Computers in Biology and Medicine 146 (2022) 105691

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Lung nodule diagnosis and cancer histology classification from computed tomography data by convolutional neural networks: A survey☆

Selene Tomassini, Nicola Falcionelli, Paolo Sernani, Laura Burattini *, Aldo Franco Dragoni

Department of Information Engineering, Engineering Faculty, Università Politecnica delle Marche, Ancona, Italy

ARTICLE INFO

Keywords:
Computed tomography
Convolutional neural network
Deep learning
Lung cancer histology classification
Lung nodule diagnosis
Supervised learning

ABSTRACT

Lung cancer is among the deadliest cancers. Besides lung nodule classification and diagnosis, developing non-invasive systems to classify lung cancer histological types/subtypes may help clinicians to make targeted treatment decisions timely, having a positive impact on patients' comfort and survival rate. As convolutional neural networks have proven to be responsible for the significant improvement of the accuracy in lung cancer diagnosis, with this survey we intend to: show the contribution of convolutional neural networks not only in identifying malignant lung nodules but also in classifying lung cancer histological types/subtypes directly from computed tomography data; point out the strengths and weaknesses of slice-based and scan-based approaches employing convolutional neural networks; and highlight the challenges and prospective solutions to successfully apply convolutional neural networks for such classification tasks. To this aim, we conducted a comprehensive analysis of relevant Scopus-indexed studies involved in lung nodule diagnosis and cancer histology classification up to January 2022, dividing the investigation in convolutional neural network-based approaches fed with planar or volumetric computed tomography data. Despite the application of convolutional neural networks in lung nodule diagnosis and cancer histology classification is a valid strategy, some challenges raised, mainly including the lack of publicly-accessible annotated data, together with the lack of reproducibility and clinical interpretability. We believe that this survey will be helpful for future studies involved in lung nodule diagnosis and cancer histology classification prior to lung biopsy by means of convolutional neural networks.

1. Introduction

Cancer is the principal cause of death in the world [1]. Lung cancer is the most common and among the deadliest cancers [2,3], killing more people than bladder, brain, breast, colorectal, prostate, and stomach cancers [4–6]. Indeed, it accounts for about 1.8 million new cases and more than 1.4 million deaths every year in the world [1,7]. Nodule visual appearance is mighty varied with subtle peculiarities in shape, texture, and size [8]. In particular, nodules with a size equal or bigger than 3 mm are referred to as lung nodules, and larger ones are likely to become malignant (*i.e.,* cancerous) [9]. In agreement with the differentiation based on the histological size of lung cancer cells under a microscope, about 15% of all lung cancers is Small Cell Lung Cancer (SCLC) and up to 85% is Non-Small Cell Lung Cancer (NSCLC) [3,10,11], as depicted in Fig. 1. SCLC is a mass located mostly along the long axis of the bronchus [7]; whereas, NSCLC is a mass that can infiltrate and encase the structures of the mediastinum [12]. NSCLC has three histological subtypes: ADenoCarcinoma (ADC), Squamous Cell Carcinoma (SCC), and Large Cell Carcinoma (LCC) [3,11,13]. According to the reclassification pointed out by *Travis et al.* [14], the diagnosis of LCC is restricted to surgically-resected cancers with unclear immunohistochemical or morphological differentiation [10,14]. Thus, ADC and SCC account for approximately 90% of NSCLC histological subtypes [15–17]. ADC is the most common NSCLC histological subtype [13]. It begins in the submucosal glands and is located mostly along the outer edges of the lungs, often showing a star-like contour [10,18]. In the majority of cases, it appears like a mass smaller than 3 cm [19]. ADC can be further categorized in invasive ADC (iADC) or in situ ADC (sADC) according to the mass growth [7]. SCC is the second most common NSCLC histological subtype [13]. It begins in the squamous cells and is usually located in the middle of the lungs,

☆ This research was supported by the projects "Using 3D convolutional neural networks for the recognition of lung cancer histotypes directly from computed tomography scans" funded by Cariverona Foundation, Italy, and "Deep learning for early medical diagnosis: A novel methodology for different clinical scenarios" funded by Department of Information Engineering, Engineering Faculty, Università Politecnica delle Marche, Ancona, Italy.

* Corresponding author.
E-mail addresses: s.tomassini@pm.univpm.it (S. Tomassini), n.falcionelli@staff.univpm.it (N. Falcionelli), p.sernani@staff.univpm.it (P. Sernani), l.burattini@univpm.it (L. Burattini), a.f.dragoni@univpm.it (A.F. Dragoni).

https://doi.org/10.1016/j.compbiomed.2022.105691
Received 12 February 2022; Received in revised form 26 May 2022; Accepted 31 May 2022
Available online 6 June 2022

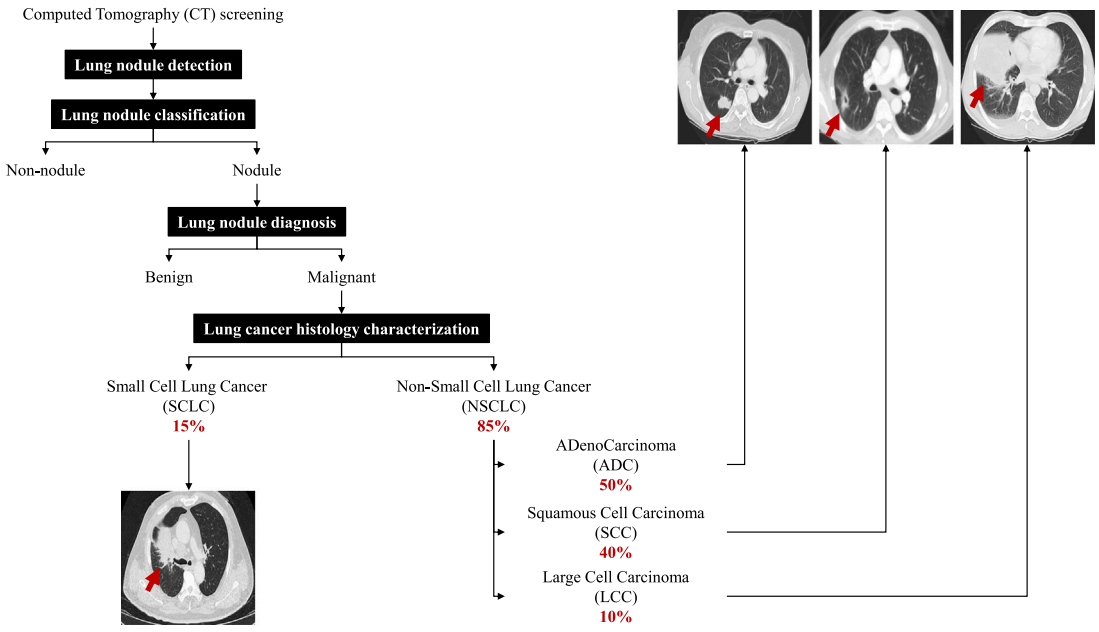

Fig. 1. Non-invasive lung cancer computer-aided diagnosis. After being detected, lung nodules are classified as nodules or non-nodules. Next, suspicious lung nodules are diagnosed as malignant or benign. Then, malignant lung nodules (*i.e.,* lung cancers) are classified according to the histological type/subtype. NSCLC accounts for about 85% of all lung cancers. ADC and SCC are the most common NSCLC histological subtypes (90%).

often presenting central necrosis, with a wall thickness typically larger than 1.5 cm [10]. Its diameter may exceed 4 cm [19]. Microscopically, ADC shows acinar, cribriform, lepidic, micropapillary, papillary, and solid histopathological growth patterns; whereas, SCC shows clusters of polyhedral cells, presence of abundant eosinophilic cytoplasm, and keratin pearl formation [1].

Lung cancer tends to be diagnosed only at its critical stage. However, early lung cancer diagnosis is fundamental to lift the associated survival rate by improving therapeutic decisions [2,20]. It is estimated that the five-year survival rate of diseased patients increases more than 50% as a result of early lung cancer diagnosis and timely treatment [21,22]. As diagnostic imaging modalities in lung cancer prognosis, pulmonologists and radiologists recommend exams such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET) [21,23]. MRI and PET have limitations in detecting lung nodules [24]. MRI, in particular, may miss small lung metastases [3]. CT (especially in the form of low-dose CT) is responsible for a substantial decrease in advanced stage lung cancer and, thereby, reduces mortality from lung cancer significantly [25]. Besides being the most sensitive to small calcified lung masses [5,26,27], CT has other advantages, as it has high spatial resolution, low noise and distortion, it is rapid, non-invasive, affordable, and widely available [21,23,28]. Furthermore, it allows to obtain a 3D characterization of the thorax, as each lung nodule is assessed and additional information about other lung structures can be retrieved [29]. Typically, benign lung nodules have an elliptical shape; whereas, cancerous ones have an irregular structure with a rough texture [30]. However, lung cancer is among the most frequently misdiagnosed diseases by CT scan visualization only [30]. Indeed, considerable efforts are required by radiologists to examine CT scans, search for lung nodules (particularly for smaller ones), and determine whether they are malignant according to their shape, texture, and size, as interpretation heavily depends on radiologists' professional experience [8,31,32]. Under ideal circumstances, radiologists spend up to 5 min per lung nodule [31]. With the presence of destabilizing factors such as distraction, fatigue, and inter- and intra-observer variability, radiologists could end up with wrong or even missing diagnoses, overlooking potentially malignant lung nodules [4,33,34]. Besides lung nodule detection, classification (*i.e.,* false positive reduction and nodule/non-nodule classification), and diagnosis (*i.e.,* malignant/benign classification), lung cancer histology characterization is crucial for clinical treatment options [3,10,35] (Fig. 1), as the effectiveness of chemotherapy and immunotherapy treatments as well as the risk of complications are different for each lung cancer histological type/subtype [36,37]. Lung cancer histological types/subtypes are characterized by some specific alterations that allow to quite easily discriminate them at the molecular level [13]. Nevertheless, it is highly challenging to precisely identify lung cancer histological types/subtypes in terms of their morphological characteristics [16,17]. As a result, in patients suspected of having lung cancer based on CT findings, it is recommended to obtain an adequate tissue sample to accurately identify the histological type/subtype for definitive lung cancer diagnosis [12,36,38]. By inspecting CT scans, suspicious lung nodules are identified. Subsequently, lung biopsy is performed and the microscopic structure of the excised tissue sample is analyzed [16]. Lung biopsy, mainly in the form of transthoracic fine needle biopsy, is the first choice for peripherally-located lung cancers [38]. For lung cancers located in proximity to airways or blood vessels as well as for deeply-located ones, lung biopsy is highly challenging [16]. Therefore, lung biopsy shows the potential to determine the status and type of the excised tissue sample [23], but it is highly invasive with potential clinical implications [29], being strongly discouraged in patients with unfavorable clinical situations [11,16]. Moreover, excising a small tissue sample may not exactly characterize the suspicious lung nodule due to its heterogeneity [10,23,39]. As a result, oncologists could fail in determining lung cancer histology [35]. Thus, developing non-invasive modalities to classify lung cancer histological types/subtypes may not only help clinicians to make targeted treatment decisions timely but also prevent patients from undergoing lung biopsy [11,16,40].

