeled images tested for pulmonary edema.

Airway diseases

However, among airways diseases classification, Zucker et al. [185] aimed to investigate the hypothesis that DCNN can facilitate automated Brasfield scoring of CXRs and concluded its promising accuracy, similar to or exceeding that of board-certified pediatric radiologists, except for the air trapping and large lesions subfeatures. Note that Brasfield scoring system is specific to cystic fibrosis evaluation using CXRs.

Adding to that, Zhao et al. [111] proposed a two-stage 2D/3D neural network and a linear programming based tracking algorithm for airway segmentation, followed by a bronchus classification algorithm based on segmentation results.

Pneumothorax

Targeting pneumothorax pulmonary disease classification, Wang et al. [186] proposed a deep-learning-based image classification method for using CXRs. It is composed of a DCNN that features a network in network (NIN) for data cleaning and random histogram equalization data augmentation processing. The method's efficiency was validated as experiments yielded 98.44% AUC and 99.06% on ZJU-2 and ChestXray14 datasets, respectively.

Lung cancer

Targeting lung cancer in its different forms with classification task is a novel approach with plenty of contribution.

At the level of CXR, the contribution of Pesce et al. [145] was updated in 2019 [146]. It basically proposed two architectures for lung nodules detection from chest radiographs using visual attention networks; the first is CNN with attention feedback CONAF and the other is recurrent with annotation feedback RAMAF, accompanied with an NLP system for tagging images automatically to be validated. For classification evaluation, CONAF achieved highest accuracy, F1 score, sensitivity and precision when compared lesions to normal only (0.85, 0.85, 0.78 and 0.92, respectively), and highest accuracy and F1 score only when compared lesions to all others (0.76 and 0.67, respectively). Yet, it recorded 0.74 sensitivity and 0.6 precision when compared to all others.

At the same level, Takemiya et al. [187] targeted pulmonary nodules detection using Region-CNN on chest radiographs. They first selected candidate regions from X-rays by selective search, then applied CNN classifier to classify nodules from non-nodule opacities.

However, among CT image modality, more contributions were published.

Hussein et al. [188] targeted lung nodules classification as malignant or non-malignant using CNNs for feature extraction, taking into consideration the notable variations in appearance between nodules. The model proposed is a multi-view CNN that used first Median Intensity projection to produce three 2D patches each corresponding to a certain dimension. A tensor is formed by concatenating the three patches which serve as different input image channels. Data augmentation is applied after then, and the CNN network extracts the features from augmented input images to finally undergo Gaussian Process regression obtaining the malignancy score. High level attributes achieve 86.58% regression accuracy (0.59 SEM%) while adding CNN to the attributes increases regression accuracy to 92.31% (1.59 SEM%).

Besides, Luckehe and von Voigt [189] aimed to simplify the image based on an evolutionary algorithm EA, in order to focus on relevant parts in classification using CNN. This showed that even though 50% of the pixels were simplified, meaningfulness was preserved and run-time was enhanced.

In addition, Shen et al. [190] targeted lung nodule malignancy classification specifically in CT scans through a hierarchical semantic convolutional neural network (HSCNN). They provided a network of two output levels: a low-level that quantifies five diagnostic features used by radiologists and can explain how the model interprets images professionally; and a high-level malignancy prediction score. The high-level output takes information from low-level tasks altogether with representations learned by convolutional layers, and use them to predict the malignancy score. The proposed model's experimental results surpassed common 3D CNN approaches and showed notable advance in the interpretability of the model. Yet, the features selected did not cover all semantic ones, and labels of LIDC dataset were not reflecting pathological diagnosis.

DFCNet based on deep fully convolutional neural network was proposed by Masood et al. [142], targeting pulmonary cancer detection and stage classification in CT images. Cancer detected with the help of metastasis information obtained from MBAN (Medical Body Area Network—IoT) was then classified into one of four lung cancer stages. The accuracy of DFCNet recorded 84.58% while 77.6% for CNN, and the proposed model was considered as generic that could cover different cancer types.

da Silva et al. [191] used paper swarm optimization PSO algorithm to optimize the convolutional neural network's hyper-parameters such as pooling type, number of batches training and dropout probabilities. The PSO involvement aimed to reduce the false-positives' rate and eliminate the need for manual search. The proposed method was compared to different conventional and deep learning techniques and recorded the best performance rates especially sensitiv-

ity. The results were 97.62% for accuracy, 95.5% for AUC, 98.64% for specificity and 92.2% for sensitivity.

On the same side, Shen et al. [192] presented a multi-crop CNN that automatically extracts nodule remarkable information by employing a pooling strategy that crops different regions from convolutional feature maps. After then, it effectively applies max-pooling different times. In addition to lung nodule malignancy classification, the proposed model characterizes semantic attributes and diameter of nodules, which are potentially helpful in modeling nodule malignancy. The highest classification accuracy obtained was 87.14% where MC-CNN had 64 layers, while AUC's best record (0.93) was at 16 layers of MC-CNN. Specificity and sensitivity recorded 77% and 93%, respectively.

Liu et al. [193] presented a 3d CNN trained from scratch, to classify pulmonary nodule malignancy. Different combinations of traditional machine learning models and 3D CNNs were used to create ensembles for classification. Results showed that models involving 3D CNNs—whether single or ensemble—outperformed those which do not. Furthermore, an ensemble that joins traditional models to 3D CNN yields complementary information enhancing its performance. AUC recorded for single 3D CNN was 73.2% while that for ensemble model including 3D CNN was 78%.

Also, Ciompi et al. [194] targeted the automatic pulmonary nodule management by applying deep learning on lung cancer screening using CT scans. It proposed a model based on multi-stream multi-scale convolutional networks that aim to classify all nodule types without the need for any information. The system eliminated need for additional pre-processing such as segmentation of nodules and their sizes by learning to analyze arbitrary number of 2D views of a given nodule and forming a 3D representation of it. Experiments show a performance better than that of classical machine learning approaches (accuracy between 78% and 79.5%), and within the inter-observer variability. More training data can enhance some tests' recall and precision values.

Adding to contributions, Wang et al. [195] targeted nodule classification using chest radiography through a proposed deep feature fusion process involving non-medical and handcrafted features willing to reduce false-positive rate. Experiments showed that the fusion of both deep model features and handcrafted ones outperform using only single handcrafted features as sensitivities and specificities recorded were, respectively, 69.3% and 96.2% for deep fusion, and 62% and 95.4% for single handcrafted features. Authors look forward to build a big dataset from clinical data to train the network in the future.

Liao et al. [196] targeted malignancy of pulmonary nodules proposing a 3D deep neural network consisting of two parts: a 3D region proposal network for nodule detection, and another one that selects the top five nodules based on the detection confidence evaluating their probability to be cancerous. Both parts are modified U-net models, and the final architecture came first in Data Science Bowl 2017 competition. It recorded 81.42% accuracy at 0.5 threshold, and 69.76% at threshold 1. Moreover, the AUC recorded 87% at threshold 0.5. Ranking 41st out of 1972 teams in the Kaggle Data Science Bowl 2017, Kuan et al. [144] presented a framework for computer-aided lung cancer diagnosis. It aimed to detect the nodules in 3D CAT scans and then classify them as malignant or not, to finally assign a cancer probability based on the results. Log-loss valued 0.52712 where only 4 features were used in the competition (number of nodules, mean, std, and sum of softmax output). For additional features, results for the Nodule Classifier NC recorded 0.632 sensitivity, 0.582 specificity, 0.474 F1 score, and 0.578 log-loss. The combination of detector with nodule classifier achieved better results in all metrics.

Hussein et al. [197] proposed a framework that aims to classify pulmonary nodules whether malignant or not. They employed transfer learning over 3D CNN in order to enhance the characterization of nodules. First, the architecture fine-tunes 3D CNNs involving 6 attributes in addition to a malignancy label. Each attribute and label are passed into a different 3D CNN where each consists of 5 convolution, 5 max pooling and 2 fully connected layers. Fusion of features comes next, then finally the multiplication with coefficient vector to obtain malignancy score. The proposed model achieved 91.26% accuracy upon experimenting. Including PET scans with CT seems promising on the level of diagnostic accuracy improvement.

Moreover, [198] developed an eye-tracking interface that is out of scope of this survey. The eye-tracking data and a CAD system are unified using an algorithm that involves graph-based clustering and sparsification, in order to interpret gaze patterns both quantitatively and qualitatively. Furthermore, segmentation and suspicious areas diagnosis are performed by an incorporated deep learning multi-task platform. Tests were made on low dose chest CT scans and specifically lung cancer screening, but showed a possibility of generalizing the framework to cover more complex complications and different image modalities. The best accuracy recorded for classification was 97%, and the DSC for segmentation was 91%. The common limitation of availability and abundance of training data was present too.

Besides, Xu et al. [199] targeted lung cancer treatment response prediction using deep learning on serial CT imaging. Two local datasets of stage-3 non-small cell lung cancer patients' CT scans were customized to train and validate the efficiency of a CNN with RNN model. The datasets involved images from pre-treatment and post-treatment follow-up (1, 3 and 6 months later). It was noted that each additional follow-up scan added for training has enhanced the model performance. Result of the model was grouping images according to mortality risk and evidence of the ability to

integrate multiple-time-points scans into the deep learning approach.

Another contribution was by Byun et al. [200], which targeted ground-glass nodules (GGN) classification in chest CT scans. The methodology proposed starts with image augmentation and background removal to enhance the input image and then uses a GGN-Net classifier that classifies GGNs into three classes using multiple input images, and the classification performance is evaluated according to input images types. The proposed model achieved 82.79% accuracy which was higher than single input images by 10.35%, 13.79%, and 6.90% for intensity-based, texture-enhanced and shape-enhanced images, respectively.

Srivastava and Purwar [201] aimed to simplify the deep learning classification process of lungs' CT images by embedding six external shape-based features; viz. solidity, circularity, discrete Fourier transform of radial length function, histogram of oriented gradient, moment, and histogram of active contour image. Experiments recorded 95.26% precision average and 69.56% recall average for two databases.

Also, Ogawa et al. [173] aimed to study the impact of augmentation on binary classification using DCNNs when applied to chest radiographs. The model was trained using images augmented by many operations, either alone or combined: rotation, horizontal and vertical flipping, brightness variation and Gaussian blur. Augmentation improved the accuracy of the network model, and the best record was achieved when rotation and horizontal flipping were applied together (91%).

Last but not least, Xie et al. [202] targeted benign-malignant lung nodule classification on chest CT scans. The proposed model is built up from two parts: an adversarial auto-encoder-based unsupervised reconstruction network and a supervised classification network. The two parts of the model are connected by learnable transition layers for adaption. An extension of the model was applied to characterize each nodule's overall features and was experimented using LIDC-IDRI dataset recording 92.53% accuracy and 95.81% AUC.

Same as in detection, lung cancer was mostly the target in classification. However, general thoracic disease were also targeted by a good number of approaches relatively. The image modalities used in each targeted topic are almost the same as those used in detection too; CXR for general thoracic diseases, PE and infections, whereas CT was used for ILD and lung cancer. Airway diseases were approached from both perspectives (CXR and CT) in a couple of approaches, respectively. General thoracic diseases included classification of anteroposterior/posteroanterior views, normal/diseases images, age prediction, or multiple diseases altogether. In lung cancer, classification was mainly for malignancy or ground-glass nodules types.

