Journal of Medical Systems (2019) 43: 181
https://doi.org/10.1007/s10916-019-1327-0

IMAGE & SIGNAL PROCESSING

An Appraisal of Nodule Diagnosis for Lung Cancer in CT Images

Guobin Zhang[1] · Zhiyong Yang[1] · Li Gong[1] · Shan Jiang[1,2] · Lu Wang[1] · Xi Cao[1] · Lin Wei[1] · Hongyun Zhang[1] · Ziqi Liu[1]

Received: 23 February 2019 / Accepted: 8 May 2019 / Published online: 15 May 2019

Abstract

As "the second eyes" of radiologists, computer-aided diagnosis systems play a significant role in nodule detection and diagnosis for lung cancer. In this paper, we aim to provide a systematic survey of state-of-the-art techniques (both traditional techniques and deep learning techniques) for nodule diagnosis from computed tomography images. This review first introduces the current progress and the popular structure used for nodule diagnosis. In particular, we provide a detailed overview of the five major stages in the computer-aided diagnosis systems: data acquisition, nodule segmentation, feature extraction, feature selection and nodule classification. Second, we provide a detailed report of the selected works and make a comprehensive comparison between selected works. The selected papers are from the IEEE Xplore, Science Direct, PubMed, and Web of Science databases up to December 2018. Third, we discuss and summarize the better techniques used in nodule diagnosis and indicate the existing future challenges in this field, such as improving the area under the receiver operating characteristic curve and accuracy, developing new deep learning-based diagnosis techniques, building efficient feature sets (fusing traditional features and deep features), developing high-quality labeled databases with malignant and benign nodules and promoting the cooperation between medical organizations and academic institutions.

Keywords Lung cancer · Computer-aided diagnosis system · Nodule diagnosis · CT images · Nodule classification

Introduction

Cancer is a leading public health problem across the world, and a report provided by the American Cancer Society (ACS) estimates that approximately 1,735,350 cancer cases will occur in 2018 in the United States [1]. Lung cancer is the leading cause of cancer deaths in both males and females, with an estimated death rate of 26% and 25%, respectively, and a 5-year relative survival rate of 18% [1, 2]. One major reason for the low survival rate is that approximately 70% of patients miss timely and effective treatment at the early stage [3]. Fortunately, the early diagnosis of lung cancer will greatly increase this rate to 52% [4].

To data, several imaging modalities (as shown in Fig. 1), such as computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET), have been used by radiologists for lung cancer diagnosis. Among them, the most sensitive diagnostic modality to diagnose lung cancer at the early stage is low-dose CT because of its high spatial resolution, cost-effectiveness, wide availability and noninvasiveness [5, 6]. With the advancement of the CT technique, the mortality of lung cancer has significantly decreased [7].

Pulmonary nodules are rounded-opacity and circumscribed parenchymal lesions measuring less than 3 to 4 cm in diameter [8]. However, not all pulmonary nodules are malignant (as shown in Fig. 2). In clinical practice, the diagnosis of pulmonary nodules is still performed by radiologists based on the nodules' information (texture, intensity and shape, etc.). They need to analyze a large number of CT images and differentiate malignant nodules from benign ones efficiently and precisely. However, this is extremely subjective and challenging. Due to destabilizing factors, such as misinterpretation, experience and distraction, as well as perceptual error, it is difficult for radiologists to precisely classify benign and malignant pulmonary nodules in CT images. Currently, invasive biopsy shows the potential to determine the status of pulmonary nodules. This is a highly invasive procedure because biopsies need to

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Shan Jiang
shanjmri@tju.edu.cn

[1] School of Mechanical Engineering, Tianjin University, Tianjin 300350, China

[2] Centre for advanced Mechanisms and Robotics, Tianjin University, 135 Yaguan Road, Jinnan District, Tianjin 300350, China

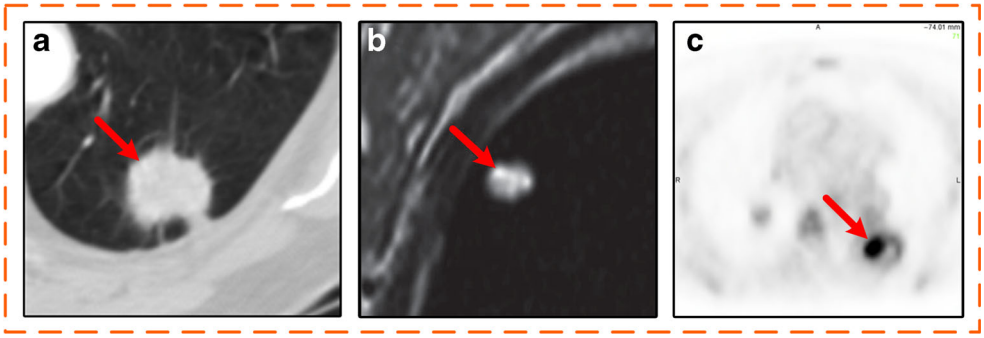

Fig. 1 Different imaging modalities of pulmonary nodules. **a** CT, **b** MRI, **c** PET

be taken from the nodules. Additionally, only extracting a small portion of the nodule may not accurately characterize the nodule due to its spatial heterogeneity. Therefore, advanced pulmonary nodule detection and diagnosis systems are needed to help radiologists overcome these challenges, interpret diagnostic data and make decisions. Such systems are known as computer-aided diagnosis (CAD) systems.