For lung nodule detection, classification, and diagnosis from CT data, several good reviews have been published from 2018 to 2021 [2,5,21,23,41]. In 2018, *Zhang et al.* [21] reviewed both feature engineering algorithms and emerging Convolutional Neural Network (CNN)-based models from a technique-driven perspective, including studies developed up to May 2018. In 2019, *Monkam et al.* [2] summarized CNN-based models developed in 2018, dividing the analysis according to the CNN type. *Zhang et al.* [23] provided a collection of studies up to December 2018 on both feature engineering and Deep Learning (DL) algorithms, dividing the analysis according to the main stages of a

computer-aided diagnosis system. Respectively in 2020 and 2021, also *Cao et al.* [5] and *Gu* et al., [41] based the analysis on the main stages of a computer-aided diagnosis system while reviewing DL algorithms, mainly CNN-based models, developed in 2019 and up to November 2020. For publicly-available lung cancer CT datasets, a good survey by *Adiraju et al.* [26] has been published in 2021. However, there is a lack of investigation targeted at discriminating the classification approaches from a data-driven perspective. Moreover, to the best of our knowledge, reviews devoted to lung cancer histology classification from CT data by means of CNN-based models are not yet existing. In this regard, this survey aims to illustrate some recent advancements in CNNs applied to the investigation of lung cancer from CT data. Specifically, we conducted a comprehensive analysis of slice-based and scan-based approaches using CNNs for lung nodule diagnosis and cancer histology classification from CT data. With this survey, we intend to:

- Show the potential of CNNs and their contribution not only in identifying malignant lung nodules but also in classifying lung cancer histological types/subtypes directly from CT data;
- Point out the strengths and weaknesses of slice-based and scan-based approaches employing CNNs;
- Highlight the challenges and prospective solutions to successfully apply CNNs for such classification tasks.

2. Lung cancer computer-aided diagnosis

Artificial intelligence has been showing remarkable success in medical image analysis, especially due to the rapid progress and outstanding performance of DL-based clinical decision support systems [2,42]. DL algorithms for lung cancer computer-aided diagnosis are more and more used to assist (not replace) clinicians in the decision-making process [2,4,21]. They may improve the diagnostic accuracy while reducing the workload of both radiologists and oncologists [3,5], unambiguously outperforming standard Machine Learning (ML) algorithms [4,24].

2.1. Deep learning in lung cancer diagnosis

Before the expansion of DL, feature engineering algorithms were extensively used in the lung cancer domain [21,22]. Intensity-based, texture-based, and morphology-based lung cancer features were manually extracted from CT data, and injected to traditional ML classifiers such as support vector machine and random forest. *Keshani et al.* [43] employed a support vector machine classifier based on 2D stochastic and 3D anatomical features for lung nodule classification from CT images. Also *Han et al.* [44] employed a feature-based support vector machine classifier in combination with rule-based filtering operations for the same classification task. *Liu et al.* [45] employed a random forest classifier based on 22 shape and texture features of lung nodules. A random forest classifier was also employed by *Gong et al.* [46] to classify lung nodules, previously detected with a 3D tensor filtering approach combined with a local image feature analysis. Despite the quite good results achieved by these studies, traditional ML classifiers have several limitations [41]. For instance, it is difficult to use a support vector machine in case of multi-classification problems and large-scale train samples [47]. Moreover, traditional ML classifiers need manual feature extraction to achieve good results. Manual feature extraction is a time-consuming and non-trivial process, especially in the main scenarios of medical image analysis: complexity of the diagnosis and limited *"a priori"* knowledge [42]. Indeed, despite the clinicians' experience, there is not an extensive knowledge about which quantitative image features predict an output [42]. Specifically, manual feature extraction is extremely challenging for what concerns lung nodule features [41]. With the explosion of DL due to the great empowerment of computational resources, some researchers have started to replace feature engineering algorithms with DL ones [22,24]. DL algorithms do not require too much manual intervention and are characterized by a high degree of automation [37,48], as they can adaptively learn the optimal representation in a fully data-driven way, without relying on manually-extracted lung nodule features [8,20,42]. Furthermore, the base knowledge of DL algorithms from unrelated fields can be transferred to the lung cancer domain easier than the base knowledge of standard ML algorithms [41]. Thus, DL algorithms have many advantages in analyzing lung cancer data.

2.2. Convolutional neural networks in lung cancer diagnosis

Among the several DL algorithms, CNNs have become the leading choice for lung nodule detection, classification, and diagnosis from CT data [2,5,21–24,41]. In particular, the employment of CNNs such as AlexNet, Densely-connected convolutional Network (DenseNet), GoogLeNet with Inception units, Residual Network (ResNet), Visual Geometry Group Network (VGGNet), and their derivatives is the most prevalent [24,49]. CNNs are implemented with convolution and pooling strategies; thus, they can simultaneously manage feature construction, feature selection, and prediction modeling, performing an end-to-end analysis from input data to predictions [16]. Automatically-extracted features from CNNs preserve the spatial information with the convolutional kernel operations on input data [39]. This is advantageous in contextual recognition, domain adaptation, fine-grained recognition, and texture attribute recognition [50]. CNNs are also less human-dependent [39]. This factor significantly reduces the bias. As a result, they are extremely-powerful and labor-saving DL algorithms, also responsible for the significant improvement of the accuracy in lung cancer diagnosis [5,51,52].

The majority of CNN-based models relies on slices. As a CT scan is inherently 3D, such models convert the volumetric information into 2D multi-channel data. Inspecting the CT scan slice by slice may result in the loss of valuable information [3,53], as the inherent spatial relationship of slices is neglected [54]. Therefore, incorporating the volumetric information by keeping the slice order may positively impact the learning process. In this regard, CNN-based models that take as input a group of slices as a whole (*e.g.,* lung nodule volume, sequence of lung slices, lung volume, sequence of CT slices, or CT volume), referred to as scan-based approaches, are likely to capture richer spatial information and extract more discriminative features than CNN-based models that take as input slices separately (*i.e.,* slice-based approaches), as schematically depicted in Fig. 2. Thus, scan-based approaches have the potential to investigate the overall information in a more detailed way, resulting in more reliable clinical judgments [51,55]. They include 3D CNNs as well as less computationally costly models such as time-distributed 2D CNNs coupled with Recurrent Neural Network (RNN) modules for integrating the spatial coherence of slices. In this case, RNN exploits its internal state to analyze spatial sequences, being able to connect the already-processed information to the present one [10].

3. Search strategy and selection criteria

The search strategy that we followed to identify and select the state-of-the-art articles included the following steps:

- Definition of the problem;
- Choice of the articles which satisfy all the selection criteria;
- Extraction of the most salient information from each selected article;
- Analysis of the extracted information.

The preliminary selection criteria were:

- For lung nodule diagnosis-related studies, we restricted time to the last two years to avoid a complete overlap with existing reviews. In this guise, we searched for Scopus-indexed works (Q1 and Q2 scientific articles in at least one of the following sectors:

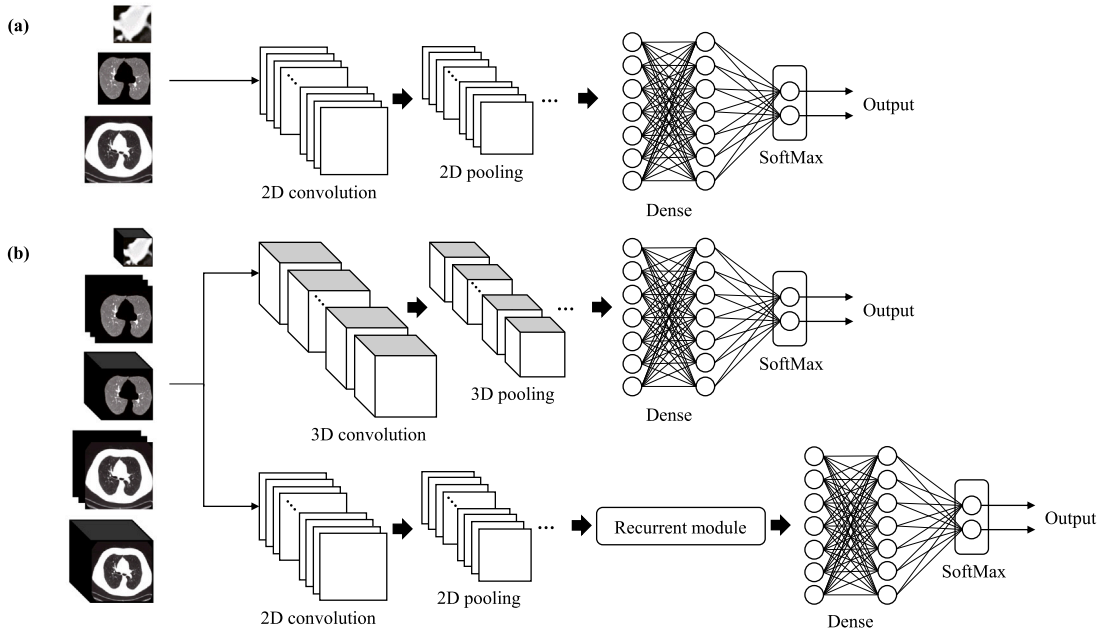

Fig. 2. A schematic illustration of slice-based (**a**) and scan-based (**b**) approaches using CNNs. Slice-based approaches are fed with planar data (*e.g.,* lung nodule slices, lung slices, or CT slices, separately) and include 2D CNNs. Scan-based approaches are fed with volumetric data (*e.g.,* lung nodule volumes, sequence of lung slices, lung volumes, sequence of CT slices, or CT volumes) and include 3D CNNs as well as time-distributed 2D CNNs coupled with recurrent modules for spatial information integration.