Classification approaches using deep learning methods have recorded good performance values; highly dependent on pre-processing techniques, external information involved, pooling methods used, multi-resolution models, and detection performance. Architectures were based on CNNs, sometimes comparing different architectures or applying them with certain edits. Examples on the architectures are AlexNet, GoogLeNet, DenseNet, and region attention feedback networks. Moreover, the choice of training and validation datasets was proved to be critical, similar to the impact of training dataset which are abundantly available. Classifiers are improving their generalizability to cover many diseases at once, and urgency level for treatment has become a target to predict as well.

This part presented an overview of deep learning. Then, it highlighted several surveys on deep learning application to pulmonary medical imaging analysis. Next, it summed up the contributions to the aforementioned topic, categorizing them on the task level (registration, image enhancement, segmentation, detection, feature extraction and classification). In depth of the tasks, a closer look was taken to targeted diseases, used image modalities, and basic adopted architectures. Finally, conclusions were drawn as a result of the discussions. Extensive work has been done to pulmonary image analysis, and eventually continues to be done until clinically approved.

4 Deep learning and pulmonary medical imaging analysis: COVID-19

In December 2019, Huang et al. [204] reported the first occurrences of 2019 novel coronavirus (COVID-19), which was then recognized by WHO [3] on December 31st. Later, in March of 2020, it was declared as pandemic, having infected more than 7 billion people worldwide by June 2020. As fast as that, deep learning applications were shifted from all possible domains to emphasize on the humane catastrophe.

4.1 COVID-19

It is the infectious disease caused by the most recently discovered coronavirus, which belongs to a large family of viruses that may cause illness in animals or humans. In humans, several coronaviruses are known to cause respiratory infections ranging from the common cold to more severe diseases such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS). Fever, tiredness and dry cough are the most common symptoms of COVID-19.

4.2 Medical imaging and COVID-19

Basically, the diagnosis of COVID-19 is based on polymerase chain reaction (PCR) tests, while the usage of medical imaging as a diagnostic test for COVID-19 was controversial. Most radiological societies did not recommended CT screening for COVID-19 detection, especially that it has similar features to several pneumonia types. By the end of March, Simpson et al. [205] anticipated a potential use of CT screening in clinical management and proposed four categories of standardized CT reporting language of COVID-19. Later in April, a study by Mahmood et al. [206] of 12270 patients recommended patients undergo CT screening to detect COVID-19 the earliest, in order to prohibit or avoid the speedy spread of infection.

And even though CT scans were favored over chest X-rays for COVID-19 detection, another perspective was that X-rays are cheaper, more available, and even portable (PCXR), minimizing the chances for patients to move around and spread the virus. This was the case with the classification scenarios suggested by Pereira et al. [207] for identification of COVID-19.

Portability, bedside evaluation capability, and possibility to repeat the examination during follow-up have questioned the possible diagnostic and prognostic role of lung ultrasound (LUS) and later application of deep learning to it. A study by Soldati et al. [208] stated that LUS are urgently needed and suggested a comparison with chest X-ray and/or lung CT scan to help design a diagnostic workup suitable to the technological and human resources.

On another side, the statement of the Italian Society of Medical and Interventional Radiology about CT and AI usage is suspected or COVID-19 positive patients [209] recommended chest X-rays as a first-line imaging tool, CT as an additional tool and US as a monitoring tool, prioritizing the sanitation of scan equipment after suspicion or detection of COVID-19 positive patients. Moreover, it supported the research on the use of AI as a diagnosis and prognosis decision support system excluding AI-CT scans combinations, insisting that CT scans should not be considered as first-line test to diagnose COVID-19.

A review on the role of imaging in the detection and management of COVID-19 by Dong et al. [210] revealed that characteristics of typical imaging and their changes play an important role in the detection and management of COVID-19. In addition, it showed that CT scans can improve the speed and accuracy of diagnosis and patient management, if made based on epidemiological history, nucleic acid detection, clinical symptoms, and laboratory findings. Not only that, but also predicted that the combination of AI and CT scans can offset medical resources limitations, speed up the diagnosis process and assess in the prognosis.

In all cases, multi-center studies will still be needed to ascertain the current findings.

4.3 Reviews on deep learning applications to COVID-19 medical imaging analysis

Researchers dedicated their time and resources to help fight the virus on several levels. And for sure, Artificial Intelligence had its big share. For instance, it was suggested as a detection mechanism of the cough of a COVID-19 patient [211]. Another example on deep learning specifically was a technique suggested for monitoring COVID-19 social distancing by Punn et al. [212]. No proposed system aimed to fully replace clinical tests or diagnosis; however, they were emerging as backup plans for peak hours and to assist inexperienced medical teams.

Technological strategies

Starting with technological strategies for controlling COVID-19 pandemic, a review by Elavarasan and Pugazhendhi [213] listed image analysis in the viable prospective technologies, under AI category. Ethical and legal issues are considered as significant challenges of image analysis, while future prospects include clinicians need to work closely with the AI research community.

Another approach was a review by Kumar et al. [214], that emphasized more on diagnosis using radiology images and patient's health condition prediction. Moreover, it stated contributions recording accuracies of applications of AI in CT diagnosis of COVID-19 ranging from 79.3% until 95%. A wide scope potential utilization of these technologies is concluded to cover clinical and cultural difficulties caused by coronavirus, but still need advances to meet the needed operational effect.

Artificial intelligence

Other reviews covered AI with COVID-19. Naudé [215] discussed the limitations, constraints and pitfalls of AI with COVID-19. Emphasizing on diagnosis and prognosis, using AI can speed it up, save lives, limit the spread of coronavirus and increase the training data needed to improve the algorithms. However, the potential of AI in this domain "isn't yet carried into practice." The limitations are usually the unavailability of enough training data, the selection probable bias, and the possibility of contamination of equipment upon imaging patients. Therefore, according to them, "no one this spring is going to be given a coronavirus diagnosis by an AI doctor."

Another review on AI applications by Kulkarni et al. [216] Summarized the main applications of AI in COVID-19 pandemic in early detection and diagnosis of the infection,

monitoring the treatment, contact tracing of the individuals, projection of cases and mortality, development of drugs and vaccines, and prevention of disease.

Shi et al. [217] covered the entire pipeline of AI medical imaging analysis techniques for COVID-19. For segmentation, all contributions were on CT images, segmenting either lungs, lung lobes, lung segments, lesions, trachea or bronchus. The methods used included U-Net, U-Net++, VB-Net and other commercial software. However, for diagnosis, contributions were performed on X-rays or CT scans, while methods included CNN, ResNet50, U-Net++, U-Net, or RF. The diagnosis methods accuracies ranged between 82.9% and 98%. A final note mentioned that these methods provide only little information about COVID-19 patients which make it a decision support tool rather than a first-line test.

Besides, Nguyen et al. [218] surveyed block-chain and AI-based solutions to combat COVID-19. The main future prospects deduced concerning medical imaging analysis solutions to COVID-19 were the development of AI models, and the combination of AI-based solutions with other technologies. Adaptive AI models were suggested for predictive modeling, patient monitoring, and in emergency departments. The high computational capability and resourceful storage are the cloud special features that would facilitate AI analytics if integrated. Thus, a highly advanced medical system is expected in the near future to combat the coronavirus-like epidemics.

The Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data was studied by Santosh [219]. It is basically the need to tools that can learn over time without having full knowledge about the data (Active Learning—AL), where learning is incremented as time proceeds (Incremental Learning-IL) allowing the model to adapt to new data. Moreover, it is wise—with respect to the authors—as a result of their research to use multimodal and multitudinal data with AI tools to be ready for dealing with COVID-19—like epidemics, where the variety in the data characteristics and from different populations can yield more consistent decisions.

Prediction models

Covering prediction models, a review by Wynants et al. [220] on prediction models for diagnosis and prognosis of COVID-19 stated that these models are at high-risk bias due to selection of control patients, exclusion of patients who lose interest in the study before it ends, and model overfitting. And the recommendation is for an urgent update of COVID-19 related prediction models, their validation, and their development by sharing data and expertise. Otherwise, the estimates are likely to be misleading. Besides, the prediction models identified in models included in the review

Table 9 Datasets used in training and validating deep learning algorithms

Dataset types	Percentage of papers from total		
Collected	30%		
Public	70%	LIDC or subset	28%
		ChestXray14	15%
		Shenzhen	5.8%
		JSRT	5.8%
		TCIA	4.7%
		Others	18.6%

are advised to be considered as candidate predictors for new efficient models.

4.4 Deep learning applications to COVID-19 medical imaging analysis

Contributions to deep learning applications for COVID-19 that were published in pre-prints until the end of May 2020 were excluded from the search. However, they might be included in the previous part. Table 10 lists the contributions described below.

4.4.1 Segmentation

Starting with segmentation, Butt et al. [221] proposed a system to screen coronavirus disease 2019. It used a 3D CNN to segment multiple candidate cubes from the CT scan after pre-processed. Then, the system collects the center image and the two neighbors of each cube. After then, classification takes place to categorize patches into COVID-19, Influenza-A-viral-pneumonia, and irrelevant-to-infection. To vote the patches from the same cube, the type and the confidence score are used. Finally, Noisy-or Bayesian function are used to calculate the overall analysis report.

Another segmentation system was proposed by Wang et al. [32], which is applied to chest X-rays, basically using a CNN to extract the feature map of the image along with the classification result, regression result, and the needed mask.

Murphy et al. [222] performed a multi-reader of an AI system based on chest X-rays and was for commercial use that targets the detection of tuberculosis. The system is of two parts: segmentation of lungs using U-Net and then a patch-based analysis by a CNN, followed by an ensemble of networks that aim to classify images as a final step. AUC recorded was of 0.81, and the system was comparable to six independent readers performance.

Table 10 Deep learning applications to medical imaging analysis on COVID-19

Authors	Image modality	Base network	Task
Butt et al. [221]	CT	3D CNN	Segmentation
Wang et al. [32]	CXR	CNN	Segmentation
Murphy et al. [222]	CXR	U-Net & CNN	Segmentation
Li et al. [223]	CT	U-Net & ResNet-50	Detection
Hasan et al. [224]	CT	CNN	Detection
Kassani et al. [225]	CXR & CT	DenseNet-121, ResNet-50, MobileNet, Xception, Inception, Inception-V3, Inception-ResNet-V2, VGG & NASNe	Detection
Hurt et al. [226]	CXR	U-Net	Detection
Loey et al. [227]	CXR	GoogleNet, AlexNet & ResNet-18	Detection
Apostolopoulos and Mpesiana [228]	CXR	VGG-19, MobileNet-V2, Inception, Xception & Inception-ResNet-V2	Detection
Luz et al. [229]	CXR	EfficientNet, MobileNet, MobileNet-V2, ResNet-50, VGG-16 and VGG-19.	Detection
Toğaçar et al. [230]	CXR	SqueezeNet & MobileNet-V2	Detection
Wu et al. [231]	CT	ResNet-50	Classification
Ardakani et al. [232]	CT	AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, & Xception	Classification
Ucar and Korkmaz [233]	CT	SqueezeNet	Classification
Farid et al. [234]	CT	CNN	Classification
Butt et al. [221]	CT	ResNet-23	Classification
Singh et al. [235]	CT	CNN	Classification
Pereira et al. [207]	CXR	CNN	Classification
Ozturk et al. [236]	CXR	CNN	Classification

4.4.2 Detection

As for detection, many contributions were recognized. Hurt et al. [226] proposed a localization system based on U-Net trained with 22k radiographs annotated by radiologists producing probability maps, which seemed to be generalizable and robust to be applied with patients of COVID-19.