The systematic investigation of CAD schemes was begun in the 1980s to detect, segment and diagnose pulmonary nodules with CT images [9–11]. CAD systems can also be referred to as two more specific computational systems: computer-aided detection (CADe) systems and computer-aided diagnosis (CADx) systems. CADe systems detect potential lesions through medical images and find special abnormalities. The CADx systems mainly aim to characterize and distinguish the lesions. There are many reviews about CADe systems [12–20] in recent years, but few about CADx systems. The existence of pulmonary nodules is not definitely equivalent to lung cancer, and the features of nodules have a complicated relationship with cancer. Therefore, the diagnosis of lung cancer requires an investigation of each suspicious nodule carefully and the integration of all information on all nodules. As "the second eyes" of radiologists, CADx systems can provide a second objective opinion to improve the performance of diagnosis by decreasing interobserver variation. Specifically, the CADx systems have already shown potential assistants in clinical practice. Hence, this work only takes into account the CADx systems. To enhance the confidence of radiologists in the CADx system as a useful tool, some rigid goals must be met, such as high area under the receiver operating characteristic curve (AUC), higher accuracy (ACC), high sensitivity and specificity, computational efficiency, low cost and software security assurance [21].

Several reviews have reported many techniques employed in the diagnosis of pulmonary nodules with CT images. Those papers introduced the databases and techniques that had been adopted at the time and compared experimental results between the proposed method and other referenced methods; see Yang et al. [22], Hu et al. [23] and Paulraj et al. [24]. Hu et al. [23] reviewed studies through July 2017, Yang et al. [22] reviewed studies through November 2018, and Paulraj et al. [24] reviewed studies through May 2016. The studies reviewed in both Yang et al. [22] and Hu et al. [23] were only based on deep learning techniques; Paulraj et al. [24] only described the various techniques used in the CADx system and did not compare different CADx systems. The main contribution of our work is to provide a comprehensive review to introduce the state-of-the-art techniques (both traditional techniques and deep learning techniques) used for pulmonary nodules diagnosing in CADx systems with CT images. The listed papers are selected from the IEEE Xplore, Science Direct, PubMed, and Web of Science databases up to December 2018.

This paper provides an appraisal of nodule diagnosis for lung cancer in CT images. We mainly focus on the proposed methods, the databases adopted and the experimental results in each selected work. Most selected papers are reported in detail. To better explore the current trends in this field, a quantitative comparison is made based on many assessment

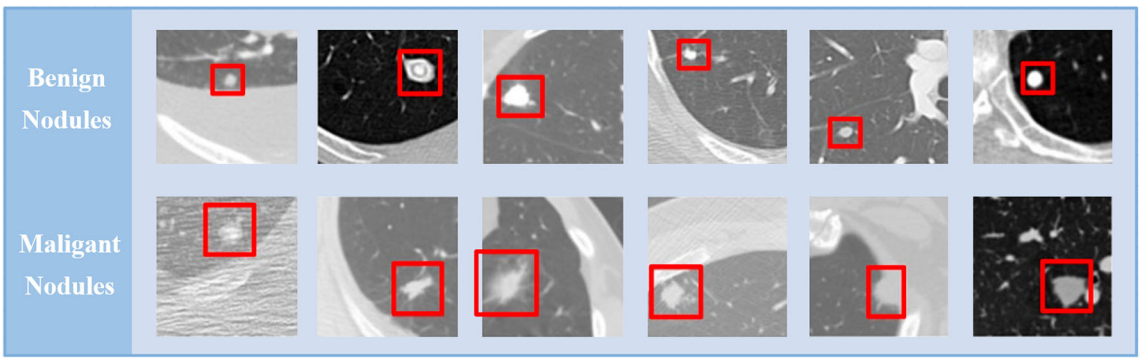

Fig. 2 Exhibition of pulmonary nodules in CT images

metrics (AUC, ACC, sensitivity, etc.). Moreover, we summarize the most popular database and feature set and the best traditional method and deep learning method for nodule classification. Further research and challenges are also reported in the Discussion. Therefore, this review will be useful for researchers to learn about the latest and most advanced techniques of pulmonary nodule diagnosis.

This review is organized as follows: Section 2 introduces the most popular structure of CADx systems, especially the five basic stages of the systems. Section 3 provides a detailed report of the selected works and makes a comprehensive comparison between selected works based on six assessment metrics. Section 4 mainly discusses and summarizes the better techniques used in nodule diagnosis and indicates the existing future challenges in this field. Finally, Section 5 concludes this paper.

Structure of the CADx system

Evidence suggests that CADx can improve the performance of radiologists in medical image interpretation and diagnosis [25] by providing quantitative support and decreasing the misinterpretation of the available data [26–28]. As potential assistants, CADx systems are usually composed of five basic stages: data acquisition, nodule segmentation, feature extraction, feature selection and nodule classification. The basic stages for a popular CADx system are shown in Fig. 3. The systems start with data acquisition. Subsequently, the nodules can be localized or segmented according to the annotations from the database. Then, pattern features are extracted and selected from the segmented nodules. From the selected features, the nodules are classified as benign or malignant. We organize this review based on the five basic stages to help readers learn about the principle and structure of the CADx systems.

Data acquisition

Among medical imaging modalities, low-dose CT is the most popular in pulmonary nodule diagnosis. A well-characterized public repository of thoracic CT scans can facilitate and stimulate the development of CADx systems. Data obtained from private databases or location hospitals are not recommended because the comparison between experimental results from different CADx systems is less persuasive. For better development, training and comparison of different CADx systems, some public databases have been built. Some public databases have achieved good performance, such as Lung Image Database Consortium (LIDC) [29, 30], Lung Image Database Consortium and Image Database Resource Initiative (LIDC–IDRI) [31, 32], LUng Nodule Analysis 2016 (LUNA16) [33] and Early Lung Cancer Action Program (ELCAP) [34, 35].