Table 1
Article search outcomes.

	Filtration by preliminary selection criteria	Filtration by language, title, abstract, and keywords	Filtration by full-text analysis	Selected articles
Lung nodule diagnosis	32	21	13	11
Lung cancer histology classification	11	10	7	7

artificial intelligence; biomedical engineering; computer science; computer science applications; health informatics; and radiology, nuclear medicine and imaging) published from January 2019 up to and including January 2022. All searches included the words "lung cancer" AND "lung nodule diagnosis" AND "malignant" AND "benign", in conjunction with ("CT" OR "computed tomography") AND ("DL" OR "deep learning") AND "supervised learning" AND "classification" AND ("CNN" OR "convolutional neural network"), using Boolean "AND" to join all words and "OR" to include synonyms;

- For lung cancer histology classification-related studies, we applied no time restriction, as there are not existing reviews on the same topic. In this guise, we searched for Scopus-indexed works (Q1 and Q2 scientific articles in at least one of the following sectors: artificial intelligence; biomedical engineering; computer science; computer science applications; health informatics; and radiology, nuclear medicine and imaging) published up to and including January 2022. In searches, we considered the terms ("SCLC" OR "small cell lung cancer") AND ("NSCLC" OR "non-small cell lung cancer") AND ("ADC" OR "adenocarcinoma") AND ("SCC" OR "squamous cell carcinoma") AND ("CT" OR "computed tomography") AND ("DL" OR "deep learning") AND "supervised learning" AND "classification" AND ("CNN" OR "convolutional neural network"), using Boolean "AND" to join all terms and "OR" to include synonyms.

In both searches, we then filtered the articles by language (English only), title, abstract, and keywords. After this further filtration, as the overall volume of studies was relatively low given the time/technical restrictions in the first case and the recency of the topic in the second one, we performed a full-text analysis of the remaining articles, selecting them according to their technical details. This resulted in preserving only 18 relevant Scopus-indexed publications that employed CNNs for such classification tasks.

Table 1 synthesizes the article search outcomes.

4. Literature descriptive analysis

We conducted a descriptive analysis of the selected studies based on the article information to assess their relevance in the lung cancer domain. Specifically, we carried out this analysis according to the year of publication, dataset, input type, input size, best-performing CNN, classes, and overall performance. Regarding the performance, we decided to report the ACCuracy (ACC) and Area Under the receiver operating characteristic Curve (AUC), as ACC represents the number of correct predictions over the total number of predictions and AUC highlights the ability to discern between the classes without depending on the discrimination threshold. Given its extreme importance in the clinical scenario, we decided to report also the SEnsitivity (SE), which reflects the number of correctly classified positive cases [54].

Sections 4.1 and 4.2 provide a comprehensive overview of the literature descriptive analysis for lung nodule diagnosis and cancer histology classification, respectively.

4.1. Lung nodule diagnosis

Relevant success has been achieved in DL-based differentiation of lung nodules from other lung lesions [4,20,32,55–59]. However, due to the heterogeneity of lung nodules, it is still challenging to obtain a satisfactory classification of their status from CT data, as not all detected lung nodules turn out to be cancerous [60]. For this reason, we focused our first investigation specifically on malignant/benign lung nodule

Table 2
Relevant Scopus-indexed studies involved in lung nodule diagnosis from CT data using CNNs.

Study	Year	Dataset	Input type	Input size (mm)	Best-performing CNN	Classes	ACC (%)	SE (%)	AUC (%)
Al-Shabi et al. [61]	2019	LIDC-IDRI	Nodule images	32 × 32	Gated dilated 2D CNN	Malignant/ benign	93	92	95
Shen et al. [8]	2019	LIDC-IDRI	Nodule volumes	52 × 52 × 52	Hierarchical semantic 3D CNN	Malignant/ benign	84	70	86
Ali et al. [62]	2020	LIDC-IDRI SPIE-AAPM-NCI	Nodule images	64 × 64	Transferable texture 2D CNN	Malignant/ benign	97 86	96 –	99 –
Lin et al. [63]	2020	LIDC-IDRI SPIE-AAPM-NCI	Nodule images	50 × 50	Taguchi optimized 2D CNN	Malignant/ benign	99 100	100 100	– –
Liu et al. [60]	2020	LIDC-IDRI	Nodule volumes	48/32/16 × 48/32/16 × 48/32/16	Multi-model 3D CNN	Malignant/ benign	90	84	94
Zhai et al. [64]	2020	LIDC-IDRI LUNA16	Nodule images	64 × 64	Multi-task 2D CNN	Malignant/ benign	– –	88 84	96 97
Zhao et al. [51]	2020	LIDC-IDRI	Nodule volumes	32 × 32 × 6	Multi-scale multi-task 3D CNN	Malignant/ benign	94	93	98
Halder et al. [30]	2021	LIDC-IDRI	Nodule images	64 × 64	Two-path morphological 2D CNN	Malignant/ benign	96	97	99
Jena et al. [65]	2021	LIDC-IDRI	Lung images	–	Region-based 2D CNN	Malignant/ benign	88	70	–
Lu et al. [66]	2021	RIDER	CT images	227 × 227	Marine predator-based 2D CNN	Malignant/ benign	93	98	–
Yu et al. [27]	2021	LIDC-IDRI	Nodule volumes	48 × 48 × 16	3D ResNet50	Malignant/ benign	87	80	91

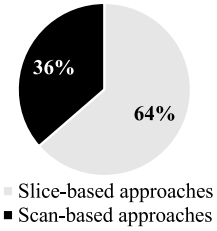

Fig. 3. Percentage of slice-based (light gray) and scan-based (black) approaches performing lung nodule diagnosis from CT data using CNNs.

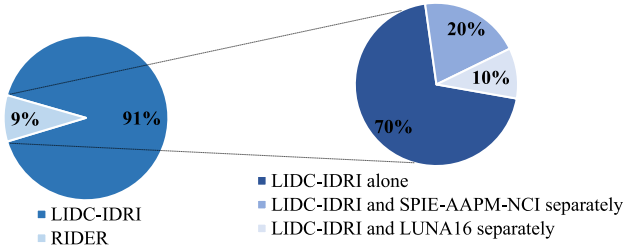

Fig. 4. Datasets employed for lung nodule diagnosis with the respective percentage. LIDC-IDRI[1] stands for Lung Image Database Consortium and Image Database Resource Initiative, SPIE-AAPM-NCI[2] stands for LUNGx Challenge of SPIE with the support of the American Association of Physicists in Medicine and the National Cancer Institute, LUNA16[3] stands for Lung Nodule Analysis 2016, and RIDER[4] stands for Reference Image Database to Evaluate therapy Response. All employed datasets are openly accessible on The Cancer Imaging Archive (TCIA) platform.

classification from CT data adopting CNNs, differentiating between slice-based (Section 4.1.1) and scan-based approaches (Section 4.1.2).

Table 2 reports 11 relevant Scopus-indexed studies involved in lung nodule diagnosis from CT data using CNNs, better detailed in the following two subsections. For each study, we indicated the publication year, the employed dataset, the input type (lung nodule images, lung images, CT images, lung nodule volumes, sequences of lung slices, lung volumes, sequences of CT slices, or CT volumes), the input size, the best-performing CNN, the classes, and the overall performance in terms of ACC, SE, and AUC. Fig. 3 depicts the percentage of slice-based and scan-based approaches hereby investigated. Fig. 4 displays which dataset has been used across the studies and the respective percentage.

4.1.1. Slice-based approaches

In 2019, *Al-Shabi et al.* [61] developed a 2D CNN-based model, named Gated dilated, to classify lung nodules as malignant or benign from CT images. Gated dilated 2D CNN exploited multiple dilated convolutions in place of max-pooling layers for catching multiple-scale features, and had a context-aware sub-network that analyzed and guided the features to a suitable dilated convolution. Besides Gated dilated 2D CNN, they implemented other CNN-based models: (1) a conventional CNN with the same number of layers and channels per layer of Gated dilated 2D CNN; (2) Gated dilated 2D CNN with no dilation; (3) Gated dilated 2D CNN with no gating; (4) a multi-crop CNN; and (5) ResNet50 and DenseNet161 trained with two different transfer learning modalities. First, they manually segmented lung nodules using the annotations of four radiologists. Next, they used a tri-linear interpolation to normalize them, ending up with an isotropic pixel resolution. Then, they extracted square regions of 32 × 32 pixels about the center of each lung nodule, and used them to feed Gated dilated 2D CNN. Through data augmentation (rotation and Gaussian blurring), they obtained 15 different images for each lung nodule and used them for training. They implemented Gated dilated 2D CNN in Python using PyTorch framework, and ran the experiments on a machine with one NVIDIA Titan X Pascal GPU. To perform the experiments, they selected 244617 images (1018 CT scans of 1010 patients) from the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI)[1] dataset of The Cancer Imaging Archive (TCIA) platform. Gated dilated 2D CNN achieved an overall ACC of 93%, SE of 92%, and AUC of 95% on test data, slightly outperforming all the other exploited CNN-based models. Results also showed a significant accuracy improvement in detecting medium-sized nodules (5 to 12-mm diameter).

In 2020, *Ali et al.* [62] designed a transferable texture 2D CNN to improve lung nodule status classification performance from CT images, and evaluated it on the LIDC-IDRI[1] (244617 images) and the LUNGx Challenge of SPIE with the support of the American Association of

[1] https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI.

Physicists in Medicine (AAPM) and the National Cancer Institute (NCI)[2] (22489 images) datasets of TCIA platform, separately. They incorporated an energy layer in the model architecture to extract texture features from the convolutional layer. The inclusion of the energy layer reduced the number of model parameters, decreasing the computational complexity and memory requirements. They first performed data augmentation by using operations like image translation, random rotation, and flip. To manage data heterogeneity, they converted all images to Hounsfield Unit (HU) scale and transformed them to a range of (0, 1) from (−1000, 500 HU). After this, they extracted a Region Of Interest (ROI) around each lung nodule by acquiring the coordinates and slice number from the associated XML file. At the end of the preprocessing step, they used as input images centered at lung nodules with a shape of 64 × 64 pixels. Then, they exploited transfer learning to train their model, running the experiments on a machine with one NVIDIA Titan Xp 12 GB GPU and one 16 GB RAM. The transferable texture 2D CNN achieved the best performance on the LIDC-IDRI dataset, with an overall ACC of 97%, SE of 96%, and AUC of 99% on test data.