On another side, Loey et al. [227] tested all three algorithms AlexNet, GoogleNet and ResNet-18 on three different scenarios: one containing four dataset classes, another one including three classes, and a third scenario including two classes. GoogleNet is selected as the best deep transfer model of the first scenario (80.6% testing accuracy), AlexNet is selected for the second (%85.2 testing accuracy), and GoogleNet for the third scenario (100% for testing accuracy and 99.9% for validation accuracy).

As for comparison too, Apostolopoulos and Mpesiana [228] experimented VGG-19, MobileNet-v2, Inception, Xception, and Inception-ResNet-V2 with transfer learning on Covid-19. The best confusion metrics were given by MobileNet-V2 and VGG-19.

Li et al. [223] developed a 3D deep learning tool for COVID-19 detection using CT scans. First, the 3D CT exam is pre-processed and lungs are extracted as ROI using U-Net. Then the second part of the tool is based on a ResNet-50 network that aims to generate features, that then will be max-pooled. Finally, the feature maps is fed into a fully connected layer and activated using soft-max functions to generate the probability score for COVID-19, CAP and non-pneumonic. Both local 2D features and global 3D ones are being extracted at the first place. AUC recorded is 0.96.

A COVID-19 detection system for X-rays and CT scans is proposed by Kassani et al. [225] and tested using different backbone networks. The best achieved accuracies were by DenseNet121 (99%) followed by a hybrid learner based on ResNet-50 trained by LightGBM (98%). The other included back-bone networks are MobileNet, Xception, Inception-V3, Inception-ResNet-V2, VGG and NASNe.

Luz et al. [229] also experimented different networks on chest X-ray images in order to detect COVID-19. An EfficientNet is proposed of five families that differ in input image shape. It is basically made up of convolutional layers followed by pooling layers and Mobile Inverted Bottlneck Conv self-made blocks. Its performance is then compared to MobileNet, MobileNet-V2, ResNet-50, VGG-16 and VGG-19. Results showed that EfficientNet had fewer parameters than ResNet-50 and VGG-16, with high accuracy and sensitivity (highest recorded was for EfficientNet B3: 93.9% accuracy and 96.8% sensitivity)

In addition, Hasan et al. [224] suggested a framework that uses histogram thresholding to isolate the background of CT lung scan, which then undergoes feature extraction using CNN and Q-deformed entropy algorithm. Then the features extracted are classified using long short-term memory neural network (LSTM NN). The achieved accuracy was 99.68% using the collected dataset.

Toğaçar et al. [230] also proposed a framework for COVID-19 detection using chest X-rays, starting with fuzzy color technique as a pre-processing step to restructure the data classes, then stack them with original images. After then, the stacked dataset is trained with SqueezeNet and MobileNetV2, then outputs are processed using an optimization method. SVM is finally used to classify the results achieving 98.3% accuracy.

4.4.3 Classification

As for classification, Wu et al. [231] proposed the fusion of deep learning networks for COVID-19 detection using the maximum lung regions in axial, coronal and sagittal views of CT scans. The lung regions are first segmented using threshold segmentation; then, the model based on ResNet-50 is trained yielding the three branch network output feature maps. A fully connected layer receives the output. AUC recorded 0.732, while the accuracy achieved 70% in validation test. In testing, AUC and accuracy achieved 0.819 and 76%, respectively.

Ozturk et al. [236] also proposed a framework for the same objective, based on 17 convolution layers and different filtering. The accuracy achieved for binary classification (COVID vs. No-Findings) recorded 87.02% and for multi-class (COVID vs NO-Findings-Pneumonia) recorded 98.08%.

Besides, Ardakani et al. [232] compared ten neural networks on CT scans for classification of COVID-19: AlexNet, VGG-16, VGG-19, SqueezeNet, GoogleNet, MobileNet-V2, ResNet-18, ResNet-50, ResNet-101, and Xception. ResNet-101 appeared to be of high sensitivity and could be considered in radiology department to facilitate operations.

Ucar and Korkmaz [233] proposed a framework using chest X-ray that is based on SqueezeNet tuned for coronavirus with Bayesian optimization additive. It is an easy-to-implement deep learning model which has an accuracy performance of 98.3% (among normal, pneumonia and Covid cases), and 100% for the single recognition of COVID-19 (among other classes).

Moreover, Farid et al. [234] compared the performance of proposed feature extractor and proposed stack hybrid classification using CT scans, with a CNN model. The result was reduced false-negative rate and showed a relatively high overall accuracy with more accurate results for the favor of the proposed model.

Butt et al. [221] proposed a deep learning framework for the classification of COVID-19 also. It was experiments on two levels: first where a traditional ResNet-23-based network is the back-bone network, and second where a location-

attention mechanism is concatenated in the fully connected layer. The second network's performance surpassed the first one, recording an overall accuracy of 86.7%.

CNN was used in Singh et al. [235] contribution on CT scans to classify COVID-19 infected or non-infected patients. The initial parameters were tuned using a multi-objective differential evolution (MODE). The results surpassed other models (i.e., ANN, ANFIS,…) by 1.9789% in terms of accuracy.

5 Discussion

Applications of deep learning technology are developing day after day in various domains, such as audio processing, text analysis and natural language processing, physical sciences and many other [10]. Emphasizing on its application to medical imaging analysis, there was a notable advance after 2017 as several pulmonary diseases and medical concerns were targeted from different perspectives.The noticeable advance was due to coronavirus epidemic by the end of 2019, which rang the bells for clinical trials time to start putting all the Maths into real-time enforcement.

Previous challenges

Machine learning, broader family of deep learning, has faced many challenges, being applied to medical imaging domain. Previously, de Bruijne [237] highlighted five challenges to be addressed in upcoming research; improvement of data access, making use of image modalities and data in processing pipeline, interpretation of results and application of models to clinical practice, and training of robust models with little training data. In addition, de Bruijne [237] spotted the light on future research directions including learning from weak labels, coping with variant imaging protocols and finally improving results interpretations. Learning from weak labels was a concern because the majority of algorithms in the past episode (2015–2017) employed supervised learning methods according to Ker et al. [4]. Another reason is the vast use of CNNs stated by Litjens et al. [2], which requires a large amount of relevant training data. Later, Lundervold and Lundervold [9] added to the requirements issues related to data access, privacy and data protection, especially that medical data are sensitive and mostly anonymized. Although many solutions were proposed since then, the objectives remained to have complex conditions that require further studies.

Weak labels?

Call for work on unsupervised or weak-labeled data grabbed the attention of few at the level of chest research, such as Tang et al. [179] who classified four lung diseases

with weak-supervision, and Xie et al. [202] who combined an adversarial auto-encoder-based unsupervised reconstruction network with a supervised classification one. Another approach that targeted registration using weak labels was that of de Vos et al. [83].

The need for unsupervised learning actually emerged the most with COVID-19, as in epidemics, there is no time for radiologists to dedicate for datasets building and management.

Datasets

In fact, approaches were more focused on supervised learning, which brings us to training/validation-data availability. Litjens et al. [2] clarified that the core problem was not the unavailability of imaging data, as most of western hospitals were already equipped with Picture Archiving and Communication systems (PACs). Yet, the issue was not their availability, but their structure and relevance to the training objectives. Datasets used by contributions from 2017 until now were featured in Table 1. Publicly available datasets were used by approximately 70% of the cases for training, validation or both. The other 30% of cases used privately collected institutionally authorized datasets which allows the customization of models to certain areas needs. However, 7.9% of the cases use both collected and publicly available datasets, either for comparison reasons or to avoid overfitting. LIDC and ChestXray14 achieved the highest datasets usage percentages (28% and 15%) due to the large number of images available with corresponding annotations. Other notable public datasets used were Shenzhen, JSRT and TCIA. An important dataset CheXpert [38] was published in 2019 and is of interest of many contributions. More information about most remarkable public datasets used is available in Table 9.

On another side, COVID-19's most popular dataset was published by Cohen et al. [238], who made frontal view X-rays available after collecting them from different sites and publications, and added annotations as time proceeds. CT scans were also made available by Zhao et al. [239] from 216 patients and 463 non-COVID-19 persons. This does not exclude the self-collected datasets by some aforementioned contributions.

Results interpretability and training data

Interpretability of results and model behavior can take place on many property levels according to Molnar [240]: accuracy, consistency, stability, certainty, and confidence. Accuracy and confidence interval were used in many approaches such as that of Purkayastha et al. [178] and Gong et al. [152]. To evaluate certainty, some approaches calcu-

lated "precision, recall and F-score." Others calculated "area under ROC," and sometimes approaches calculated both.

For stability, many approaches compared the performance of their models using different datasets whether for training, feature learning, validation, or all together. Examples of approaches in training, validation and all together are, respectively: Pan et al. [62], Ho et al. [133] and Gozes and Greenspan [166]. Some also experimented on different dataset sizes such as Shen et al. [192]. Performance was better when validated on internal datasets rather than external ones (using subset of dataset used for training) by Ho et al. [133]; however, there is an optimal training dataset size for each situation. Pediatrics were introduced to chest diseases diagnosis as suggested by Candemir and Antani [7]; then, the model was trained with pediatrics and adults images resulting with good performance of the deep learning classifier by Kim et al. [171]. However, the addition of adults images to training data did not increase the accuracy a lot (only increased 1.6%). Data augmentation was also used as one of the efficient pre-processing technique in many approaches to avoid over-fitting such as that of Zhao et al. [138], Setio et al. [148], Ho et al. [133]. Bone shadow exclusion such as that of Gordienko et al. [99] and normalization of pixel values like in that of Raju et al. [182] also improved classification when applied to input images as pre-processing procedures.

As expected with respect to applying deep learning to COVID-19, the more Imaging data available, the better the results are. However, the lack of enough training data has made researches move on with small available datasets, and apply augmentation when possible, ending up with a uncertain results, and a call for more dataset entries as soon as possible.

Networks architecture

Previously, end-to-end trained CNNs were seen to become the standard practice being integrated by most approaches into image analysis pipelines, replacing handcrafted machine learning methods research [2]. Similarly, their application fields expanded. Few contributions to chest were based on auto-encoders such as those of Xie et al. [202] and Nemoto et al. [163], while many were based on U-nets such as segmentation approaches of Ahn et al. [88] and Furutani et al. [103]. Likewise, other CNN architectures were used as FCNN by Dou et al. [143], ResNet by Kim et al. [171], and DenseNet by Gozes and Greenspan [166]. Moreover, combinations of architectures took place as by Gozes and Greenspan [164] where ResNet and DenseNet were integrated, and comparisons of multiple architectures were explicit as by Gozes and Greenspan [133], Ayan and Unver [131], Pan et al. [62], and Wang et al. [184]. The comparison of multiple architectures' application on same task is interpretation on consistency level. At the same level, results of experiments were compared sometimes to radiologists' diagnosis and classification, showing the competency of deep learning, such as the approach of Rajpurkar et al. [130] where model's performance was compared to 4 radiologists using only frontal radiographs and that of Ardila et al. [153] where model's performance was compared to 6 radiologists.