In April 2000, the National Cancer Institute (NCI) issued a request for applications (RFAs) entitled “Lung Image Database Resource for Imaging Research” to convene five institutions to form the LIDC. Guided by the premise that “public-private partnerships are essential to accelerating scientific discovery for human health” and their successes in this realm [36], the Image Database Resource Initiative (IDRI) was created by the Foundation for the National Institutes of Health (FNIH) to further develop the LIDC. Through the IDRI, eight medical imaging companies provided additional resources to expand the LIDC database, and the database that combined those CT scans was referred to as the LIDC-IDRI database [37]. The LIDC-IDRI database further advanced by the FNIH with active participation from the Food and Drug Administration (FDA) and became the largest open-source pulmonary nodule image database in the world [38]. The LIDC-IDRI database contains 1018 CT scans with slice thicknesses varying from 1.25 to 3 mm and reconstruction intervals between 0.625 and 3 mm. For each scan, an XML file is included, containing the marking results of a two-phase image annotation process. The XML files were made by four professional radiologists and include certain characteristics (roundness, texture, malignancy, etc.), the location and the contour of each nodule. Additionally, all nodules were marked by each radiologist independently and are divided into three categories: “nodule < 3 mm,” “nodule > or = 3 mm,” and “non-nodule > or = 3 mm” initially. Each nodule is indicated with one of 5 malignancy levels: Level 1: highly unlikely for cancer; Level 2: moderately unlikely for cancer; Level 3: intermediate likelihood; Level 4: moderately suspicious for cancer; Level 5: highly suspicious for cancer. The LUNA16 database was extracted from the LIDC-IDRI database, which contains 888 CT scans with relevant annotations. The ELCAP database was built in 2003 to evaluate the performance of different CAD systems with 50 CT scans and related annotations.

Nodule segmentation

The accurate segmentation of pulmonary nodules is an essential step for the subsequent feature extraction step. Nodule segmentation mainly aims to remove irrelevant information (trachea, bronchi, pulmonary vessels, etc.) and recover significant information during data acquisition. To segment the nodules, many works used the annotation supplied by the four different radiologists to extract nodule contours. It is interesting to note the different degrees of variation in the annotations provided by different radiologists, as shown in Fig. 4. Filho et al. [39] and Costa et al. [40] chose the largest area to represent the instance of the nodules. Kumar et al. [41] used annotations of all radiologists to represent one nodule for better classification accuracy. Different techniques have been proposed to segment pulmonary nodules in this review: optimal threshold [42–44], Otsu algorithm [45, 46], pixel-based

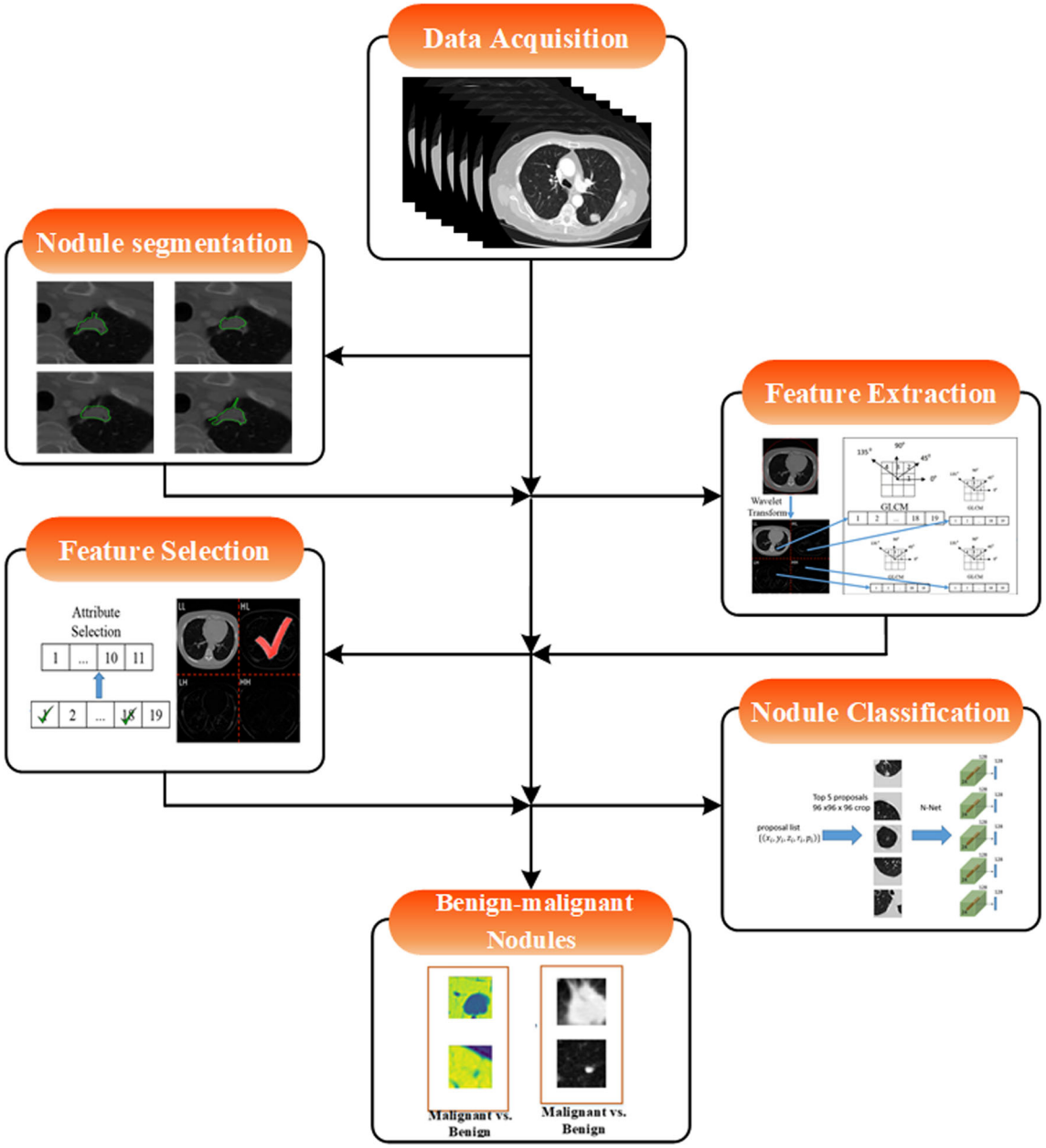

Fig. 3 The basic stages of a popular CADx system

segmentation method [47], Watershed transform [48], Hough transform [49], improved random walk algorithm [50], 3D region-growing algorithm and multiscale 3D dot-filtering algorithm [51], and thresholding and connected component analysis [52]. In particular, the segmentation of pulmonary nodules is not needed for pulmonary nodule diagnosis in most