Lin et al. [63] proposed a 2D CNN with Taguchi parametric optimization to automatically classify lung nodules as malignant or benign from CT images. In the Taguchi method, which is a statistical method using an orthogonal array to optimize process parameters, they selected 36 experiments and 8 control factors of mixed levels in order to determine the optimal parameter combination and also improve the model performance. Next, they extracted lung nodule images of 50 × 50 pixels, and set the optimization parameters to train the neural network for malignant/benign classification. Specifically, they trained and validated their model on the LIDC-IDRI[1] (244617 images from 1018 CT scans) and the LUNGx SPIE-AAPM-NCI[2] (22489 images from 68 CT scans) datasets of TCIA platform, separately. Results of 2D CNN with Taguchi parameter optimization reached an overall ACC of 99% and SE of 100% on the LIDC-IDRI dataset, and ACC of 100% and SE of 100% on the LUNGx SPIE-AAPM-NCI dataset. Performance also showed that their framework resulted to be 7% and 5% more accurate than the original 2D CNN (without Taguchi parametric optimization) on both datasets.

Zhai et al. [64] presented a multi-task 2D CNN to discriminate malignant from benign lung nodules on CT images. They first resampled each image to a unified value using the spline interpolation. They also rescaled the pixel intensity range to (0, 1) from (−1000, 400 HU), and performed lung parenchyma segmentation. Next, they decomposed each nodule cube of 64 × 64 × 64 voxels to nine different 2D views of 64 × 64 pixels. At that point, they augmented train data using random image translation, rotation, and flip. Then, they built a multi-task 2D CNN for each view, composed by two branches: one for malignant/benign classification (main task) and the other for image reconstruction (auxiliary task). The final classification result was obtained by fusing the outputs of nine multi-task 2D CNNs. They performed the experiments employing images from the LIDC-IDRI[1] (175 malignant and 266 benign lung nodules from 1018 CT scans) dataset and one of its subsets, the Lung Nodule Analysis 2016 (LUNA16)[3] (1120 lung nodules from 888 CT scans) dataset of TCIA platform, separately. Their multi-task 2D CNN achieved an overall SE of 88% and 84%, and AUC of 96% and 97% on test data, respectively.

In 2021, *Halder et al.* [30] designed a framework using adaptive morphology-based operations combined with Gabor filter, named two-path morphological 2D CNN, for lung nodule status classification from CT images. Specifically, they selected 2600 lung nodule slices, 1300 benign and 1300 malignant, from the LIDC-IDRI[1] dataset of TCIA platform. They used different morphology-based operations to filter lung nodules. Next, they used Gabor filter to capture the texture variations of lung nodules. The two-path morphological 2D CNN consisted of two paths, both using VGGNet: the first path with Gabor filter-based weight initialization and convolution operations, and the second path with adaptive morphology-based weight initialization and image morphology-based operations. The input of both paths was a lung nodule image of 64 × 64 pixels. Also, they performed train data augmentation (translation, horizontal and vertical flip operations, and 90°, 180° and 270° rotation), ending up with a total of 10374 lung nodule images. They implemented the two-path morphological 2D CNN in Python on Google platform, using Google Colab Pro. Their two-path morphological 2D CNN achieved an overall ACC of 96%, SE of 97%, and AUC of 99% on train data.

Jena et al. [65] proposed a neural network with multiple layer latent variables to classify lung nodules as malignant or benign from CT images. First, they employed images of the LIDC-IDRI[1] dataset of TCIA platform to carry out the experiments. Second, they removed the noise using Gaussian and Wiener filters. Third, they segmented the ROI using the region-growing segmentation, which selected the salient pixels and merged the adjacent ones to obtain larger regions. Next, they manually extracted effectual and sharp features (*e.g.,* area, perimeter, entropy, intensity, etc.) for detecting lung nodules. Then, they reduced the dimensionality of the featured output using a deep Gaussian mixture model and performed the classification with a region-based 2D CNN. They carried out the experiments in the MATLAB 2018a environment, with 8 GB RAM. Although their model achieved an overall ACC of 88% and SE of 70% on test data, they planned to modify the feature extraction step exploiting fully-automatic feature extraction algorithms in order to possibly lift the performance.

Lu et al. [66] presented a 2D CNN with the metaheuristic-based marine predator algorithm, used for solving the optimization problems, to discriminate cancerous lung nodules from benign ones on CT images. They selected CT images (32 subjects with NSCLC) from the Reference Image Database to Evaluate therapy Response (RIDER)[4] dataset of TCIA platform. They first pre-processed input data employing the median filter. After noise removal, they performed intensity normalization to (0, 1) using the min–max normalization method. Next, they rescaled all images to 227 × 227 pixels, and injected them to the neural network. They implemented their framework on MATLAB software, and compared model performance with other pre-trained 2D CNNs (ResNet18, GoogLeNet, AlexNet, and VGG19). Results showed that the marine predator algorithm-based 2D CNN outperformed all the other models, achieving an overall ACC of 93% and SE of 98% on test data.

4.1.2. Scan-based approaches

In 2019, *Shen et al.* [8] designed an interpretable hierarchical semantic 3D CNN to determine whether a lung nodule observed on a CT scan was malignant. The hierarchical semantic 3D CNN took as input raw cubes centered at lung nodules, belonging to the LIDC-IDRI[1] dataset of TCIA platform, and generated two output levels: the first predictive level provided intermediate outputs (diagnostic semantic features) and the second represented the final lung nodule malignancy prediction score. The intermediate outputs provided not only interpretations about what the model learned from raw data but also additional information to make the final malignancy prediction task more accurate by using jump connections. Moreover, they introduced a cost function to train the model as a whole. First of all, they converted all CT scans to HU scale and normalized them to a range of (0, 1) from (–1000, 500 HU). Next, they extracted a cube of 40 × 40 × 40 voxels for each lung nodule candidate. Then, they rescaled all cubes to a fixed shape before using them to feed the neural network. Additionally, they performed 3D data augmentation (translation of the lung nodule position within 4 mm or flip of the lung nodule cube along one axis)

[2] https://wiki.cancerimagingarchive.net/display/Public/LUNGx+SPIE-AAPM-NCI+Lung+Nodule+Classification+Challenge.

[3] https://luna16.grand-challenge.org/Download/.

[4] https://wiki.cancerimagingarchive.net/display/Public/RIDER+Lung+CT.

during the training process. To capture more lung nodule morphology while reducing the input data dimensionality, they set the lung nodule cube to 52 × 52 × 52 voxels. To compare the performance of the model in lung nodule status classification, they also implemented a 3D CNN as baseline, trained and validated with the same data split. They implemented both hierarchical semantic 3D CNN and baseline 3D CNN in Python using TensorFlow framework and Keras library, and ran the experiments on a machine with one NVIDIA Titan X Pascal 12 GB GPU and one 32 GB RAM. The interpretable hierarchical semantic 3D CNN achieved an overall ACC of 84%, SE of 70%, and AUC of 86% on test data. All metric assessments and receiver operating characteristic plots showed that their hierarchical semantic 3D CNN achieved better performance than the baseline 3D CNN.

In 2020, *Liu et al.* [60] presented a 3D CNN-based multi-model ensemble learning to identify malignant from benign lung nodules on CT scans, selected from the LIDC-IDRI[1] dataset of TCIA platform. Specifically, they designed multiple independent neural networks to simulate different expert behaviors, and exploited ensemble learning to fuse the results. The multi-model ensemble learning 3D CNN consisted of three different types of architectures respectively based on VGGNet (3D multi-model VGGNet), ResNet (3D multi-model ResNet), and InceptionNet (3D multi-model IncepNet), further divided into three substructures of different input sizes (the inputs were cubes centered at lung nodules of 48 × 48 × 48 voxels, 32 × 32 × 32 voxels, and 16 × 16 × 16 voxels). In order for the model to better handle low-contrast lung nodules, they performed five image enhancement techniques: (1) histogram equalization; (2) adaptive histogram equalization; (3) gamma transformation; (4) logarithmic transformation; and (5) intensity stretch transformation. Next, to make the model better focus on lung nodule location, shape, and intensity, they concatenated the intensity image corresponding to the lung nodule mask with the original and enhanced images, forming a three-channel 3D input. Then, they augmented the train set by performing 90° rotation and flip along the three axes. They implemented the multi-model ensemble learning 3D CNN in Python using PyCharm as integrated development environment and Keras library, and ran the experiments on a machine with three NVIDIA GTX-1080Ti 11 GB GPUs. Their multi-model ensemble learning 3D CNN achieved an overall ACC of 90%, SE of 84%, and AUC of 94% on test data.

Zhao et al. [51] proposed a multi-scale multi-task 3D CNN to classify lung nodules as malignant or benign from CT scans. They selected 1004 lung nodules (450 malignant and 554 benign) from the LIDC-IDRI[1] dataset of TCIA platform, regulated the size of each voxel to 1 mm^3, and extracted volumes at two different scales. Specifically, the first volume had a size of 32 × 32 × 6 voxels and the second of 64 × 64 × 12 voxels, then adjusted to a volume of 32 × 32 × 6 voxels. To perform the experiments, they built a 3D CNN that combined the features of the two different scale volumes, followed by a multi-task learning that realized both malignant/benign and attribute classification of lung nodules. They also proposed a new loss function. They implemented the multi-scale multi-task 3D CNN in Python using PyTorch framework, and ran the experiments on a machine with four NVIDIA GTX 1080 8 GB GPUs and four 32 GB RAMs. Their multi-scale multi-task 3D CNN achieved an overall ACC of 94%, SE of 93%, and AUC of 98% on test data.

In 2021, *Yu et al.* [27] presented a framework based on 3D Res U-Net lung nodule segmentation and 3D ResNet50 malignant/benign classification networks, both applied on CT scans. They first converted CT scans to HU scale and resampled them to a voxel spacing of 3 × 1.5 × 1.5 mm (axial, coronal, and sagittal planes). Next, they binarized CT scans, and removed disturbing regions such as bed frame and air. Then, they filled the holes in the lung parenchyma and repaired the lung mask through morphological techniques. At that point, they designed 3D Res U-Net to automatically segment lung nodules. According to the predicted lung nodule mask, they extracted the ROIs of 48 × 48 × 16 voxels and used them as input to 3D ResNet50. They implemented both 3D Res U-Net and 3D ResNet50 in Python using PyTorch framework, and ran the experiments on a machine with one NVIDIA RTX 2080Ti GPU, selecting 971 CT scans (1507 malignant and 1478 benign nodules) from the LIDC-IDRI[1] dataset of TCIA platform. Classification results showed that 3D ResNet50 achieved an overall ACC of 87%, SE of 80%, and AUC of 91% on test data.