Multi-view CNNs were proposed by Hussein et al. [188], multi-task learning was used by Chen et al. [168], and multi-crop CNN by Shen et al. [192]. Two-dimensional views were used to form a 3D representation of a given nodule by Ciompi et al. [194], and vice versa, 3D CT scans were used to synthesize X-rays (2D) when they are unavailable by Chang and Moturu [149]. 3D CNNs were preferred over 2D for volumetric medical image analysis efficiency-wise in many approaches, one of them is that of Huang et al. [150].

Models involving CNNs outperformed traditional learning ones, yet, their combination enhanced performance of the classifier [193]. Priory information, especially demographic features, also improved the performance of deep learning models as concluded by Heo et al. [132]. Sometimes, these information were integrated through NLP systems as by Annarumma et al. [128]. On top of that, the role of handcrafted features was found, by Wang et al. [195], to be useful on the level of performance enhancement when fused with deep learned features.

Fine-tuning was put in comparison against pre-training of deep learning networks, as Lopes and Valiati [165] validated the usefulness and power of pre-trained networks refuting the saying that fine-tuning is better. On the other hand, Xu et al. [127] assured the importance of proper design rather than the fine-tuning, but still agreed that extra training enhances the performance of deep learning model.

Because contributions were vigorously developed in the 3 years past to COVID-19 pandemic, it was time by 2020 to apply all past methods and study their performance for the new virus. CNNs were basically used to segment lungs, U-Nets were used to extract features, and most of the contributions emphasized on comparing the efficiency of popular architectures like AlexNet, MobileNet, ResNet, GoogleNet, SqueezeNet, Inception, Xception, VGG, and DenseNet with their various versions and layers count. Few contributions came up with customized networks, like EfficientNet by Luz et al. [229].

New applications

New applications to medical deep learning on chest were introduced the past 4 years, such as to age prediction learning model by Karargyris et al. [175] that aims to predict the age of a patient from his medical scan, where a gap between the predicted age and the real one indicates a health concern. Another approach by Wong et al. [176] classifies the normal scans to disregard them prioritizing diagnosis of abnormal

ones. In an advanced contribution, a deep learning classifier by Annarumma et al. [128] aims to predict the urgency of cases to allow treating worse first. Databases built along years with follow-up scans of patients allow involvement of mortality risk prediction, such as the one used in the contribution of Xu et al. [199], and relating to available training data, inhale/exhale scans allow training of deep learning models that segment lungs and assess COPD diseases more efficiently as concluded by Kitahara et al. [102]. In addition, frontal chest radiographs can be classified into anteroposterior or posteroanterior views by the classifier of Kim et al. [171].

Image modalities

Segmentation applications based on deep learning methodologies were mostly on CT scans for airways diseases and lung cancer (nodules and tumors), whereas CXRs were used more for organ segmentation. On detection, lung cancer was the most approached target, more often in CT scans, but with recent approaches on CXRs. But for infections and multiple pathologies, use of CXR was dominant. On the other hand, classification of general thoracic diseases on CXR images and that of lung cancer on CT scans were of higher interest than other diseases.

When tackling the dilemma of information extracted from medical images concerning coronavirus, we notice that two parties exist: one encouraging the use of CT scans and one which does not. However, published contributions by the end of May 2020 showed an approximately equal share for each of both image modalities. Yet, CT scans were mostly used for classification, while CXRs were used for detection, feature extraction and segmentation.

6 Conclusion

As choosing the datasets is critical for training models, and over-fitting is probable especially with the varying imaging technologies and the wide variety of diseases, research is missing the third world population that might formulate different imaging features and is of more need to technological aid, especially with the present shortage of health institutions, medical staff and PACs. Unsupervised learning can thus be the solution, even though training data is becoming more relevant. It is early to have public datasets fully satisfying the needs of deep learning models, particularly for developing countries, but promising contributions can be expanded to different image modalities and populations, as in [119] which do self-supervised learning covering segmentation and classification in 2D and 3D versions can be expanded.

Many new ideas are applied to deep learning on chest, and the scope of research interest isn't just expanding to fulfill the tasks more conveniently, but going in depth of clinical needs, proving the support scope it can provide by experiments done with available datasets. Interpretability is still in debate, and Molnar [240] suggest pixel-level labeling to make better use of feature visualization and understand what is really inside the box. Many clinical applications of deep learning-based models have started to take their place, as that by Hwang et al. [241], and Foch Hospital in Suresnes, France [242]. For COVID-19 diagnosis and prognosis, it is still questionable when it will be implemented, but efforts put in this domain make it very promising, especially with it being the very vast chance to prove past work done with deep learning. Nevertheless, research direction of deep learning on chest medical imaging foresees remarkable advances and a flourishing future full of achievements especially with the unexpected coronavirus stepping into the world since the beginning of 2020.

Compliance with ethical standards

Conflict of interest Not applicable.

Availability of data Not applicable.

Code availability Not applicable.

References

1. Lo, S.C., Lou, S.L., Lin, J.S., Freedman, M.T., Chien, M.V., Mun, S.K.: Artificial convolution neural network techniques and applications for lung nodule detection. IEEE Trans. Med. Imaging **14**(4), 711–718 (1995)

2. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)

3. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/

4. Ker, J., Wang, L., Rao, J., Lim, T.: Deep learning applications in medical image analysis. IEEE Access **6**, 9375–9389 (2018)

5. Goel, N., Yadav, A., Singh, B.M.: Medical image processing: a review. In: 2016 Second International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH), pp. 57–62. IEEE (2016)

6. Zhang, G., Jiang, S., Yang, Z., Gong, L., Ma, X., Zhou, Z., Bao, C., Liu, Q.: Automatic nodule detection for lung cancer in CT images: a review. Comput. Biol. Med. **103**, 287–300 (2018)

7. Candemir, S., Antani, S.: A review on lung boundary detection in chest x-rays. Int. J. Comput. Assist. Radiol. Surg. **14**(4), 563–576 (2019). https://doi.org/10.1007/s11548-019-01917-1

8. Lee, S.M., Seo, J.B., Yun, J., Cho, Y.H., Vogel-Claussen, J., Schiebler, M.L., Gefter, W.B., van Beek, E.J., Goo, J.M., Lee,

K.S., et al.: Deep learning applications in chest radiography and computed tomography. J. Thorac. Imaging **34**(2), 75–85 (2019)

9. Lundervold, A.S., Lundervold, A.: An overview of deep learning in medical imaging focusing on MRI. Z. Med. Phys. **29**(2), 102–127 (2018)

10. Hatcher, W.G., Yu, W.: A survey of deep learning: platforms, applications and emerging research trends. IEEE Access **6**, 24411–24432 (2018)

11. Sahiner, B., Pezeshk, A., Hadjiiski, L.M., Wang, X., Drukker, K., Cha, K.H., Summers, R.M., Giger, M.L.: Deep learning in medical imaging and radiation therapy. Med. Phys. **46**(1), e1–e36 (2019)

12. Haskins, G., Kruger, U., Yan, P.: Deep learning in medical image registration: a survey. arXiv preprint arXiv:1903.02026 (2019)

13. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (2018)

14. Kamilaris, A., Prenafeta-Boldu, F.X.: Deep learning in agriculture: a survey. Comput. Electron. Agric. **147**, 70–90 (2018)

15. Zhang, C., Patras, P., Haddadi, H.: Deep learning in mobile and wireless networking: a survey. IEEE Commun. Surv. Tutor. **21**, 2224–2287 (2019)

16. Cui, L., Yang, S., Chen, F., Ming, Z., Lu, N., Qin, J.: A survey on application of machine learning for internet of things. Int. J. Mach. Learn. Cybern. **9**(8), 1399–1417 (2018)

17. Peng, L., Peng, M., Liao, B., Huang, G., Li, W., Xie, D.: The advances and challenges of deep learning application in biological big data processing. Curr. Bioinform. **13**(4), 352–359 (2018)

18. Khan, S., Yairi, T.: A review on the application of deep learning in system health management. Mech. Syst. Signal Process. **107**, 241–265 (2018)

19. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)

20. Liu, B., Chi, W., Li, X., Li, P., Liang, W., Liu, H., Wang, W., He, J.: Evolving the pulmonary nodules diagnosis from classical approaches to deep learning aided decision support: three decades development course and future prospect. arXiv preprint arXiv:1901.07858 (2019)

21. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)

22. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

23. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)

24. Shin, H.C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R.M.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging **35**(5), 1285–1298 (2016)

25. Wang, G.: A perspective on deep imaging. IEEE Access **4**, 8914–8924 (2016)

26. Miranda, E., Aryuni, M., Irwansyah, E.: A survey of medical image classification techniques. In: 2016 International Conference on Information Management and Technology (ICIMTech), pp. 56–61. IEEE (2016)

27. Merjulah, R., Chandra, J.: Segmentation technique for medical image processing: a survey. In: 2017 International Conference on Inventive Computing and Informatics (ICICI), pp. 1055–1061. IEEE (2017)

28. Erickson, B.J., Korfiatis, P., Akkus, Z., Kline, T.L.: Machine learning for medical imaging. Radiographics **37**(2), 505–515 (2017)

29. Shen, D., Wu, G., Suk, H.I.: Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. **19**(1), 221–248 (2017). https://doi.org/10.1146/annurev-bioeng-071516-044442

30. Meyer, P., Noblet, V., Mazzara, C., Lallement, A.: Survey on deep learning for radiotherapy. Comput. Biol. Med. **98**, 126–146 (2018)

31. Latif, J., Xiao, C., Imran, A., Tu, S.: Medical imaging using machine learning and deep learning algorithms: a review. In: 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pp. 1–5. IEEE (2019)

32. Wang, B., Wu, Z., Khan, Z.U., Liu, C., Zhu, M.: Deep convolutional neural network with segmentation techniques for chest x-ray analysis. In: 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1212–1216. IEEE (2019)

33. Altaf, F., Islam, S., Akhtar, N., Janjua, N.K.: Going deep in medical image analysis: concepts, methods, challenges and future directions. arXiv preprint arXiv:1902.05655 (2019)

34. Thaler, S., Menkovski, V.: The role of deep learning in improving healthcare. In: Data Science for Healthcare, pp. 75–116. Springer (2019)

35. Cheplygina, V., de Bruijne, M., Pluim, J.P.: Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med. Image Anal. **54**, 280–296 (2019)

36. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **68**(6), 394–424 (2018)

37. Luna16—grand challenge. https://luna16.grand-challenge.org/data/

38. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., et al.: Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison. Proc. AAAI Conf. Artif. Intell. **33**, 590–597 (2019)

39. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE Conference on IEEE Computer Vision and Pattern Recognition (CVPR), pp. 3462–3471 (2017)