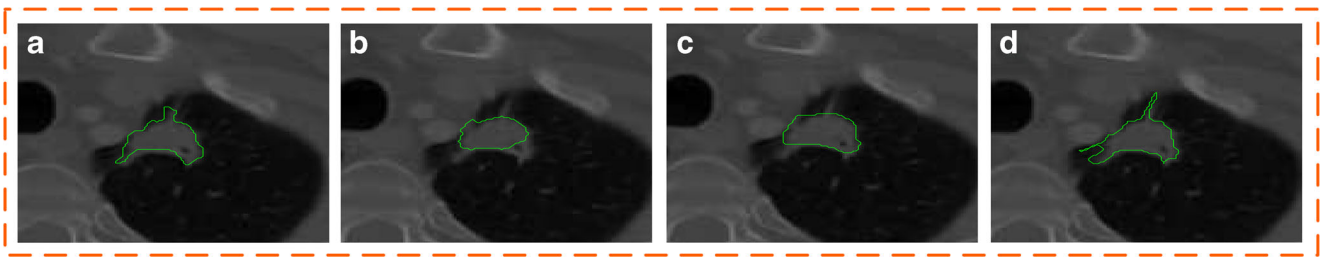

Fig. 4 Examples of annotations provided by four different radiologists for the same nodule in a single slice of CT from the LIDC-IDRI database [41]

deep learning techniques. Small 3D or 2D patches containing nodules will be extracted from the raw lung scans and input to the network individually to learn deep features, so only the center of the pulmonary nodule is necessary.

Feature extraction and selection

After segmentation, the pulmonary nodules are subjected to the feature extraction stage. Feature extraction is an important step in representing pulmonary nodules. Generally, a large number of both 2D and 3D image features are computed for all images, and a subset of the features with the best diagnostic performance is selected for the final classification. The obtained features are mainly divided into two types: deep features and traditional features [53]. Deep features are extracted by deep neural networks, and traditional features (such as shape, texture, intensity, size, margin and morphology, etc.) are mainly calculated by feature descriptors. Feature selection is also essential to obtain the most representative features to improve the classification accuracy. The main purposes of feature extraction and selection are as follows: (1) reducing the dimensionality of the input data and removing irrelevant information, the probability of which increases with the dimension of the feature space; (2) reducing the chances of overfitting, the probability of which increases with the dimension of the feature space [54].

A useful and efficient feature set plays a significant role in nodule diagnosis. With the advancement of image processing techniques, a number of image features have been explored to characterize pulmonary nodules. As shown in Fig. 5, the features used most often include deep features [53, 55–79], texture-based features [39–42, 45, 47, 49–53, 69, 71, 73, 80–84], shape (geometry)-based features [39, 42, 43, 47, 48, 50, 51, 53, 64, 69, 71, 82, 84,], intensity-based features [42–44, 50, 51, 80, 82, 83, 85, 86], size-based features [66, 67, 71, 80, 84], margin-based features [48, 52, 86] and morphology-based features [73, 80, 86].

Of all features, deep features are the most important because deep neural networks can learn richer features in more detail. Apart from deep features, texture-based features are the most widely used to represent pulmonary nodules because texture-based features can reflect the inner structure of nodules efficiently and offer much information to describe the nodules in a CT scan. Gray-level cooccurrence matrix (GLCM) features [42, 50–53, 62, 71], local binary pattern (LBP) features [50, 69, 82], Haralick features [52, 71, 73, 81], run-length features [62] and wavelet features [49, 62] are useful texture-based features for lung cancer analysis. Histograms of oriented gradients (HOGs) [69] and Fourier shape descriptors [53] have been found to be powerful features for shape (geometry)-based classification. Intensity-based features are often used as the primary source of image information, such as the mean gradient of the region boundary, average intensity, and density uniformity. Size-based features (aspect ratio and area of bounding box, etc.) and margin-based features (margin sharpness and margin gradient, etc.) are the most fundamental and intuitive features. The circularity and ratio of the semiaxis are typical morphological features. Moreover, several works [53, 64, 71] utilized hybrid-feature descriptors for pulmonary nodule classification. Xie et al. [53] fused texture, shape and deep model-learned information to build a learning model. Yuan et al. [64] fused the statistical features from multi-view multi-scale convolutional neural networks (CNNs) and geometrical features from Fisher vector (FV) encodings to classify pulmonary nodule types. Kaya et al. [71] combined a group of handcrafted features (morphology, color and texture) and deep features from CNN for nodule classification.