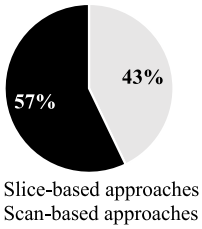

Fig. 5. Percentage of slice-based (light gray) and scan-based (black) approaches performing lung cancer histology classification from CT data using CNNs.

4.2. Lung cancer histology classification

Aforementioned studies just focused on diagnosing the status of lung nodules. In differential diagnosis of lung cancer, however, accurate classification of lung cancer histological types/subtypes is required [39].

Conventionally, researchers aimed to classify lung cancer histological types/subtypes by analyzing microscopic images, and 2D CNN is the most utilized DL algorithm [37,42,67,68]. Nevertheless, lung cancer histology classification using routinely-acquired CT data may have significant implications for diagnostic and therapeutic decisions [39]. Some studies explored the potential of radiomics-based ML algorithms in lung cancer histology classification from CT data [69–75]. However, these classification architectures relied upon pre-defined radiomics features to train the model. By adopting different radiomics features, it is difficult for clinicians to select the appropriate feature set, making the application of these models not straightforward in the clinical practice [17]. The major limitation of radiomics, indeed, is the lack of standardization in the acquisition of the features, that implies the lack of reproducibility [76,77]. In particular, CT-derived radiomics features strictly depend on the number of gray levels and on the voxel size [77]. Such dependencies demonstrate that CT-derived radiomics feature applicability is highly influenced by the number of intensity bins and by the voxel size choice [77]. Thus, a more automatic feature detection system is essential for improving the diagnostic performance. For this reason, we focused our second investigation specifically on lung cancer histology classification directly from CT data adopting CNNs, differentiating between slice-based (Section 4.2.1) and scan-based approaches (Section 4.2.2).

Table 3 reports 7 relevant Scopus-indexed studies involved in lung cancer histology classification from CT data using CNNs, better detailed in the following two subsections. For each study, we indicated the publication year, the employed dataset, the input type (lung cancer images, lung images, CT images, lung cancer volumes, sequences of lung slices, lung volumes, sequences of CT slices, or CT volumes), the input size, the best-performing CNN, the classes, and the overall performance in terms of average ACC and AUC, and SE for each pathologically-proven class. Fig. 5 depicts the percentage of slice-based and scan-based approaches hereby investigated. Fig. 6 displays which dataset has been used across the studies and the respective percentage.

4.2.1. Slice-based approaches

In 2020, *Pang et al.* [40] proposed a DenseNet with a meta-algorithm classifier (adaptive boosting) to classify SCLC, ADC, and SCC from 2222 CT images of a private data collection (Shandong Provincial Hospital). Among them, 96 were SCLC, 1985 ADC, and 141 SCC images. First, they adopted the histogram equalization method to obtain a uniform distribution in the gray range, enhancing the contrast of each image

Table 3
Relevant Scopus-indexed studies involved in lung cancer histology classification from CT data using CNNs.

Study	Year	Dataset	Input type	Input size (mm)	Best-performing CNN	Classes	ACC (%)	SE (%)	AUC (%)
Guo et al. [11]	2020	Private	Cancer volumes	64 × 64 × 64	End-to-end 3D CNN	SCLC/ ADC/SCC	72	-(SCLC)/ -(ADC)/-(SCC)	84
Moitra et al. [35]	2020	NSCLC-Radiogenomics	Slice sequences	64 × 64	2D CNN + biLSTM	ADC/ SCC/other	96	-(ADC)/ -(SCC)/-(other)	99
Pang et al. [40]	2020	Private	Cancer images	50 × 50	DenseNet + boosting	SCLC/ ADC/SCC	90	61(SCLC)/ 100(ADC)/98(SCC)	–
Pang et al. [7]	2020	Private	Cancer images	50 × 50	VGG16-T + boosting	SCLC/iADC/ sADC/SCC	87	90(SCLC)/86(iADC)/ 76(sADC)/92(SCC)	–
Chaunzwa et al. [39]	2021	Private	Cancer images	50 × 50	Pre-trained VGG16	ADC/ SCC	69	38(ADC)/ -(SCC)	71
Liu et al. [17]	2021	Private	Cancer volumes	64 × 64 × 64	3D Capsule Net	ADC/ SCC	81	-(ADC)/ -(SCC)	85
Marentakis et al. [10]	2021	NSCLC-Radiomics	Slice sequences	299 × 299	Inceptionv3 + LSTM	ADC/ SCC	74	81(ADC)/ -(SCC)	78

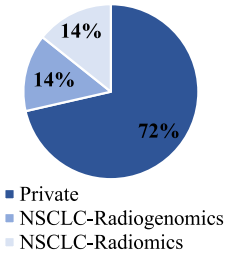

Fig. 6. Datasets employed for lung cancer histology classification with the respective percentage. Among the employed datasets, only NSCLC-Radiogenomics[5] and NSCLC-Radiomics[6] (28%) are openly accessible on The Cancer Imaging Archive (TCIA) platform.

and making it clearer. Next, they augmented and balanced the train set by using image rotation, translation, and transformation methods, ending up with a total of 3940 images centered at lung cancers, each of 50 × 50 pixels. Then, they used a DenseNet architecture to classify lung cancers. Finally, they aggregated multiple results by using adaptive boosting in order to improve the model performance. DenseNet with adaptive boosting showed better classification performance than AlexNet, DenseNet without adaptive boosting, ResNet, and VGG16. Indeed, DenseNet with adaptive boosting achieved an overall ACC of 90% and a SE of 61%, 100%, and 98% in identifying respectively SCLC, ADC, and SCC; whereas, the overall ACC of AlexNet was 64%, of DenseNet without adaptive boosting was 81%, of ResNet was 80%, and of VGG16 was 76% on test data.

Pang et al. [7] developed another classification algorithm, named VGG16-T with boosting, to non-invasively classify SCLC, iADC, sADC, and SCC from CT images. They considered 2219 CT images (125 patients, 96 SCLC, 1882 iADC, 101 sADC, and 140 SCC) from Shandong Provincial Hospital. First, they performed 1-mm resampling to a fixed isomorphic resolution. Next, they augmented and balanced train data up to 50000 CT images by random rotation, translation, and transformation operations, enhancing multiple times the classes with less samples. Then, they extracted lung nodules, ending up with input data of 50 × 50 pixels. After that, they choose VGG16 as main structure due to its high performance in CT image diagnosis, and proposed a new architecture, named VGG16-T. VGG16-T had three 1 × 1 convolutions in place of the fully-connected layer of the traditional VGG16. To avoid overfitting, they pre-trained VGG16-T on the publicly-accessible LUNA16 dataset. Since VGG16-T worked as weak classifier, they trained multiple VGG16-T models with the boosting strategy, repeated until the weak classifiers reached an acceptable ACC. They implemented VGG16-T with boosting in Python, and ran the experiments on one NVIDIA GTX 2080Ti 12 GB GPU. They found that three weak classifiers, linearly combined, were enough to achieve an overall ACC of 87% and a SE of 90%, 86%, 76%, and 92% in classifying respectively SCLC, iADC, sADC, and SCC. Also, VGG16-T with boosting performed better than VGG16-T without boosting, AlexNet, ResNet34, and DenseNet with or without Softmax weights. As well, VGG16-T with boosting gained an overall ACC of 85% by diagnosing 20 randomly-selected CT images.

In 2021, *Chaunzwa et al.* [39] presented a transfer learning approach to classify NSCLC histological subtypes (ADC and SCC) from CT images belonging to 311 patients of Massachusetts General Hospital. The preprocessing step was made of manual cancer identification (thanks to clinician-located seed points), isotropic rescaling, and density normalization with mean subtraction and linear transformation. To perform the experiments, they exploited VGG16 (pre-trained on ImageNet) coupled with different classifier types: (1) a fully-connected classifier; and (2) different ML classifiers (k-nearest neighbors, random forest, support vector machine, and linear support vector machine). Specifically, inputs of VGG16 were cancer-centered images of 50 × 50 pixels, and the model was evaluated with fine-tuning of the last convolutional, pooling, and fully-connected layers. VGG16 with fully-connected classifier achieved an overall ACC of 69% and AUC of 71% on test data, and SE was 38% as they considered ADC as positive class. Eventually, they obtained activation heatmaps using Gradient-weighted Class Activation Mapping (Grad-CAM), providing a spatial representation of the input image areas which contributed the most to the model predictions.

4.2.2. Scan-based approaches

In 2020, *Guo et al.* [11] developed two automatic classification models to distinguish lung cancer histological types/subtypes (SCLC, ADC, and SCC) from unenhanced CT scans, involving 920 patients (191 SCLC, 554 ADC, and 175 SCC) from a private data collection. First of all, the ROI was automatically delineated via threshold or clustering and, consequently, manually corrected/confirmed by three radiologists. The first model, named ProNet, was based on an end-to-end 3D CNN, implemented using the skip-connection method of ResNet. Its input consisted in volumes of 64 × 64 × 64 voxels, centered at lung cancers. They first normalized input data to (−1, 1), and performed train data augmentation. They implemented ProNet in Python with TensorFlow framework and Keras library, and ran the experiments on a machine with two NVIDIA 1080Ti GPUs, reaching an overall ACC of 72% and AUC of 84% on test data. The other model, named com_radNet, was based on radiomics and was composed of four fully-connected layers. They extracted 1743 radiomics features with PyRadiomics, starting from a mask delineated by three radiologists. After the procedure of feature selection, they retained 20 radiomics features and used them to feed com_radNet. This second model achieved an overall accuracy of 75% and AUC of 79% on test data. Results indicated that both models are able to distinguish SCLC, ADC, and SCC, but the performance of ProNet was better than the one of com_radNet. Additionally, they

created activation heatmaps with Grad-CAM to visually highlight the most salient portions on lung cancer volumes.