40. Welcome to the cancer imaging archive. https://www.cancerimagingarchive.net/

41. Oasis brains. http://www.oasis-brains.org/

42. Alzheimer's disease neuroimaging initiative. http://adni.loni.usc.edu/

43. Structured analysis of the retina. http://cecas.clemson.edu/~ahoover/stare

44. Welcome to the fastmri dataset. https://fastmri.med.nyu.edu/

45. Datasets—plco—the cancer data access system. https://biometry.nci.nih.gov/cdas/datasets/plco/

46. Qin, C., Yao, D., Shi, Y., Song, Z.: Computer-aided detection in chest radiography based on artificial intelligence: a survey. Biomed. Eng. Online **17**(1), 113 (2018)

47. Demner-Fushman, D., Kohli, M.D., Rosenman, M.B., Shooshan, S.E., Rodriguez, L., Antani, S., Thoma, G.R., McDonald, C.J.: Preparing a collection of radiology examinations for distribution and retrieval. J. Am. Med. Inform. Assoc. **23**(2), 304–310 (2015)

48. Ryoo, S., Kim, H.J.: Activities of the Korean Institute of Tuberculosis. Osong Public Health Res. Perspect. **5**, S43–S49 (2014)

49. Jaeger, S., Candemir, S., Antani, S., Wang, Y.X.J., Lu, P.X., Thoma, G.: Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quant. Imaging Med. Surg. **4**(6), 475 (2014)

50. Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y., Doi, K.: Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating

characteristic analysis of radiologists' detection of pulmonary nodules. Am. J. Roentgenol. **174**(1), 71–74 (2000)

51. Van Ginneken, B., Stegmann, M.B., Loog, M.: Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Med. Image Anal. **10**(1), 19–40 (2006)

52. Anode09—grand challenge. https://anode09.grand-challenge.org/

53. Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., 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. **38**(2), 915–931 (2011)

54. Deep lesion (2018). https://nihcc.app.box.com/v/DeepLesion

55. Yan, K., Wang, X., Lu, L., Summers, R.M.: Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. J. Med. Imaging **5**(3), 036501 (2018)

56. Genetic epidemiology of COPD (copdgene) funded by the National Heart, Lung, and Blood Institute (2019). https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000179.v6.p2

57. Requesting access to mimic-cxr. https://mimic-cxr.mit.edu/about/access/

58. Castillo, R., Castillo, E., Fuentes, D., Ahmad, M., Wood, A.M., Ludwig, M.S., Guerrero, T.: A reference dataset for deformable image registration spatial accuracy evaluation using the copdgene study archive. Phys. Med. Biol. **58**(9), 2861 (2013)

59. National lung screening trial (NLST) screening—full text view. https://clinicaltrials.gov/ct2/show/NCT00047385

60. Sorensen, L., Shaker, S.B., De Bruijne, M.: Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans. Med. Imaging **29**(2), 559–569 (2010)

61. Van Ginneken, B., Armato III, S.G., de Hoop, B., van Amelsvoort-van de Vorst, S., Duindam, T., Niemeijer, M., Niemeijer, M., Murphy, K., Schilham, A., Retico, A., Fantacci, M.E., et al.: Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the anode09 study. Med. Image Anal. **14**(6), 707–722 (2010)

62. Pan, I., Agarwal, S., Merck, D.: Generalizable inter-institutional classification of abnormal chest radiographs using efficient convolutional neural networks. J. Digit. Imaging **32**, 888–896 (2019)

63. Kermany, D., Goldbaum, M.: Labeled optical coherence tomography (OCT) and chest x-ray images for classification. Mendeley Data 2 (2018)

64. Johnson, A.E., Pollard, T.J., Berkowitz, S., Greenbaum, N.R., Lungren, M.P., Deng, C.y., Mark, R.G., Horng, S.: Mimic-cxr: A large publicly available database of labeled chest radiographs. arXiv preprint arXiv:1901.07042 (2019)

65. Prior, F., Smith, K., Sharma, A., Kirby, J., Tarbox, L., Clark, K., Bennett, W., Nolan, T., Freymann, J.: The public cancer radiology imaging collections of the cancer imaging archive. Sci. Data **4**, 170124 (2017)

66. Van Ginneken, B.: Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiol. Phys. Technol. **10**(1), 23–32 (2017)

67. Claes, A.S., Clapuyt, P., Menten, R., Michoux, N., Dumitriu, D.: Performance of chest ultrasound in pediatric pneumonia. Eur. J. Radiol. **88**, 82–87 (2017)

68. Bi, W.L., Hosny, A., Schabath, M.B., Giger, M.L., Birkbak, N.J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I.F., Mak, R.H., Tamimi, R.M., Tempany, C.M., Swanton, C., Hoffmann, U., Schwartz, L.H., Gillies, R.J., Huang, R.Y., Aerts, H.J.W.L.: Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J. Clin. **69**(2), 127–157 (2019). https://doi.org/10.3322/caac.21552

69. Domingues, I., Pereira, G., Martins, P., Duarte, H., Santos, J., Abreu, P.H.: Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET. Artif. Intell. Rev. **53**, 4093–4160 (2019)

70. Pehrson, L., Nielsen, M., Ammitzbøl, Lauridsen C.: Automatic pulmonary nodule detection applying deep learning or machine learning algorithms to the LIDC-IDRI database: a systematic review. Diagnostics **9**(1), 29 (2019). https://doi.org/10.3390/diagnostics9010029

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

72. Sivaramakrishnan, R., Antani, S., Candemir, S., Xue, Z., Abuya, J., Kohli, M., Alderson, P., Thoma, G.: Comparing deep learning models for population screening using chest radiography. In: Medical Imaging 2018: Computer-Aided Diagnosis, vol. 10575. International Society for Optics and Photonics (2018)

73. Traverso, A., Torres, E.L., Fantacci, M.E., Cerello, P.: Computer-aided detection systems to improve lung cancer early diagnosis: state-of-the-art and challenges. J. Phys. Conf. Ser. **841**, 012013 (2017)

74. Wang, H., Zhou, Z., Li, Y., Chen, Z., Lu, P., Wang, W., Liu, W., Yu, L.: Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images. EJNMMI Res. **7**(1), 11 (2017). https://doi.org/10.1186/s13550-017-0260-9

75. Kim, G.B., Jung, K.H., Lee, Y., Kim, H.J., Kim, N., Jun, S., Seo, J.B., Lynch, D.A.: Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease. J. Digit. Imaging **31**(4), 415–424 (2018). https://doi.org/10.1007/s10278-017-0028-9

76. Labaki, W.W., Han, M.K.: Artificial intelligence and chest imaging. will deep learning make us smarter? Am. J. Respir. Crit. Care Med. **197**(2), 148–150 (2018)

77. Benzaquen, J., Boutros, J., Marquette, C., Delingette, H., Hofman, P.: Lung cancer screening, towards a multidimensional approach: why and how? Cancers **11**(2), 212 (2019). https://doi.org/10.3390/cancers11020212

78. Gooßen, A., Deshpande, H., Harder, T., Schwab, E., Baltruschat, I., Mabotuwana, T., Cross, N., Saalbach, A.: Deep Learning for Pneumothorax detection and localization in chest radiographs (2019). arXiv preprint arXiv:1907.07324

79. Yang, W., Chen, Y., Liu, Y., Zhong, L., Qin, G., Lu, Z., Feng, Q., Chen, W.: Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Med. Image Anal. **35**, 421–433 (2017)

80. Stephen, O., Sain, M., Maduh, U.J., Jeong, D.U.: An efficient deep learning approach to pneumonia classification in healthcare. J. Healthc. Eng. 2019, Article ID 4180949 (2019). https://doi.org/10.1155/2019/4180949

81. Lakhani, P., Sundaram, B.: Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology **284**(2), 574–582 (2017)

82. Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., Dong, D., Gevaert, O., Tian, J.: Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. Med. Image Anal. **40**, 172–183 (2017)

83. de Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Med. Image Anal. **52**, 128–143 (2019)

84. Hering, A., van Ginneken, B., Heldmann, S.: mlvirnet: multilevel variational image registration network. In: International Con-

ference on Medical Image Computing and Computer-Assisted Intervention, pp. 257–265. Springer (2019)

85. Rana, S.: A review of medical image enhancement techniques for image processing. Int. J. Curr. Eng. Technol. **5**, 1282–1286 (2011). https://doi.org/10.14741/Ijcet/22774106/5.2.2015.121

86. Umehara, K., Ota, J., Ishimaru, N., Ohno, S., Okamoto, K., Suzuki, T., Shirai, N., Ishida, T.: Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs. In: Medical Imaging 2017: Image Processing, vol. 10133. International Society for Optics and Photonics (2017)

87. Tang, J., Sun, Z., Yang, B., Lu, W., Liu, C.: Artificial neural network based noise reduction for chest PET imaging. J. Nucl. Med. **60**(supplement 1), 246–246 (2019)

88. Ahn, C., Heo, C., Kim, J.H.: Combined low-dose simulation and deep learning for CT denoising: application in ultra-low-dose chest CT. In: International Forum on Medical Imaging in Asia 2019, vol. 11050. International Society for Optics and Photonics (2019)

89. Vidya, M., Krishnan, M., Anirudh, G., Kundeti, S.R., Vijayananda, J.: Local and global transformations to improve learning of medical images applied to chest radiographs. In: Medical Imaging 2019: Image Processing, vol. 10949. International Society for Optics and Photonics (2019)

90. Zarshenas, A., Liu, J., Forti, P., Suzuki, K.: Separation of bones from soft tissue in chest radiographs: anatomy-specific orientation-frequency-specific deep neural network convolution. Med. Phys. **46**(5), 2232–2242 (2019)

91. Lee, D., Choi, S., Kim, H.J.: High quality imaging from sparsely sampled computed tomography data with deep learning and wavelet transform in various domains. Med. Phys. **46**(1), 104–115 (2019)

92. Umehara, K., Ota, J., Ishida, T.: Application of super-resolution convolutional neural network for enhancing image resolution in chest CT. J. Digit. Imaging **31**(4), 441–450 (2018)

93. Lee, S.M., Lee, J.G., Lee, G., Choe, J., Do, K.H., Kim, N., Seo, J.B.: CT image conversion among different reconstruction kernels without a sinogram by using a convolutional neural network. Korean J. Radiol. **20**(2), 295–303 (2019)

94. Lee, D., Kim, H., Choi, B., Kim, H.J.: Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure. Phys. Med. Biol. **64**(11), 115017 (2019)

95. Zhang, Y., Miao, S., Mansi, T., Liao, R.: Task driven generative modeling for unsupervised domain adaptation: application to x-ray image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 599–607. Springer (2018)

96. Dai, W., Dong, N., Wang, Z., Liang, X., Zhang, H., Xing, E.P.: Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 263–273. Springer (2018)

97. Dong, N., Kampffmeyer, M., Liang, X., Wang, Z., Dai, W., Xing, E.: Unsupervised domain adaptation for automatic estimation of cardiothoracic ratio. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 544–552. Springer (2018)

98. Gordienko, Y., Kochura, Y., Alienin, O., Rokovyi, O., Stirenko, S., Gang, P., Hui, J., Zeng, W.: Dimensionality reduction in deep learning for chest x-ray analysis of lung cancer. arXiv preprint arXiv:1801.06495 (2018)