Additionally, several works [39, 46, 73] used a genetic algorithm (GA) for feature selection to obtain the best learning model and reduce training time. The features used in those

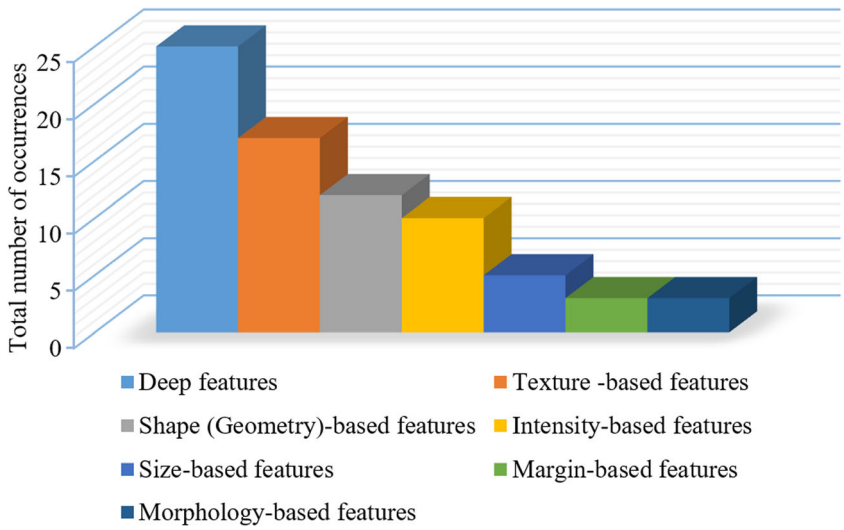

Fig. 5 Features used in the works for feature extraction and selection

works are based on texture, morphology or shape without deep feature because GA is less efficient on high-dimensional features [87].

Nodule classification

Once the pulmonary nodules are detected and segmented from the CT images, the next task is to determine whether the detected nodules are malignant or benign. This step mainly focuses on differentiating malignant nodules from benign ones efficiently and precisely. Before training the learning model, the malignant and benign samples in the training database should be classified accurately. However, the classification criteria for malignant and benign nodules are not very rigorous. As we mentioned above, each nodule was indicated with 5 malignancy levels in the LIDC-IDRI database. Most of the selected works carried out a binary classification by calculating the average level from four radiologists for each nodule and defined that if the average level was higher than 3, this nodule was considered to be malignant; if the average level was lower than 3, it was considered to be benign; and if the average level was equal to 3, it is considered to be an uncertain nodule and discarded. Two works [70, 72] denoted nodules with levels 1 to 2.5 as benign and levels 3.5 to 5 as malignant and eliminated level 3 nodules. Wang et al. [73] studied uncertain data intensively in the classification. Moreover, Liu et al. [75] and Nishio et al. [79] developed a ternary classification (benign, primary malignant, and metastatic malignant) to develop better CADx systems.

All features extracted from malignant and benign nodules will be fed to the classifier for training. To better organize the presentation of the related works, we divided the classifiers into two categories: traditional classifiers and deep learning classifiers. The main traditional classifiers and deep learning classifiers used for nodule classification are listed in Tables 1 and 2, respectively. It is obvious that support vector machine (SVM) is the most popular traditional classifier and most deep learning classifiers are based on CNN. Moreover, we find that a deep learning classifier is more frequently used than a traditional classifier. However, it is worth noting that there are also some limitations in adopting deep learning for pulmonary nodule classification: (1) A large amount of training data is necessary for nodule diagnosis using supervised learning schemes because a small amount of training data may lead to overfitting and convergence issues. (2) The ratio of positive to negative cases in some public databases is highly imbalanced. (3) It is quite time consuming to train a deep and accurate model, even with the support of powerful GPU hardware. Therefore, researchers should pay more attention to the existing limitations and adopt corresponding solutions, such as data augmentation, balancing positive and negative samples, and cropping the desired size of target objects.

Detailed report of selected works

This paper provides an appraisal of nodule diagnosis for lung cancer in CT images. To present a comprehensive appraisal, the selected papers are taken from the IEEE Xplore, Science Direct, PubMed, and Web of Science databases up to December 2018. We applied the following inclusion criteria to optimize the survey: pulmonary nodule detection works were removed, nodule diagnosis methods with CT images were selected, repeated works were removed, and innovative works were chosen. Additionally, some effective keywords were selected for further survey, such as lung cancer, computer-aided diagnosis system, nodule diagnosis, CT images, medical image processing, deep learning technique and nodule classification. Finally, 58 articles were selected from the last four years. Most of them are reported in detail in the following paragraphs. We compared the performances of different nodule diagnosis techniques with six assessment metrics: (1) nodule size, (2) number of nodules used, (3) AUC, (4) ACC, (5) sensitivity and (6) specificity. Equations (1), (2) and (3) calculate the ACC, sensitivity and specificity, respectively. Where the *TP* is true positive, *TN* is true negative, *FP* is false positive and *FN* is false negative. These assessment metrics are mentioned in most of the selected papers. The comparison of the selected works is reported in Table 3.

$$ACC = \frac{TP + TN}{TP + FN + FP + FN} \tag{1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

In 2018, Costa et al. [40] proposed a method for the classification of malignant and benign lung nodules. They adopted the mean phylogenetic distance (MPD) and taxonomic diversity index as texture descriptors. As in [39], GA in conjunction with the SVM was applied to select the best training model. Experiments on the LIDC-IDRI database with 1405 (1011 benign and 394 malignant) nodules showed promising results, with the best sensitivity of 93.42%, specificity of 91.21%, ACC of 91.81%, and AUC of 0.94.