Moitra et al. [35] presented a DL algorithm that is an ensemble of 2D CNN and bidirectional Long Short-Term Memory (biLSTM), and evaluated it on CT scans belonging to 211 patients of the publicly-available NSCLC-Radiogenomics[5] dataset (285411 CT images) of TCIA platform. Their study included not only ADC and SCC, but also a third non-specified NSCLC histological subtype. They first resized original slices to 64 × 64 pixels from 128 × 128 pixels in order to reduce the execution load, and used them to feed the neural network. In the learning process, the output of the last convolutional layer of the 2D CNN was flattened and injected to the biLSTM, deputed to fuse the spatial information of adjacent slices. They implemented their model in Python, using CPU cores of an Intel(R) Core (TM) i5-3230 m CPU @ 2.60 GHz processor. Ensemble 2D CNN with biLSTM reached an overall ACC of 96% and AUC of 99% on test data.

In 2021, *Liu et al.* [17] built and compared three frameworks to classify NSCLC histological subtypes as ADC or SCC from CT scans. To perform the experiments, they collected CT scans of 72 ADC and 54 SCC patients from a private data collection (Third Medical Center of PLA General Hospital). For every patient, they semi-automatically segmented the volume containing the lung cancer with the ITK-SNAP software and the help of an expert radiologist. Next, they resampled each volume to an isotropic voxel dimension (1 mm^3) by means of linear and nearest-neighbor interpolations. Then, they developed three models: (1) a Capsule Net model; (2) a CNN modified to take as input volumetric data; and (3) four radiomics-based ML classifiers (random forest, logistic regression, logistic regression with l1 regularization, and logistic regression with principal component analysis). Specifically, Capsule Net and the modified CNN shared the same training strategy, were implemented in Python using PyTorch framework, and took in input isotropic volumes of 64 × 64 × 64 pixels encapsulating lung cancers. Radiomics-based ML classifiers took in input 107 radiomics features, extracted from the lung cancer volumes with PyRadiomics, for each patient. Results showed that Capsule Net reached the best classification performance, with an overall ACC of 81% and AUC 85% on test data. The classification performance of the modified CNN (overall ACC of 75%) was comparable to those of radiomics-based ML classifiers.

Marentakis et al. [10] investigated the potential of both ML and DL algorithms to classify NSCLC histological subtypes as ADC or SCC from CT scans of the publicly-available NSCLC-Radiomics[6] dataset (102 patients, 48 ADC and 54 SCC) of TCIA platform. Specifically, they experimented four different models: (1) two radiomics-based ML classifiers (k-nearest neighbors and support vector machine); (2) four pre-trained 2D CNNs (AlexNet, Inceptionv3, InceptionResNetv2, and ResNet101) with fine-tuning; (3) a time-distributed 2D CNN coupled with a Long Short-Term Memory (LSTM) module; and (4) two joint models (time-distributed 2D CNN with LSTM and k-nearest neighbors or support vector machine). Models (1) and (2) ignored potentially-relevant information about the spatial coherency of slices; whereas, models (3) and (4) took it into account. Regarding CNN-based models, they cropped slices centered at lung cancers according to the receptive field of each specific CNN. Next, they applied a linear mapping, corresponding to the lung window. Then, they performed a three-fold replication of the intensity channel, as the employed CNNs required an input with three channels. The model that reached the best performance was Inceptionv3 with LSTM, achieving an overall ACC of 74% and AUC of 78%, and SE was 81% considering ADC as positive class. The performance achieved by Inceptionv3 with LSTM resulted to be better than expert radiologists' outcomes up to 25%. A notable finding was that adding radiomics to the best-performing model did not show any further performance improvement.

[5] https://wiki.cancerimagingarchive.net/display/Public/NSCLC+Radiogenomics.

[6] https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics.

5. Discussion

With this survey, we aimed to illustrate some recent advancements in CNNs applied to lung nodule diagnosis and cancer histology classification from CT data. We obtained this outcome by collecting 18 relevant Scopus-indexed studies and categorizing them into slice-based and scan-based approaches, based on whether they took planar or volumetric CT data as input (Fig. 2).

In the proposed article collection, the majority of slice-based approaches [7,30,39,40,61–66] outperformed scan-based ones [8,10,11,17,27,35,51,60] (Table 2 and Table 3). Nevertheless, it is not so clear how much performance is gained by employing slice-based approaches for both lung nodule diagnosis and cancer histology classification from CT data. Being designed for 2D input data, slice-based approaches might be not well-suited for medical imaging when dealing with volumetric data such as CT scans [78]. Conversely, scan-based approaches can better preserve the spatial information. Doing so would likely result in fairer diagnostic results. However, scan-based approaches using 3D CNNs require large computational resources to be trained. A solution to this problem could be the application of scan-based approaches using hybrid models such as time-distributed 2D CNNs combined with RNN modules, as done by *Moitra et al.* [35] and *Marentakis et al.* [10] for lung cancer histology characterization. From Table 3, it can be noticed that the performance achieved by *Moitra et al.* [35] and *Marentakis et al.* [10] is competitive with the one achieved by the studies that employed slice-based approaches. Hybrid models have the potential to process a sequence of slices as a whole using a much smaller number of parameters than 3D CNNs. Thus, time-distributed 2D CNNs combined with RNN modules can reduce the overall complexity of the model and the required computational/memory power while preserving the information about the spatial coherence of slices, which is an important piece of information that should be taken into account when dealing with CT data.

Regardless the specific CNN type, the main limitation of DL is that supervised-classification algorithms require large amounts of labeled data for analysis [5,52,78]. Labeling medical images is quite challenging, as it requires in-depth domain knowledge [2,48]. Moreover, annotating lots of medical data is a time-consuming and non-trivial procedure [48]. In the lung cancer domain, it is extremely difficult to collect numerous data from patients with the respective pathologically-proven ground truth [52]. As a result, the cardinality of available datasets is relatively small. Hence, there is the need of strategies to learn effectively from small datasets by means of CNNs. There is also the need of joint efforts among hospitals, technical personnel, and society to increase the amount of annotated data while preserving the patients' privacy. Besides these needs, data augmentation/synthetic data generation techniques and/or transfer learning approaches with fine-tuning may be essential to prevent overfitting. In the proposed article collection, the majority of researchers performed data augmentation [7,8,11,30,40,60–62,64] and/or used pre-trained CNNs [7,30,39,40,62,74]. Data augmentation may also be useful to recover class imbalance, which is another common issue to handle. A framework developed with a non-balanced dataset may have the consequence of assigning all items to a single majority class, apparently outperforming other approaches [78]. Reducing class imbalance through data augmentation would destroy the clinical prevalence of classes. Consequently, the real clinical scenario would be neglected. A more clinically-correct approach might be to maintain the clinical prevalence of classes and implement alternative strategies to handle the class imbalance problem. Another data-related challenge relies on data heterogeneity. Different datasets are made up of samples having lots of heterogeneities, mainly due to different medical scanners, acquisition settings, and years. These heterogeneous characteristics constitute one of the leading factors of the low generalizability of CNN-based models. However, if on the one hand heterogeneity is a problem not easy to manage, on the other hand it could make the study less dataset-biased; thus, more

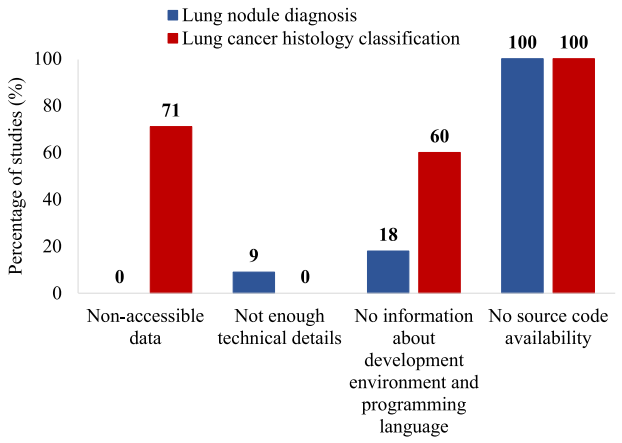

Fig. 7. Percentage of lung nodule diagnosis-related (blue) and lung cancer histology classification-related (red) studies non-entirety reproducible in terms of non-accessibility of data (employment of private datasets), absence of necessary technical details, no information about the development environment and programming language, and non-availability of the source code of the implementation.

realistic. No one among the investigated studies mixed two different datasets in performing the experiments. To manage data heterogeneity, one possible solution could be taking advantage of some specific pre-processing steps, such as resampling to a common pixel-spacing and slice-thickness, HU scale conversion, and intensity normalization prior to data dimensionality standardization. To alleviate the low generalizability, the solution could be developing more flexible methods that can be validated not only on heterogeneous data but even on data coming from different imaging modalities (*i.e.,* multi-modal approaches).

We also observed that several of the investigated studies are not entirely reproducible because they either used a private data collection [7,11,17,39,40], did not give all the necessary technical details (*e.g.,* the input data size [65]), or mentioned neither the development environment nor the programming language [10,39,40,63,64]. Moreover, no research groups made the source code of the implementation available. Making the source code fully accessible under copyright may be advantageous for the entire scientific community, as other researchers could work on it, knowledge could be shared, and significant improvements might be achieved. Fig. 7 displays the percentage of lung nodule diagnosis-related and lung cancer histology classification-related studies not entirety reproducible.

Finally, a very heartfelt challenge relies on the interpretability of the results [2,5]. DL algorithms are considered as black boxes, as their success or failure is hard to interpret [48,78]. Due to their opacity, the interpretation of the model inner states is not as straightforward as with decision trees or standard object-oriented code [79]. However, the interpretability of the results is essential for clinicians in order to cut down any doubts and concerns. Results as well as diagnosis scores without any interpretation do not significantly help them in making the final diagnosis and planning an accurate treatment [5]. One solution could be to add interpretability modules such as Grad-CAM for visual explanations [80], in order to allow the clinicians to understand which regions are the most involved in the decision-making process. Doing so would result in the definition of more accurate and reliable clinical judgments. Among the analyzed studies, only *Shen et al.* [8], *Guo et al.* [11], and *Chaunzwa et al.* [39] allowed a sort of interpretation of the model outputs, highlighting the CT portions where the neural network had focused the most during the decision-making process.

6. Perspectives

From the carried-out investigation, we observed an increasing interest in studies employing CNNs for lung nodule diagnosis and cancer histology classification from CT data. Nevertheless, there is the necessity to explore the lung cancer domain deeper. Scan-based approaches have the potential to make full use of the fundamental information contained in volumetric data such as CT scans, but their success is often subjective to the amount of data and computational resources. Larger and well-annotated publicly-available datasets as well as more scalable architectures and the addition of interpretability modules could help in further improving the performance and the trust towards such automatic decision-making techniques. Furthermore, it could be worthwhile to design a diagnostic system with continuous learning ability, able to support real-time contexts. One possible perspective for developing continuous learning systems is to build a CNN-based model with cloud computing techniques, as done by *Halder et al.* [30]. Implementing a framework with cloud computing would guarantee its machine-independent reproducibility. Clinical records can be delivered to cloud storage in real-time and training datasets can be changed continuously, so that the framework can be trained in an adaptively way on a cloud backend [81].