99. Gordienko, Y., Gang, P., Hui, J., Zeng, W., Kochura, Y., Alienin, O., Rokovyi, O., Stirenko, S.: Deep learning with lung segmentation and bone shadow exclusion techniques for chest x-ray analysis of lung cancer. In: International Conference on Theory

and Applications of Fuzzy Systems and Soft Computing, pp. 638–647. Springer (2018)

100. Hooda, R., Mittal, A., Sofat, S.: Lung segmentation in chest radiographs using fully convolutional networks. Turk. J. Electr. Eng. Comput. Sci. **27**(2), 710–722 (2019)

101. Huynh, H.T., Anh, V.N.N.: A deep learning method for lung segmentation on large size chest x-ray image. In: 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF), pp. 1–5. IEEE (2019)

102. Kitahara, Y., Tanaka, R., Roth, H.R., Oda, H., Mori, K., Kasahara, K., Matsumoto, I.: Lung segmentation based on a deep learning approach for dynamic chest radiography. In: Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950. International Society for Optics and Photonics (2019)

103. Furutani, K., Hirano, Y., Kido, S.: Segmentation of lung region from chest x-ray images using u-net. In: International Forum on Medical Imaging in Asia 2019, vol. 11050. International Society for Optics and Photonics (2019)

104. Skourt, B.A., El Hassani, A., Majda, A.: Lung CT image segmentation using deep neural networks. Procedia Comput. Sci. **127**, 109–113 (2018)

105. Gerard, S.E., Reinhardt, J.M., Pulmonary lobe segmentation using a sequence of convolutional neural networks for marginal learning. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1207–1211. IEEE (2019)

106. Wang, W., Chen, J., Zhao, J., Chi, Y., Xie, X., Zhang, L., Hua, X.: Automated segmentation of pulmonary lobes using coordination-guided deep neural networks. arXiv preprint arXiv:1904.09106 (2019)

107. Cui, H., Liu, X., Huang, N.: Pulmonary vessel segmentation based on orthogonal fused u-net++ of chest CT images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 293–300. Springer (2019)

108. Yun, J., Park, J., Yu, D., Yi, J., Lee, M., Park, H.J., Lee, J.G., Seo, J.B., Kim, N.: Improvement of fully automated airway segmentation on volumetric computed tomographic images using a 2.5 dimensional convolutional neural net. Med. Image Anal. **51**, 13–20 (2019)

109. Nadeem, S.A., Hoffman, E.A., Saha, P.K.: A fully automated CT-based airway segmentation algorithm using deep learning and topological leakage detection and branch augmentation approaches. In: Medical Imaging 2019: Image Processing, vol. 10949. International Society for Optics and Photonics (2019)

110. Qin, Y., Chen, M., Zheng, H., Gu, Y., Shen, M., Yang, J., Huang, X., Zhu, Y.M., Yang, G.Z.: Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 212–220. Springer (2019)

111. Zhao, T., Yin, Z., Wang, J., Gao, D., Chen, Y., Mao, Y.: Bronchus segmentation and classification by neural networks and linear programming. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 230–239. Springer (2019)

112. Wang, C., Hayashi, Y., Oda, M., Itoh, H., Kitasaka, T., Frangi, A.F., Mori, K.: Tubular structure segmentation using spatial fully connected network with radial distance loss for 3D medical images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 348–356. Springer (2019)

113. Wang, W., Lu, Y., Wu, B., Chen, T., Chen, D.Z., Wu, J.: Deep active self-paced learning for accurate pulmonary nodule segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 723–731. Springer (2018)

114. Jin, D., Xu, Z., Tang, Y., Harrison, A.P., Mollura, D.J.: Ct-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 732–740. Springer (2018)

115. Mobiny, A., Van Nguyen, H.: Fast capsnet for lung cancer screening. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 741–749. Springer (2018)

116. Hu, S., Worrall, D., Knegt, S., Veeling, B., Huisman, H., Welling, M.: Supervised uncertainty quantification for segmentation with multiple annotations. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 137–145. Springer (2019)

117. Wang, D., Li, M., Ben-Shlomo, N., Corrales, C.E., Cheng, Y., Zhang, T., Jayender, J.: Mixed-supervised dual-network for medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 192–200. Springer (2019)

118. Tang, H., Zhang, C., Xie, X.: Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 266–274. Springer (2019)

119. Zhou, Z., Sodha, V., Siddiquee, M.M.R., Feng, R., Tajbakhsh, N., Gotway, M.B., Liang, J.: Models genesis: Generic autodidactic models for 3D medical image analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 384–393. Springer (2019)

120. Moriya, T., Oda, H., Mitarai, M., Nakamura, S., Roth, H.R., Oda, M., Mori, K.: Unsupervised segmentation of micro-CT images of lung cancer specimen using deep generative models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 240–248. Springer (2019)

121. Jiang, J., Hu, Y.C., Tyagi, N., Zhang, P., Rimner, A., Mageras, G.S., Deasy, J.O., Veeraraghavan, H.: Tumor-aware, adversarial domain adaptation from Ct to MRI for lung cancer segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 777–785. Springer (2018)

122. Jue, J., Jason, H., Neelam, T., Andreas, R., Sean, B.L., Joseph, D.O., Harini, V.: Integrating cross-modality hallucinated mri with CT to aid mediastinal lung tumor segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 221–229. Springer (2019)

123. Astaraki, M., Toma-Dasu, I., Smedby, Ö., Wang, C.: Normal appearance autoencoder for lung cancer detection and segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 249–256. Springer (2019)

124. Rajpurkar, P., Irvin, J., Ball, R.L., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C.P., et al.: Deep learning for chest radiograph diagnosis: a retrospective comparison of the chexnext algorithm to practicing radiologists. PLoS Med. **15**(11), e1002686 (2018)

125. Singh R., Kalra M.K., Nitiwarangkul C., Patti J.A., Homayounieh F., Padole A., Rao P., Putha P., Muse V.V., Sharma A., et al.: Deep learning in chest radiography: detection of findings and presence of change. PLoS ONE . 10/4/2018, Vol. 13, Issue 10, p1-12. 12p. https://doi.org/10.1371/journal.pone.0204155

126. Cai, J., Lu, L., Harrison, A.P., Shi, X., Chen, P., Yang, L.: Iterative attention mining for weakly supervised thoracic disease pattern localization in chest x-rays. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 589–598. Springer (2018)

127. Xu, S., Wu, H., Bie, R.: Cxnet-m1: anomaly detection on chest x-rays with image-based deep learning. IEEE Access **7**, 4466–4477 (2019)

128. Annarumma, M., Withey, S.J., Bakewell, R.J., Pesce, E., Goh, V., Montana, G.: Automated triaging of adult chest radiographs with deep artificial neural networks. Radiology **291**(1), 196–202 (2019)

129. Gerard, S.E., Patton, T.J., Christensen, G.E., Bayouth, J.E., Reinhardt, J.M.: Fissurenet: a deep learning approach for pulmonary fissure detection in CT images. IEEE Trans. Med. Imaging **38**(1), 156–166 (2019)

130. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al.: Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)

131. Ayan, E., Unver, H.M.: Diagnosis of pneumonia from chest x-ray images using deep learning. In: 2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT), pp. 1–5. IEEE (2019)

132. Heo, S.J., Kim, Y., Yun, S., Lim, S.S., Kim, J.K., Nam, C.M.N., Park, E.C., Jung, I., Yoon, J.H.: Deep learning algorithms with demographic information help to detect tuberculosis in chest radiographs in annualworkers' health examination data. Int. J. Environ. Res. Public Health **16**(2), 250 (2019)

133. Ho, T.K.K., Gwak, J., Prakash, O., Song, J.I., Park, C.M.: Utilizing pretrained deep learning models for automated pulmonary tuberculosis detection using chest radiography. In: Asian Conference on Intelligent Information and Database Systems, pp. 395–403. Springer (2019)

134. Lin, Y., Su, J., Wang, X., Li, X., Liu, J., Cheng, K.T., Yang, X.: Automated pulmonary embolism detection from CTPA images using an end-to-end convolutional neural network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 280–288. Springer (2019)

135. Taylor, A.G., Mielke, C., Mongan, J.: Automated detection of moderate and large pneumothorax on frontal chest x-rays using deep convolutional neural networks: a retrospective study. PLoS Med. **15**(11), 1–15 (2018)

136. Park, S., Lee, S.M., Kim, N., Choe, J., Cho, Y., Do, K.H., Seo, J.B.: Application of deep learning-based computer-aided detection system: detecting pneumothorax on chest radiograph after biopsy. Eur. Radiol. **29**(10), 5341–5348 (2019)

137. Nam, J.G., Park, S., Hwang, E.J., Lee, J.H., Jin, K.N., Lim, K.Y., Vu, T.H., Sohn, J.H., Hwang, S., Goo, J.M., et al.: Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. Radiology **290**(1), 218–228 (2019)

138. Zhao, W., Yang, J., Ni, B., Bi, D., Sun, Y., Xu, M., Zhu, X., Li, C., Jin, L., Gao, P. et al.: Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. Cancer Med. (2019)

139. Hamidian, S., Sahiner, B., Petrick, N., Pezeshk, A.: 3D convolutional neural network for automatic detection of lung nodules in chest CT. In: Medical Imaging 2017: Computer-Aided Diagnosis, vol. 10134. International Society for Optics and Photonics (2017)

140. Cha, M.J., Chung, M.J., Lee, J.H., Lee, K.S.: Performance of deep learning model in detecting operable lung cancer with chest radiographs. J. Thorac. Imaging **34**, 86–91 (2019)

141. Jiang, H., Ma, H., Qian, W., Gao, M., Li, Y.: An automatic detection system of lung nodule based on multigroup patch-based deep learning network. IEEE J. Biomed. Health Inform. **22**(4), 1227–1237 (2018)

142. Masood, A., Sheng, B., Li, P., Hou, X., Wei, X., Qin, J., Feng, D.: Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images. J. Biomed. Inform. **79**, 117–128 (2018)

143. Dou, Q., Chen, H., Jin, Y., Lin, H., Qin, J., Heng, P.A.: Automated pulmonary nodule detection via 3D convnets with online sample filtering and hybrid-loss residual learning. In: International Con-

ference on Medical Image Computing and Computer-Assisted Intervention, pp. 630–638. Springer (2017)

144. Kuan, K., Ravaut, M., Manek, G., Chen, H., Lin, J., Nazir, B., Chen, C., Howe, T.C., Zeng, Z., Chandrasekhar, V.: Deep learning for lung cancer detection: tackling the kaggle data science bowl 2017 challenge. arXiv preprint arXiv:1705.09435 (2017)

145. Pesce, E., Ypsilantis, P.P., Withey, S., Bakewell, R., Goh, V., Montana, G.: Learning to detect chest radiographs containing lung nodules using visual attention networks. arXiv preprint arXiv:1712.00996 (2017)

146. Pesce, E., Withey, S.J., Ypsilantis, P.P., Bakewell, R., Goh, V., Montana, G.: Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. Med. Image Anal. **53**, 26–38 (2019)

147. Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P.A.: Multilevel contextual 3-d CNNs for false positive reduction in pulmonary nodule detection. IEEE Trans. Biomed. Eng. **64**(7), 1558–1567 (2017)