Filho et al. [45] developed a method for pulmonary nodule diagnosis based on CT images. Two types of techniques were used in the proposed method: image processing and pattern recognition techniques. First, each pulmonary nodule was segmented into three internal subregions using the Otsu algorithm. Next, the indexes' basic taxic weights and standardized taxic weights based on topology were used as texture descriptors for the extraction of nodule characteristics. Finally, a CNN was employed to perform the classification task. The proposed method was validated on a set of 1405 pulmonary

Table 1 Traditional classifiers used for nodule classification

Authors	Traditional classifier(s)
Akram et al. [43], Akram et al. [44], Orozco et al. [49], Dhara et al. [52], Farag et al. [82], Chen et al. [85]	SVM
Li et al. [50], Ferreira et al. [88]	Random forest (RF) algorithm
Costa et al. [40]	GA and SVM
Gong et al. [51]	SVM, naïve Bayes classifier and linear discriminant analysis
Kaya et al. [71]	Cascaded classifiers and stacking methods
Wei et al. [81]	Local kernel regression models (LKRM)/(Linear regression regularization term) LR, LKRM/Global kernel regression regularization term (KR)
Nishio et al. [89]	Gradient tree boosting (XGBoost)
Filho et al. [39], Filho et al. [46]	GA and SVM
Sweetlin et al. [47]	Ant colony optimization (ACO)-Rough dependency measure (RDM) and SVM
Wu et al. [42]	Relevance vector machine (RVM)
Firmino et al. [48]	Rule-based classifier and SVM
Kumar et al. [41]	Autoencoder and binary decision tree
Kaya et al. [84]	Rule based method

Table 2 Deep learning classifiers used for nodule classification

Authors	Deep learning classifier(s)
Filho et al. [45], Tu et al. [61], Wang et al. [69], Nishio et al. [79]	CNN
Liu et al. [72]	Deep learning architecture named dense convolutional binary-tree network (DenseBTNet)
Wang et al. [73]	Semi-supervised extreme learning machine (SS-ELM)
Zhang et al. [76]	Spatial pyramid dilated network
Zhu et al. [67]	Dual path networks (DPN) and gradient boosting machines (GBM)
Xie et al. [53]	AdaBoosted back propagation neural network (BPNN)
Shen et al. [55]	Multi-crop convolutional neural networks (MC-CNN)
Silva et al. [57]	CNN and GA
Sun et al. [62]	LeCun's model-based CNN, deep belief networks (DBN) and stacked denoising autoencoder (SDAE)
Chen et al. [63]	Convolutional autoencoder neural network (CANN)
Ciompi et al. [65]	Multi-stream multi-scale convolutional networks (ConvNets)
Liao et al. [68]	Deep neural network and leaky noisy-or gate
Zhao et al. [70]	Hybrid CNN based on LeNet and AlexNet
Nibali et al. [74]	Deep residual networks (ResNet)
Yan et al. [77]	Dual-modal supervised deep autoencoder (DSDAE) framework
Liu et al. [80], Jung et al. [78]	CNN and ensemble models
Tajbakhsh et al. [56]	Massive-training artificial neural networks (MTANNs) and CNN
Sun et al. [58]	CNN, DBNs and SDAE
Cheng et al. [59]	SDAE
Hua et al. [60]	DBN, CNN, scale invariant feature transform (SIFT) and fractal analysis
Liu et al. [75]	Multi-view convolutional neural networks (MV-CNN)
Hancock et al. [90]	Statistical learning methods
Farag et al. [82]	K-nearest neighbor (KNN)
Mao et al. [91]	Single-center classifier and multicenter Classifier

Table 3 Comparison of the selected works

Authors	Year	Database	Size (mm)	N° of nodules	AUC	ACC (%)	Sensitivity (%)	Specificity (%)
Costa et al. [40]	2018	LIDC-IDRI	3–30	1405	0.94	91.81	93.42	91.21
Li et al. [50]	2018	LIDC and private	NI	1300	0.95	90	92	83
Gong et al. [51]	2018	Private and NSCLC	NI	234	0.94, 0.90, 0.99	NI	NI	NI
Kaya et al. [71]	2018	LIDC	NI	439	NI	84.70	67.37	95.46
Wei et al. [81]	2018	LIDC-IDRI	NI	746	NI	85.0 ± 3.4, 85.4 ± 3.4	NI	NI
Chen et al. [85]	2018	Private	NI	75	NI	84	92.85	72.73
Nishio et al. [89]	2018	TCIA	NI	99	0.896	82.0	NI	NI
Filho et al. [45]	2018	LIDC-IDRI	3–30	1405	NI	93.47	92.63	90.7
Liu et al. [72]	2018	LIDC-IDRI	3–30	2001 samples	0.9360	89.50	NI	NI
Wang et al. [73]	2018	LIDC-IDRI	NI	1439	0.961	95.91	95.56	95.03
Zhang et al. [76]	2018	LIDC	NI	353	0.883	88.6	86.3	90.3
Zhu et al. [67]	2018	LIDC-IDRI	3–30	1004	NI	90.44	NI	NI
Filho et al. [39]	2017	LIDC-IDRI	3–30	1405	0.9252	NI	93.1	92.26
Sweetlin et al. [47]	2017	NI	NI	390	NI	94.36	96.69	92.35
Ferreira et al. [88]	2017	LIDC	3–30	1171	0.858	80.0	70.2	85.6
Xie et al. [53]	2017	LIDC-IDRI	≥3	2669	0.9665 ± 0.0001	89.53 ± 0.09	84.19 ± 0.09	92.02 ± 0.01
Shen et al. [55]	2017	LIDC-IDRI	3–30	2618	0.93	87.14	NI	NI
Silva et al. [57]	2017	LIDC-IDRI	3–30	3243	0.949	94.78	94.66	95.14
Sun et al. [62]	2017	LIDC-IDRI	≥3	134,668 samples	0.899 ± 0.018, 0.884 ± 0.022, 0.852 ± 0.025	NI	NI	NI
Chen et al. [63]	2017	Private	NI	4500 scans	0.98	95.00	NI	NI
Ciompi et al. [65]	2017	MILD and DLCST	NI	1991	NI	78.9	NI	NI
Liao et al. [68]	2017	LUNA and DSB	≥6	832	0.87	NI	NI	NI
Wang et al. [69]	2017	LIDC-IDRI	3–30	2910 samples	0.9702	91.75	NI	NI
Zhao et al. [70]	2017	LIDC-IDRI	≥3	743	0.877	82.23	NI	NI
Nibali et al. [74]	2017	LIDC-IDRI	3–42	831	0.9459	89.90	91.07	88.64
Yan et al. [77]	2017	Private	NI	2810	0.94	92.81 ± 0.57	91.75 ± 1.53	NI
Liu et al. [80]	2017	NLST and ELCAP	5–14	326	0.732, 0.780	NI	NI	NI
Wu et al. [42]	2016	LIDC	>3	273	NI	79.4	72.7	NI
Akram et al. [43]	2016	LIDC	3–30	836	0.9975 ± 0.0020	99.95 ± 0.03	100 ± 0	99.90 ± 0.07
Filho et al. [46]	2016	LIDC-IDRI	>3	1405	NI	93.19	92.75	93.33
Firmino et al. [48]	2016	LIDC-IDRI	3–30	1109	0.91, 0.80, 0.72, 0.67, 0.83	NI	NI	NI
Dhara et al. [52]	2016	LIDC-IDRI	3–30	891	0.9505, 0.8822, 0.8488,	NI	89.73, 82.89, 76.14	86.36, 80.73, 74.91
Tajbakhsh et al. [56]	2016	Private	3–29	489	0.8806, 0.7755	NI	NI	NI
Sun et al. [58]	2016	LIDC-IDRI	3–30	174,412 samples	NI	79.76, 81.19, 79.29	NI	NI
Cheng et al. [59]	2016	LIDC	3–30	1400	0.984 ± 0.015	94.4 ± 3.2	90.8 ± 5.3	98.1 ± 2.2
Liu et al. [75]	2016	LIDC-IDRI	NI	4294 lesions	0.981	NI	90.49	99.91
Hancock et al. [90]	2016	LIDC	≥3	2817 samples	0.949 ± 0.007	88.08 ± 1.11	NI	NI
Kumar et al. [41]	2015	LIDC	≥3	4323	NI	75.01	83.35	NI
Akram et al. [44]	2015	LIDC	3–30	84 Scans	0.9967	96.54	96.31	96.77
Orozco et al. [49]	2015	ELCAP and LIDC	2–30	106 scans	NI	82	90.90	73.91