7. Conclusions

Timely and accurate diagnosis of lung cancer is crucial to decrease the lung cancer-related deaths. To this aim, the hereby presented survey offers a comprehensive investigation of the research trends associated with lung nodule diagnosis and cancer histology classification from CT data adopting CNNs, and shows how CNN-based models are valid learning strategies for such classification tasks. Therefore, combining non-invasive CNN-based classification models with thorax CT scans could have a significant impact in lung cancer diagnosis, targeted treatment, and patient care.

We believe that this survey will be helpful for future studies involved in lung nodule diagnosis and cancer histology classification prior to lung biopsy by means of CNNs.

CRediT authorship contribution statement

Selene Tomassini: Conceptualization, Methodology, Investigation, Writing – original draft. **Nicola Falcionelli:** Methodology, Writing – review & editing. **Paolo Sernani:** Methodology, Writing – review & editing. **Laura Burattini:** Project administration, Supervision. **Aldo Franco Dragoni:** Project administration, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Prabhu, K. Prasad, A. Robels-Kelly, X. Lu, AI-based carcinoma detection and classification using histopathological images: A systematic review, Comput. Biol. Med. (2022) 105209.
[2] P. Monkam, S. Qi, H. Ma, W. Gao, Y. Yao, W. Qian, Detection and classification of pulmonary nodules using convolutional neural networks: A survey, IEEE Access 7 (2019) 78075–78091.
[3] A. Naik, D.R. Edla, Lung nodule classification on computed tomography images using deep learning, Wirel. Pers. Commun. 116 (1) (2021) 655–690.
[4] M. Winkels, T.S. Cohen, Pulmonary nodule detection in CT scans with equivariant CNNs, Med. Image Anal. 55 (2019) 15–26.
[5] W. Cao, R. Wu, G. Cao, Z. He, A comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans, IEEE Access 8 (2020) 154007–154023.
[6] W.J. Sori, J. Feng, A.W. Godana, S. Liu, D.J. Gelmecha, DFD-Net: Lung cancer detection from denoised CT scan image using deep learning, Front. Comput. Sci. 15 (2) (2021) 1–13.
[7] S. Pang, M. Fan, X. Wang, J. Wang, T. Song, X. Wang, X. Cheng, VGG16-T: A novel deep convolutional neural network with boosting to identify pathological type of lung cancer in early stage by CT images, Int. J. Comput. Intell. Syst. 13 (1) (2020) 771.

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

[9] S. Li, D. Liu, Automated classification of solitary pulmonary nodules using convolutional neural network based on transfer learning strategy, J. Mech. Med. Biol. (2021) 2140002.

[10] P. Marentakis, P. Karaiskos, V. Kouloulias, N. Kelekis, S. Argentos, N. Oikonomopoulos, C. Loukas, Lung cancer histology classification from CT images based on radiomics and deep learning models, Med. Biol. Eng. Comput. 59 (1) (2021) 215–226.

[11] Y. Guo, Q. Song, M. Jiang, Y. Guo, P. Xu, Y. Zhang, C.-C. Fu, Q. Fang, M. Zeng, X. Yao, Histological subtypes classification of lung cancers on CT images using 3D deep learning and radiomics, Academic Radiol. (2020).

[12] M.P. Rivera, A.C. Mehta, M.M. Wahidi, Establishing the diagnosis of lung cancer: Diagnosis and management of lung cancer: American college of chest physicians evidence-based clinical practice guidelines, Chest 143 (5) (2013) e142S–e165S.

[13] D.I. Suster, M. Mino-Kenudson, Molecular pathology of primary non-small cell lung cancer, Arch. Med. Res. (2020).

[14] W.D. Travis, E. Brambilla, A.G. Nicholson, Y. Yatabe, J.H. Austin, M.B. Beasley, L.R. Chirieac, S. Dacic, E. Duhig, D.B. Flieder, et al., The 2015 world health organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification, J. Thorac. Oncol. 10 (9) (2015) 1243–1260.

[15] M. Kriegsmann, C. Haag, C.-A. Weis, G. Steinbuss, A. Warth, C. Zgorzelski, T. Muley, H. Winter, M.E. Eichhorn, F. Eichhorn, et al., Deep learning for the classification of small-cell and non-small-cell lung cancer, Cancers 12 (6) (2020) 1604.

[16] Y. Han, Y. Ma, Z. Wu, F. Zhang, D. Zheng, X. Liu, L. Tao, Z. Liang, Z. Yang, X. Li, et al., Histologic subtype classification of non-small cell lung cancer using PET/CT images, Eur. J. Nucl. Med. Mol. Imag. 48 (2) (2021) 350–360.

[17] H. Liu, Z. Jiao, W. Han, B. Jing, Identifying the histologic subtypes of non-small cell lung cancer with computed tomography imaging: A comparative study of capsule net, convolutional neural network, and radiomics, Quant. Imag. Med. Surg. 11 (6) (2021) 2756.

[18] P. Hao, K. You, H. Feng, X. Xu, F. Zhang, F. Wu, P. Zhang, W. Chen, Lung adenocarcinoma diagnosis in one stage, Neurocomputing 392 (2020) 245–252.

[19] A. Panunzio, P. Sartori, Lung cancer and radiological imaging, Curr. Radiopharm. 13 (3) (2020) 238–242.

[20] M.M.N. Abid, T. Zia, M. Ghafoor, D. Windridge, Multi-view convolutional recurrent neural networks for lung cancer nodule identification, Neurocomputing (2021).

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

[22] A. Halder, D. Dey, A.K. Sadhu, Lung nodule detection from feature engineering to deep learning in thoracic CT images: A comprehensive review, J. Digit. Imag. 33 (3) (2020) 655–677.

[23] G. Zhang, Z. Yang, L. Gong, S. Jiang, L. Wang, X. Cao, L. Wei, H. Zhang, Z. Liu, An appraisal of nodule diagnosis for lung cancer in CT images, J. Med. Syst. 43 (7) (2019) 1–18.

[24] S.K. Thakur, D.P. Singh, J. Choudhary, Lung cancer identification: A review on detection and classification, Cancer Metastasis Rev. (2020) 1–10.

[25] F.C. Detterbeck, P.J. Mazzone, D.P. Naidich, P.B. Bach, Screening for lung cancer: Diagnosis and management of lung cancer: American college of chest physicians evidence-based clinical practice guidelines, Chest 143 (5) (2013) e78S–e92S.

[26] R.V. Adiraju, S. Elias, A survey on lung CT datasets and research trends, Res. Biomed. Eng. (2021) 1–16.

[27] H. Yu, J. Li, L. Zhang, Y. Cao, X. Yu, J. Sun, Design of lung nodules segmentation and recognition algorithm based on deep learning, BMC Bioinformatics 22 (5) (2021) 1–21.

[28] W. Alakwaa, M. Nassef, A. Badr, Lung cancer detection and classification with 3D convolutional neural network (3D-CNN), Int. J. Adv. Comput. Sci. Appl. 8 (8) (2017).

[29] T. Pereira, C. Freitas, J.L. Costa, J. Morgado, F. Silva, E. Negrão, B.F. de Lima, M.C. da Silva, A.J. Madureira, I. Ramos, et al., Comprehensive perspective for lung cancer characterisation based on AI solutions using CT images, J. Clin. Med. 10 (1) (2021) 118.

[30] A. Halder, S. Chatterjee, D. Dey, Adaptive morphology aided 2-pathway convolutional neural network for lung nodule classification, Biomed. Signal Process. Control 72 (2022) 103347.

[31] G.D. Rubin, Lung nodule and cancer detection in CT screening, J. Thorac. Imag. 30 (2) (2015) 130.

[32] Y. Gu, X. Lu, L. Yang, B. Zhang, D. Yu, Y. Zhao, L. Gao, L. Wu, T. Zhou, 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 (2018) 220–231.

[33] B. Zhao, Y. Tan, D.J. Bell, S.E. Marley, P. Guo, H. Mann, M.L. Scott, L.H. Schwartz, D.C. Ghiorghiu, Exploring intra-and inter-reader variability in uni-dimensional, bi-dimensional, and volumetric measurements of solid tumors on CT scans reconstructed at different slice intervals, Eur. J. Radiol. 82 (6) (2013) 959–968.

[34] P.F. Pinsky, D.S. Gierada, P.H. Nath, E. Kazerooni, J. Amorosa, National lung screening trial: Variability in nodule detection rates in chest CT studies, Radiology 268 (3) (2013) 865–873.

[35] D. Moitra, R.K. Mandal, Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network, J. Digit. Imag. 33 (4) (2020) 895–902.

[36] E. Bębas, M. Borowska, M. Derlatka, E. Oczeretko, M. Hładuński, P. Szumowski, M. Mojsak, Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis, Biomed. Signal Process. Control 66 (2021) 102446.

[37] L. Cong, W. Feng, Z. Yao, X. Zhou, W. Xiao, Deep learning model as a new trend in computer-aided diagnosis of tumor pathology for lung cancer, J. Cancer 11 (12) (2020) 3615.

[38] D. Planchard, S. Popat, K. Kerr, S. Novello, E. Smit, C. Faivre-Finn, T. Mok, M. Reck, P. Van Schil, M. Hellmann, et al., Metastatic non-small cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up, Ann. Oncol. 29 (2018) iv192–iv237.

[39] T.L. Chaunzwa, A. Hosny, Y. Xu, A. Shafer, N. Diao, M. Lanuti, D.C. Christiani, R.H. Mak, H.J. Aerts, Deep learning classification of lung cancer histology using CT images, Sci. Rep. 11 (1) (2021) 1–12.

[40] S. Pang, Y. Zhang, M. Ding, X. Wang, X. Xie, A deep model for lung cancer type identification by densely connected convolutional networks and adaptive boosting, IEEE Access 8 (2019) 4799–4805.

[41] Y. Gu, J. Chi, J. Liu, L. Yang, B. Zhang, D. Yu, Y. Zhao, X. Lu, A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning, Comput. Biol. Med. 137 (2021) 104806.