148. Setio, A.A.A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., Van Riel, S.J., Wille, M.M.W., Naqibullah, M., Sanchez, C.I., van Ginneken, B.: Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans. Med. Imaging **35**(5), 1160–1169 (2016)

149. Tiddens, H.A.: Detecting early structural lung damage in cystic fibrosis. Pediatric Pulmonology **34**(3), 228–231 (2002)

150. Huang, X., Shan, J., Vaidya, V.: Lung nodule detection in CT using 3D convolutional neural networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 379–383. IEEE (2017)

151. Gu, Y., Lu, X., Yang, L., Zhang, B., Yu, D., Zhao, Y., Gao, L., Wu, L., Zhou, T.: Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs. Comput. Biol. Med. **103**, 220–231 (2018)

152. Gong, H., Walther, A., Hu, Q., Koo, C.W., Takahashi, E.A., Levin, D.L., Johnson, T.F., Hora, M.J., Leng, S., Fletcher, J.G., et al.: Correlation between a deep-learning-based model observer and human observer for a realistic lung nodule localization task in chest CT. In: Medical Imaging 2019: Image Perception, Observer Performance, and Technology Assessment, vol. 10952. International Society for Optics and Photonics (2019)

153. Ardila, D., Kiraly, A.P., Bharadwaj, S., Choi, B., Reicher, J.J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., et al.: End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat. Med. **25**(6), 954 (2019)

154. Wang, J., Gao, R., Huo, Y., Bao, S., Xiong, Y., Antic, S.L., Osterman, T.J., Massion, P.P., Landman, B.A.: Lung cancer detection using co-learning from chest CT images and clinical demographics. In: Medical Imaging 2019: Image Processing, vol. 10949. International Society for Optics and Photonics (2019)

155. Wang, B., Qi, G., Tang, S., Zhang, L., Deng, L., Zhang, Y.: Automated pulmonary nodule detection: high sensitivity with few candidates. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 759–767. Springer (2018)

156. Khosravan, N., Bagci, U.: S4nd: single-shot single-scale lung nodule detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 794–802. Springer (2018)

157. Zhu, W., Vang, Y.S., Huang, Y., Xie, X.: Deepem: deep 3D convnets with EM for weakly supervised pulmonary nodule detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 812–820. Springer (2018)

158. Yan, K., Bagheri, M., Summers, R.M.: 3D context enhanced region-based convolutional neural network for end-to-end lesion detection. In: International Conference on Medical Image

Computing and Computer-Assisted Intervention, pp. 511–519. Springer (2018)

159. Bhatia, S., Sinha, Y., Goel, L.: Lung cancer detection: a deep learning approach. In: Soft Computing for Problem Solving, pp. 699–705. Springer (2019)

160. Winkels, M., Cohen, T.S.: Pulmonary nodule detection in CT scans with equivariant CNNs. Med. Image Anal. **55**, 15–26 (2019)

161. Zhang, Z., Yang, J., Zhao, J.: An automatic detection model of pulmonary nodules based on deep belief network. Int. J. Wirel. Mobile Comput. **16**(1), 7–13 (2019)

162. Gonzalez, G., Ash, S.Y., Vegas-Sanchez-Ferrero, G., Onieva Onieva, J., Rahaghi, F.N., Ross, J.C., Diaz, A., San Jose Estepar, R., Washko, G.R.: Disease staging and prognosis in smokers using deep learning in chest computed tomography. Am. J. Respir. Crit. Care Med. **197**(2), 193–203 (2018)

163. Nemoto, M., Ushifusa, K., Tanaka, A., Yamada, T., Kimura, Y., Hayashi, N.: Pilot study to generate image features by deep autoencoder for computer-aided detection systems. In: International Forum on Medical Imaging in Asia 2019, vol. 11050. International Society for Optics and Photonics (2019)

164. Chen, B., Li, J., Guo, X., Lu, G.: Dualchexnet: dual asymmetric feature learning for thoracic disease classification in chest x-rays. Biomed. Signal Process. Control **53**, 101554 (2019)

165. Lopes, U., Valiati, J.F.: Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. Comput. Biol. Med. **89**, 135–143 (2017)

166. Gozes, O., Greenspan, H.: Deep feature learning from a hospital-scale chest x-ray dataset with application to TB detection on a small-scale dataset. arXiv preprint arXiv:1906.00768 (2019)

167. Liang, G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Comput. Methods Prog. Biomed. **187**, 104964 (2019)

168. Chen, S., Qin, J., Ji, X., Lei, B., Wang, T., Ni, D., Cheng, J.Z.: Automatic scoring of multiple semantic attributes with multi-task feature leverage: a study on pulmonary nodules in CT images. IEEE Trans. Med. Imaging **36**(3), 802–814 (2017)

169. Abiyev, R.H., Ma'aitah, M.K.S.: Deep convolutional neural networks for chest diseases detection. J. Healthc. Eng 2018, Article ID 4168538. https://doi.org/10.1155/2018/4168538 (2018)

170. Hwang, E.J., Park, S., Jin, K.N., Im, Kim J., Choi, S.Y., Lee, J.H., Goo, J.M., Aum, J., Yim, J.J., Cohen, J.G., et al.: Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. JAMA Netw. Open **2**(3), e191095–e191095 (2019)

171. Kim, T.K., Paul, H.Y., Wei, J., Shin, J.W., Hager, G., Hui, F.K., Sair, H.I., Lin, C.T.: Deep learning method for automated classification of anteroposterior and posteroanterior chest radiographs. J. Digit. Imaging **32**, 925–930 (2019)

172. Tang, Y., Tang, Y., Han, M., Xiao, J., Summers, R.M.: Abnormal chest x-ray identification with generative adversarial one-class classifier. arXiv preprint arXiv:1903.02040 (2019)

173. Ogawa, R., Kido, T., Mochizuki, T.: Effect of augmented datasets on deep convolutional neural networks applied to chest radiographs. Clin. Radiol. **74**, 697–701 (2019)

174. Calli, E., Sogancioglu, E., Scholten, E.T., Murphy, K., van Ginneken, B.: Handling label noise through model confidence and uncertainty: application to chest radiograph classification. In: Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950. International Society for Optics and Photonics (2019)

175. Karargyris, A., Kashyap, S., Wu, J.T., Sharma, A., Moradi, M., Syeda-Mahmood, T.: Age prediction using a large chest x-ray dataset. In: Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950. International Society for Optics and Photonics (2019)

176. Wong, K.C., Moradi, M., Wu, J., Syeda-Mahmood, T.: Identifying disease-free chest x-ray images with deep transfer learning. In:

Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950. International Society for Optics and Photonics (2019)

177. Ma, C., Wang, H., Hoi, S.C.: Multi-label thoracic disease image classification with cross-attention networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 730–738. Springer (2019)

178. Purkayastha, S., Buddi, S.B., Nuthakki, S., Yadav, B., Gichoya, J.W.: Evaluating the implementation of deep learning in librehealth radiology on chest x-rays. In: Science and Information Conference, pp. 648–657. Springer (2019)

179. Tang, R., Tushar, F.I., Han, S., Hou, R., Rubin, G.D., Lo, J.Y.: Classification of chest ct using case-level weak supervision. In: Medical Imaging 2019: Computer-Aided Diagnosis; vol. 10950. International Society for Optics and Photonics (2019)

180. Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H.C., Roth, H., Papadakis, G.Z., Depeursinge, A., Summers, R.M., et al.: Holistic classification of ct attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Comput. Methods Biomech. Biomed. Eng. Imaging Vis. **6**(1), 1–6 (2018)

181. Kim, G.B., Jung, K.H., Lee, Y., Kim, H.J., Kim, N., Jun, S., Seo, J.B., Lynch, D.A.: Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease. J. Digit. Imaging **31**(4), 415–424 (2018)

182. Hogeweg, L., Sánchez, C.I., Maduskar. P., Philipsen, R., Story, A., Dawson, R., Theron, G., Dheda, K., Peters-Bax, L., Van Ginneken, B.: Automatic detection of tuberculosis in chest radiographs using a combination of textural, focal, and shape abnormality analysis. IEEE Trans. Med. Imaging **34**(12), 2429–2442 (2015)

183. Zech, J.R., Badgeley, M.A., Liu, M., Costa, A.B., Titano, J.J., Oermann, E.K.: Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med. **15**(11), e1002683 (2018)

184. Kato, S., Nakamoto, T., Iizuka, M.: Early diagnosis and estimation of pulmonary congestion and edema in patients with left-sided heart diseases from histogram of pulmonary CT number. Chest **109**(6), 1439–1445 (1996)

185. Zucker, E.J., Barnes, Z.A., Lungren, M.P., Shpanskaya, Y., Seekins, J.M., Halabi, S.S., Larson, D.B.: Deep learning to automate Brasfield chest radiographic scoring for cystic fibrosis. **19**(1), 131–138 (2020)

186. Wang, Y., Sun, L.L., Jin, Q.: Enhanced diagnosis of pneumothorax with an improved real-time augmentation for imbalanced chest x-rays data based on dcnn. IEEE/ACM Trans. Comput. Biol. Bioinform. (2019)

187. Takemiya, R., Kido, S., Hirano, Y., Mabu, S.: Detection of pulmonary nodules on chest x-ray images using R-CNN. In: International Forum on Medical Imaging in Asia 2019, vol. 11050. International Society for Optics and Photonics (2019)

188. Hussein, S., Gillies, R., Cao, K., Song, Q., Bagci, U.: Tumor-net: lung nodule characterization using multi-view convolutional neural network with Gaussian process. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 1007–1010. IEEE (2017)

189. Luckehe, D., von Voigt, G.: Evolutionary image simplification for lung nodule classification with convolutional neural networks. Int. J. Comput. Assist. Radiol. Surg. **13**, 1499–1513 (2018)

190. Shen, S., Han, S.X., Aberle, D.R., Bui, A.A., Hsu, W.: An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. Expert Syst. Appl. **128**, 84–95 (2019)

191. da Silva, G.L.F., Valente, T.L.A., Silva, A.C., de Paiva, A.C., Gattass, M.: Convolutional neural network-based pso for lung nodule false positive reduction on ct images. Comput. Methods Programs Biomed. **162**, 109–118 (2018)

192. Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., Zang, Y., Tian, J.: Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recogn. **61**, 663–673 (2017)

193. Liu, S., Xie, Y., Jirapatnakul, A., Reeves, A.P.: Pulmonary nodule classification in lung cancer screening with three-dimensional convolutional neural networks. J. Med. Imaging **4**(4), 041308 (2017)

194. Ciompi, F., Chung, K., Van Riel, S.J., Setio, A.A.A., Gerke, P.K., Jacobs, C., Scholten, E.T., Schaefer-Prokop, C., Wille, M.M., Marchianò, A., et al.: Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci. Rep. **7**, 46479 (2017)

195. Wang, C., Elazab, A., Wu, J., Hu, Q.: Lung nodule classification using deep feature fusion in chest radiography. Comput. Med. Imaging Graph. **57**, 10–18 (2017)

196. Liao, F., Liang, M., Li, Z., Hu, X., Song, S.: Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network. IEEE Trans. Neural Netw. Learn. Syst. **30**, 3484–3495 (2019)

197. Hussein, S., Cao, K., Song, Q., Bagci, U.: Risk stratification of lung nodules using 3d CNN-based multi-task learning. In: International Conference on Information Processing in Medical Imaging, pp. 249–260. Springer (2017)