NI Not informed

nodules with diameters between 3 and 30 mm from the LIDC-IDRI database, including 1011 benign 394 malignant nodules, and it achieved a sensitivity of 92.63%, a specificity of 90.7% and an ACC of 93.47%.

Li et al. [50] proposed an improved Random Forest (RF) algorithm to classify benign and malignant pulmonary nodules in CT images. Intensity-based, texture-based and geometry-based features were extracted from the segmented nodules to characterize nodules. Finally, the extracted features were combined to generate an effective feature vector and fed into an improved RF classifier for classification. A total of 1000 nodules were randomly selected from the LIDC database to train the classifier, and 300 nodules from the General Hospital of Guangzhou Military Command (GHGMC) dataset were used to evaluate the classifier. The experiments on the LIDC dataset and the GHGMC dataset demonstrated that the proposed method achieved promising results, with a sensitivity of 0.92, a specificity of 0.83, an ACC of 0.90 and an AUC of 0.95.

Gong et al. [51] mainly aimed to develop a CADx system for pulmonary nodule classification and explored the effects of different image training datasets on CADx scheme performance. A total of 243 nodules from the private database and the NSCLC database were involved in this study, and all nodules were divided into three datasets: all nodules, benign and stage I malignant nodules, and benign and stage III malignant nodules. A set of 66 3D heterogeneity-based, shape-based and texture-based features were extracted, and 10 features were selected by the Relief-F feature selection method [92]. Three machine learning models, SVM, naïve Bayes classifier and linear discriminant analysis, were applied to train and test using a leave-one-case-out (LOCO) cross-validation method [93]. The experimental results showed that three classifiers achieved an average AUC of 0.94, 0.90 and 0.99 using three different datasets, respectively.

Zhu et al. [67] designed two 3D DPNs for fully automatic pulmonary nodule detection and classification in CT images. For nodule detection, a U-net-like encoder-decoder structure was used to learn nodule features, and a 3D Faster R-CNN was designed for nodule detection. For nodule classification, GBM with 3D DPN features was proposed. Dual path connection integrated the advantages of residual learning [94] (the ability of feature reuse) and dense connection [95] (the ability to exploit new features) into a unified structure that obtained success on the ImageNet dataset. For nodule classification, a GBM with 3D constructed features (deep dual path features, nodule size and raw nodule CT pixels) was proposed. The proposed classification subnetwork was validated on a set of 1004 nodules from the LIDC-IDRI database with 10-fold patient-level cross validation. The results showed that DPN achieved an ACC of 90.44%. Additionally, the classification subnetwork achieved a nodule-level diagnosis ACC of 92.74% and a patient-level diagnosis ACC of 82.31%.

Kaya et al. [71] proposed cascaded classifiers and three stacking approaches for pulmonary nodule classification. In their study, deep features and hand-crafted features were used to define the nodules. The deep features were obtained from two different areas: the first area was the region of the nodule itself, and the second was the bounding box area of the nodule. The deep features were extracted using Alexnet [96], and the hand-crafted features were extracted based on nodule size, shape, and texture defined by Kaya and Can [84]. Both deep features and hand-crafted features were fed into a CNN for training. In the first step of classification, several base classifiers (SVM, KNN, Adaboost, and RF) were used to classify nodule characteristics. In the second step, the results obtained from the first step were combined for malignancy classification. For better classification, stacking methods were employed. The experimental results showed that the cascaded classifier achieved the best classification ACC of 84.70%, sensitivity of 67.37%, and specificity of 95.46%.