[42] S. Wang, D.M. Yang, R. Rong, X. Zhan, J. Fujimoto, H. Liu, J. Minna, I.I. Wistuba, Y. Xie, G. Xiao, Artificial intelligence in lung cancer pathology image analysis, Cancers 11 (11) (2019) 1673.

[43] M. Keshani, Z. Azimifar, F. Tajeripour, R. Boostani, Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system, Comput. Biol. Med. 43 (4) (2013) 287–300.

[44] H. Han, L. Li, F. Han, B. Song, W. Moore, Z. Liang, Fast and adaptive detection of pulmonary nodules in thoracic CT images using a hierarchical vector quantization scheme, IEEE J. Biomed. Health Inf. 19 (2) (2014) 648–659.

[45] J.-k. Liu, H.-y. Jiang, M.-d. Gao, C.-g. He, Y. Wang, P. Wang, H. Ma, et al., An assisted diagnosis system for detection of early pulmonary nodule in computed tomography images, J. Med. Syst. 41 (2) (2017) 1–9.

[46] J. Gong, J.-y. Liu, L.-j. Wang, X.-w. Sun, B. Zheng, S.-d. Nie, Automatic detection of pulmonary nodules in CT images by incorporating 3D tensor filtering with local image feature analysis, Phys. Med. 46 (2018) 124–133.

[47] V.K. Chauhan, K. Dahiya, A. Sharma, Problem formulations and solvers in linear SVM: A review, Artif. Intell. Rev. 52 (2) (2019) 803–855.

[48] J. Ma, Y. Song, X. Tian, Y. Hua, R. Zhang, J. Wu, Survey on deep learning for pulmonary medical imaging, Front. Med. (2019) 1–20.

[49] Y. Zhou, X. Xu, L. Song, C. Wang, J. Guo, Z. Yi, W. Li, The application of artificial intelligence and radiomics in lung cancer, Precis. Clin. Med. 3 (3) (2020) 214–227.

[50] U. Djuric, G. Zadeh, K. Aldape, P. Diamandis, Precision histology: How deep learning is poised to revitalize histomorphology for personalized cancer care, NPJ Precis. Oncol. 1 (1) (2017) 1–5.

[51] J. Zhao, C. Zhang, D. Li, J. Niu, Combining multi-scale feature fusion with multi-attribute grading, a CNN model for benign and malignant classification of pulmonary nodules, J. Digit. Imag. 33 (4) (2020) 869–878.

[52] S. Zhang, F. Han, Z. Liang, J. Tan, W. Cao, Y. Gao, M. Pomeroy, K. Ng, W. Hou, An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets, Comput. Med. Imaging Graph. 77 (2019) 101645.

[53] P. Sahu, D. Yu, M. Dasari, F. Hou, H. Qin, A lightweight multi-section CNN for lung nodule classification and malignancy estimation, IEEE J. Biomed. Health Inf. 23 (3) (2018) 960–968.

[54] S. Tomassini, N. Falcionelli, P. Sernani, H. Müller, A.F. Dragoni, An end-to-end 3D convLSTM-based framework for early diagnosis of Alzheimer's disease from full-resolution whole-brain sMRI scans, in: 2021 IEEE International Symposium on Computer-Based Medical Systems, IEEE, 2021, pp. 74–78.

[55] K. Mehta, A. Jain, J. Mangalagiri, S. Menon, P. Nguyen, D.R. Chapman, Lung nodule classification using biomarkers, volumetric radiomics, and 3D CNNs, J. Digit. Imag. (2021) 1–20.

[56] H. Jung, B. Kim, I. Lee, J. Lee, J. Kang, Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method, BMC Med. Imag. 18 (1) (2018) 1–10.

[57] H. Peng, H. Sun, Y. Guo, 3D multi-scale deep convolutional neural networks for pulmonary nodule detection, Plos One 16 (1) (2021) e0244406.

[58] P.S. Mittapalli, V. Thanikaiselvan, Multiscale CNN with compound fusions for false positive reduction in lung nodule detection, Artif. Intell. Med. 113 (2021) 102017.

[59] H. Zhang, Y. Peng, Y. Guo, Pulmonary nodules detection based on multi-scale attention networks, Sci. Rep. 12 (1) (2022) 1–14.

[60] H. Liu, H. Cao, E. Song, G. Ma, X. Xu, R. Jin, C. Liu, C.-C. Hung, Multi-model ensemble learning architecture based on 3D CNN for lung nodule malignancy suspiciousness classification, J. Digit. Imag. 33 (5) (2020) 1242–1256.

[61] M. Al-Shabi, H.K. Lee, M. Tan, Gated-dilated networks for lung nodule classification in CT scans, IEEE Access 7 (2019) 178827–178838.
[62] I. Ali, M. Muzammil, I.U. Haq, A.A. Khaliq, S. Abdullah, Efficient lung nodule classification using transferable texture convolutional neural network, IEEE Access 8 (2020) 175859–175870.
[63] C.-J. Lin, S.-Y. Jeng, M.-K. Chen, Using 2D CNN with Taguchi parametric optimization for lung cancer recognition from CT images, Appl. Sci. 10 (7) (2020) 2591.
[64] P. Zhai, Y. Tao, H. Chen, T. Cai, J. Li, Multi-task learning for lung nodule classification on chest CT, IEEE Access 8 (2020) 180317–180327.
[65] S.R. Jena, S.T. George, D.N. Ponraj, Lung cancer detection and classification with DGMM-RBCNN technique, Neural Comput. Appl. (2021) 1–17.
[66] X. Lu, Y. Nanehkaran, M. Karimi Fard, A method for optimal detection of lung cancer based on deep learning optimized by marine predators algorithm, Comput. Intell. Neurosci. 2021 (2021).
[67] Y. Li, D. Chen, X. Wu, W. Yang, Y. Chen, A narrative review of artificial intelligence-assisted histopathologic diagnosis and decision-making for non-small cell lung cancer: Achievements and limitations, J. Thorac. Dis. 13 (12) (2021) 7006.
[68] F. Xing, Y. Xie, H. Su, F. Liu, L. Yang, Deep learning in microscopy image analysis: A survey, IEEE Trans. Neural Netw. Learn. Syst. 29 (10) (2017) 4550–4568.
[69] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nature Commun. 5 (1) (2014) 1–9.
[70] W. Wu, C. Parmar, P. Grossmann, J. Quackenbush, P. Lambin, J. Bussink, R. Mak, H.J. Aerts, Exploratory study to identify radiomics classifiers for lung cancer histology, Front. Oncol. 6 (2016) 71.
[71] C. Shen, Z. Liu, M. Guan, J. Song, Y. Lian, S. Wang, Z. Tang, D. Dong, L. Kong, M. Wang, et al., 2D and 3D CT radiomics features prognostic performance comparison in non-small cell lung cancer, Transl. Oncol. 10 (6) (2017) 886–894.
[72] X. Zhu, D. Dong, Z. Chen, M. Fang, L. Zhang, J. Song, D. Yu, Y. Zang, Z. Liu, J. Shi, et al., Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer, Eur. Radiol. 28 (7) (2018) 2772–2778.
[73] H. Liu, B. Jing, W. Han, Z. Long, X. Mo, H. Li, A comparative texture analysis based on NECT and CECT images to differentiate lung adenocarcinoma from squamous cell carcinoma, J. Med. Syst. 43 (3) (2019) 59.
[74] S.R. Digumarthy, A.M. Padole, R.L. Gullo, L.V. Sequist, M.K. Kalra, Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? Medicine 98 (1) (2019).
[75] F. Yang, W. Chen, H. Wei, X. Zhang, S. Yuan, X. Qiao, Y.-W. Chen, Machine learning for histologic subtype classification of non-small cell lung cancer: A retrospective multicenter radiomics study, Front. Oncol. 10 (2020).
[76] R. Thawani, M. McLane, N. Beig, S. Ghose, P. Prasanna, V. Velcheti, A. Madabhushi, Radiomics and radiogenomics in lung cancer: A review for the clinician, Lung Cancer 115 (2018) 34–41.
[77] R. Reiazi, E. Abbas, P. Famiyeh, A. Rezaie, J.Y. Kwan, T. Patel, S.V. Bratman, T. Tadic, F.-F. Liu, B. Haibe-Kains, The impact of the variation of imaging parameters on the robustness of computed tomography radiomic features: A review, Comput. Biol. Med. 133 (2021) 104400.
[78] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, P.H. Abreu, Using deep learning techniques in medical imaging: A systematic review of applications on CT and PET, Artif. Intell. Rev. 53 (6) (2020) 4093–4160.
[79] Z. Salahuddin, H.C. Woodruff, A. Chatterjee, P. Lambin, Transparency of deep neural networks for medical image analysis: A review of interpretability methods, Comput. Biol. Med. 140 (2022) 105111.
[80] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: Visual explanations from deep networks via gradient-based localization, in: 2017 IEEE International Conference on Computer Vision, IEEE, 2017, pp. 618–626.
[81] A. Masood, P. Yang, B. Sheng, H. Li, P. Li, J. Qin, V. Lanfranchi, J. Kim, D.D. Feng, Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT, IEEE J. Transl. Eng. Health Med. 8 (2019) 1–13.

Selene Tomassini received the M.S. degree in biomedical engineering in December 2018 from Università Politecnica delle Marche, Ancona, Italy, where she is currently pursuing the Ph.D. degree in information engineering. Her research interests mainly include deep learning algorithms for biomedical signal, image, and video analysis.

Nicola Falcionelli received the Ph.D. degree in information engineering in April 2020 from Università Politecnica delle Marche, Ancona, Italy, where he is currently a post-doc research fellow. His research interests range in the spectrum of artificial intelligence, from symbolic approaches to statistical learning techniques.

Paolo Sernani received the Ph.D. degree in information engineering in March 2016 from Università Politecnica delle Marche, Ancona, Italy, where he is currently a post-doc research fellow. His research interests include deep learning for image and video analysis, multi-agent systems, expert systems, and decision-support systems.

Laura Burattini is currently in charge as full professor and president of the unified council of the biomedical engineering degree at Università Politecnica delle Marche, Ancona, Italy, where she teaches "bioengineering" and "biomedical signal and data processing". Her scientific interests range in the spectrum of biomedical engineering, focusing mainly on biomedical signal and image analysis.

Aldo Franco Dragoni is currently in charge as associate professor at Università Politecnica delle Marche, Ancona, Italy, where he teaches "artificial intelligence", "dedicated operating systems", and "fundamentals of computer sciences". His scientific interests concern several aspects of artificial intelligence, from knowledge-based approaches to advanced hybrid systems.