198. Khosravan, N., Celik, H., Turkbey, B., Jones, E.C., Wood, B., Bagci, U.: A collaborative computer aided diagnosis (C-CAD) system with eye-tracking, sparse attentional model, and deep learning. Med. Image Anal. **51**, 101–115 (2019)

199. Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., Mak, R.H., Aerts, H.J.: Deep learning predicts lung cancer treatment response from serial medical imaging. Clin. Cancer Res. **25**(11), 3266–3275 (2019)

200. Byun, S.H., Jung, J., Hong, H., Song, Y.S., Kim, H., Park, C.M.: Ground-glass nodule classification with multiple 2.5-dimensional deep convolutional neural networks in chest CT images. In: International Forum on Medical Imaging in Asia 2019, vol. 11050. International Society for Optics and Photonics (2019)

201. Srivastava, V., Purwar, R.K.: Classification of CT scan images of lungs using deep convolutional neural network with external shape-based features. J. Digit. Imaging **33**, 252–261 (2019)

202. Xie, Y., Zhang, J., Xia, Y.: Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT. Med. Image Anal. **57**, 237–248 (2019)

203. Song, Q., Zhao, L., Luo, X., Dou, X.: Using deep learning for classification of lung nodules on computed tomography images. J. Healthc. Eng. 2017 (2017)

204. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet **395**(10223), 497–506 (2020)

205. Simpson, S., Kay, F.U., Abbara, S., Bhalla, S., Chung, J.H., Chung, M., Henry, T.S., Kanne, J.P., Kligerman, S., Ko, J.P., et al.: Radiological Society of North America Expert Consensus Statement on reporting chest CT findings related to COVID-19. endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. Radiol. Cardiothorac. Imaging **2**(2), e200152 (2020)

206. Mahmood, A., Gajula, C., Gajula, P., et al.: Covid 19 diagnostic tests: a study of 12,270 patients to determine which test offers the most beneficial results. Surg. Sci. **11**(04), 82 (2020)

207. Pereira, R.M., Bertolini, D., Teixeira, L.O., Silla Jr., C.N., Costa, Y.M.: Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. Comput. Methods Prog. Biomed. 105532 (2020)

208. Soldati, G., Smargiassi, A., Inchingolo, R., Buonsenso, D., Perrone, T., Briganti, D.F., Perlini, S., Torri, E., Mariani, A., Mossolani, E.E., et al.: Is there a role for lung ultrasound during the COVID-19 pandemic? J. Ultrasound Med. (2020)

209. Neri, E., Miele, V., Coppola, F., Grassi, R.: Use of CT and artificial intelligence in suspected or Covid-19 positive patients: statement of the Italian Society of Medical and Interventional Radiology. La Radiol. Med. **125**, 505–508 (2020)

210. Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F., et al.: The role of imaging in the detection and management of COVID-19: a review. IEEE Rev. Biomed. Eng. (2020)

211. Imran, A., Posokhova, I., Qureshi, H.N., Masood, U., Riaz, S., Ali, K., John, C.N., Nabeel, M.: Ai4covid-19: Ai enabled preliminary diagnosis for Covid-19 from cough samples via an app. arXiv preprint arXiv:2004.01275 (2020)

212. Punn, N.S., Sonbhadra, S.K., Agarwal, S.: Monitoring covid-19 social distancing with person detection and tracking via fine-tuned yolo v3 and deepsort techniques. arXiv preprint arXiv:2005.01385 (2020)

213. Elavarasan, R.M., Pugazhendhi, R.: Restructured society and environment: a review on potential technologica strategies to control the COVID-19 pandemic. Sci. Total Environ. 138858 (2020)

214. Kumar, A., Gupta, P.K., Srivastava, A.: A review of modern technologies for tackling covid-19 pandemic. Diabetes Metab. Syndrome Clin. Res. Rev. **14**, 569–573 (2020)

215. Naudé, W.: Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. AI Soc 1–5 (2020)

216. Kulkarni, S., Seneviratne, N., Baig, M.S., Khan, A.H.A.: Artificial intelligence in medicine: where are we now? Acad. Radiol. **27**(1), 62–70 (2020)

217. Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y., Shen, D.: Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. IEEE Rev. Biomed. Eng (2020)

218. Nguyen, D., Ding, M., Pathirana, P.N., Seneviratne, A.: Blockchain and AI-based solutions to combat coronavirus (COVID-19)-like epidemics: a survey (2020)

219. Santosh, K.: AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multitudinal/multimodal data. J. Med. Syst. **44**(5), 1–5 (2020)

220. Wynants, L., Van Calster, B., Bonten, M.M., Collins, G.S., Debray, T.P., De Vos, M., Haller, M.C., Heinze, G., Moons, K.G., Riley, R.D., et al.: Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal . BMJ 369:m1328. https://doi.org/10.1136/bmj.m1328 (2020)

221. Butt, C., Gill, J., Chun, D., Babu, B.A.: Deep learning system to screen coronavirus disease 2019 pneumonia. Appl. Intell. **2020**, 1 (2019)

222. Murphy, K., Smits, H., Knoops, A.J., Korst, M.B., Samson, T., Scholten, E.T., Schalekamp, S., Schaefer-Prokop, C.M., Philipsen, R.H., Meijers, A., et al.: COVID-19 on the Chest Radiograph: A Multi-Reader Evaluation of an AI System. Radiology 201874 (2020)

223. Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., et al.: Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology 200905 (2020)

224. Hasan, A.M., AL-Jawad, M.M., Jalab, H.A., Shaiba, H., Ibrahim, R.W., AL-Shamasneh, A.R.: Classification of Covid-19 coronavirus, pneumonia and healthy lungs in CT scans using q-deformed entropy and deep learning features. Entropy **22**(5), 517 (2020)

225. Kassani, S.H., Kassasni, P.H., Wesolowski, M.J., Schneider, K.A., Deters, R.: Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: A machine learning-based approach. arXiv preprint arXiv:2004.10641 (2020)

226. Hurt, B., Kligerman, S., Hsiao, A.: Deep learning localization of pneumonia: 2019 coronavirus (Covid-19) outbreak. J. Thorac. Imaging **35**(3), W87–W89 (2020)

227. Loey, M., Smarandache, F., M Khalifa, N.E.: Within the lack of chest Covid-19 x-ray dataset: a novel detection model based on gan and deep transfer learning. Symmetry **12**(4), 651 (2020)

228. Apostolopoulos, I.D., Mpesiana, T.A.: Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys. Eng. Sci. Med. **43**, 635–640 (2020)

229. Luz, E., Silva, P.L., Silva, R., Moreira, G.: Towards an efficient deep learning model for covid-19 patterns detection in x-ray images. arXiv preprint arXiv:2004.05717 (2020)

230. Toğaçar, M., Ergen, B., Cömert, Z.: Covid-19 detection using deep learning models to exploit social mimic optimization and structured chest x-ray images using fuzzy color and stacking approaches. Comput. Biol. Med. 103805 (2020)

231. Wu, X., Hui, H., Niu, M., Li, L., Wang, L., He, B., Yang, X., Li, L., Li, H., Tian, J., et al.: Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. Eur. J. Radiol. 109041 (2020)

232. Ardakani, A.A., Kanafi, A.R., Acharya, U.R., Khadem, N., Mohammadi, A.: Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. Comput. Biol. Med. 103795 (2020)

233. Ucar F., Korkmaz D., Covidiagnosis-net: Deep bayes-squeezenet based diagnostic of the coronavirus disease: (covid-19) from x-ray images. Med. Hypotheses **2020**, 109761 (2019)

234. Farid, A.A., Selim, G.I., Awad, H., Khater, A.: A novel approach of CT images feature analysis and prediction to screen for corona virus disease (COVID-19). Int. J. Sci. Eng. Res. **11**(3), 1–9 (2020)

235. Singh, D., Kumar, V., Kaur, M.: Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. Eur. J. Clin. Microbiol. Infect. Dis. **39**, 1379–1389 (2020)

236. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of COVID-19 cases using deep neural networks with x-ray images. Comput. Biol. Med. **121**, 103792 (2020)

237. de Bruijne, M.: Machine learning approaches in medical image analysis: from detection to diagnosis. Med. Image Anal. **33**, 94–37 (2016)

238. Cohen, J.P., Morrison, P., Dao, L.: Covid-19 image data collection. arXiv preprint arXiv:2003.11597 (2020)

239. Zhao, J., Zhang, Y., He, X., Xie, P.: Covid-ct-dataset: a ct scan dataset about Covid-19. arXiv preprint arXiv:2003.13865 (2020)

240. Molnar, C.: Interpretable machine learning (2019). https://christophm.github.io/interpretable-ml-book/

241. Hwang, E.J., Nam, J.G., Lim, W.H., Park, S.J., Jeong, Y.S., Kang, J.H., Hong, E.K., Kim, T.M., Goo, J.M., Park, S., et al.: Deep learning for chest radiograph diagnosis in the emergency department. Radiology **293**(3), 573–580 (2019b)

242. Vitard, A.: L'hôpital foch mise sur l'intelligence artificielle pour créer des radiologues augmentés (2020). https://www.usine-digitale.fr/article/l-hopital-foch-mise-sur-l-intelligence-artificielle-pour-creer-des-radiologues-augmentes.N923009

Hanan Farhat received her BS and ME in computers and telecommunications engineering from Lebanese International University, Khyara, Lebanon (with honor), in 2014 and 2016, respectively. Then, she received the MS degree in networking and telecommunications from the

joint program between Lebanese University and St. Joseph University of Beirut in 2017. She is now a PhD fellow at the faculty of electrical and electronics engineering at the Saint Joseph University of Beirut (CIMTI-ESIB-USJ) where she works on applying machine learning in medical domain. Her research is funded by the joint support of the National Center of Scientific Research, Beirut, and St. Joseph University of Beirut. Hanan is a former Labs Instructor at Lebanese International University (2016-2020). She taught programming and networking labs. Link to Research gate: https://www.researchgate.net/profile/Hanan_Farhat2

George E. Sakr   received the BE degree (with distinction) in electrical and electronics engineering from Lebanese University in 2005 and the PhD degree in Machine Learning from the American University of Beirut in 2011. Between 2011 and 2015, George was a Lecturer at the American university of Beirut and the Lebanese American University as well as a post-doctoral fellow at the Medical Center of the American University of Beirut where he worked on applying machine learning for precision medicine. Since 2015, George has been an assistant professor at St Joseph University of Beirut. His research interests include artificial intelligence, machine learning, human machine interfacing, and application of machine learning for precision medicine. He is a senior member of the IEEE. He has more than 30 journal and conference publications in the domain of machine learning. Link to research gate: https://www.researchgate.net/profile/George_Sakr

Rima Kilany   is an Associate Professor at the "Ecole Supérieure des Ingénieurs de Beyrouth" of Saint-Jospeh University, Lebanon, since 1994. She pursued her PhD degree in Information Systems in Telecom ParisTech, Paris, France. Research interests: Semantic web, Integration Teaching activity Design Patterns, Software Engineering, Data Integration, Middleware, Enterprise Integration Link to research gate (publications, reads and citations): https://www.researchgate.net/profile/Rima_Kilany/2