Liu et al. [72] presented a novel end-to-end deep learning architecture named DenseBTNet for pulmonary nodule diagnosis in CT images. Not only did the study introduce center-crop operation into the DenseNet, but the DenseBTNet could split isolated transition layers of the DenseNet and merge them with dense blocks, then adjust the feature-map transition mode to compact the model. The proposed method was tested on the LIDC-IDRI database with 2001 (1361 benign and 640 malignant) samples using five-fold cross validation. The experimental results showed that DenseBTNet-C (the DenseBTNet with compression factor less than 1) achieved the best performance, with an ACC of 89.50%, and DenseBTNet-BC (both bottleneck and compression transition layers were used in DenseBTNet) achieved the highest AUC of 0.9360.

Wang et al. [73] proposed a CADx system for pulmonary nodule classification based on SS-ELM. The authors pointed out that treating uncertain class data as benign or malignant class was unreasonable. To verify this idea, the feature model was established based on 14 Haralick features and 20 morphological features. In addition, GA was employed for feature selection to obtain promising experimental results and reduce training time. After that, the feature model was put into four classifiers: ELM, SVM, probabilistic neural network (PNN) [97] and multilayer perceptron (MLP) [98] for better comparison. The experiments were based on 1439 nodules from the LIDC-IDRI database using 5-fold cross validation. Additionally, compared with ELM, the pulmonary nodule CAD system based on SS-ELM achieved the best performance with an ACC of 95.91%, a sensitivity of 95.56%, a specificity of 95.03% and an AUC of 0.961.

Zhang et al. [76] proposed a novel 3D spatial pyramid dilated convolution network for the classification of malignant and benign pulmonary nodules. The advantage of the proposed network was that using dilated convolution alleviated the loss of small bits of information and feature map

resolution. The proposed method was validated on the LIDC database with 353 nodules using 5-fold cross validation, and it achieved the best ACC of 88.6%, sensitivity of 86.3%, specificity of 90.3% and AUC of 0.883.

Wei et al. [81] proposed a novel spectral clustering (SC) algorithm [99, 100] LKRM with out-of-sample extension to differentiate unlabeled benign and malignant lung nodules. In the algorithm, a linear regression regularization term (LR) or a global kernel regression regularization term (KR) was integrated into the Laplacian matrix to tackle the out-of-sample problem. The Haralick texture features were calculated to represent each nodule. The proposed method was evaluated on a set of 746 (371 benign and 375 malignant) nodules from the LIDC-IDRI database. From the testing clustering performance, LKRM/LR achieved an average ACC of 0.850 ± 0.034, and LKRM/KR achieved an average ACC of 0.854 ± 0.034.

Nishio et al. [89] mainly aimed to evaluate a conventional CADx system for pulmonary nodule classification. In this study, SVM and XGBoost were used to compare the classification performance. Additionally, the Tree Parzen Estimator (TPE) was used for Bayesian optimization for parameters of SVM and XGBoost. The experiment was performed on the Cancer Imaging Archive (TCIA) with 99 (62 malignant and 37 benign) nodules. The results showed that the XGBoost classifier achieved better performance with an ACC of 0.82 and an AUC of 0.896.

In 2017, Filho et al. [39] proposed a method to differentiate between the patterns of malignant and benign pulmonary nodules in CT images. For better feature extraction, phylogenetic diversity was used by means of particular indexes: intensive quadratic entropy, extensive quadratic entropy, average taxonomic distinctness, total taxonomic distinctness, and pure diversity indexes. Finally, the GA proposed by [101] was employed to select the most significant features, and the SVM was used for classification of pulmonary nodules into malignant and benign. In the test stage, a set of 1405 (394 malignant and 1011 benign) nodules from the LIDC database with diameters between 3 and 30 mm were applied. The proposed method achieved an ACC of 92.52%, a sensitivity of 93.1%, and a specificity of 92.26%.

Sweetlin et al. [47] designed a CADx system to diagnose pulmonary nodules. After nodule segmentation, texture-based, shape-based and run length-based features were extracted from the nodules. Subsequently, the cosine similarity measure (CSM) and RDM with ACO were used to select two subsets of features. Then, the selected features were fed to two classifiers, namely, SVM and NB, to train the learning model using 10-fold cross validation. Finally, 390 nodules were used for the experiment. From the results, it was obvious that the SVM classifier with the feature subsets chosen by ACO-RDM achieved the best ACC of 94.36%, sensitivity of 96.69% and specificity of 92.35%.

Shen et al. [55] proposed an end-to-end MC-CNN for pulmonary nodule malignancy suspiciousness classification using CT images. A multi-crop pooling operation was proposed to produce multi-scale features. The nodule semantic attributes (subtlety and margin) and nodule diameter were also characterized to potentially assist researchers in modeling nodule malignancy. A total of 2618 nodules with diameters between 3 and 30 mm from the LIDC-IDRI database were used to evaluate the MC-CNN performance. The proposed method achieved promising results in both classification ACC (87.14%) and AUC score (0.93).

Silva et al. [57] proposed a deep learning technique jointly with a GA to classify pulmonary nodules as malignant or benign without computing the traditional features. The first step was to divide each nodule into two subregions with maximum inter-class variance using the Otsu algorithm [102] based on particle swarm optimization (PSO) [103]. The following steps were performed using each two-dimensional CT slice as an individual sample and resizing all nodules' and subregions' slices to 28×28. Finally, the diagnosis was completed using the evolutionary CNN. The proposed method was validated on a set of 21,631 nodules (sizes between 3 and 30 mm) from the LIDC-IDRI database, and it achieved the best sensitivity of 94.66%, specificity of 95.14%, ACC of 94.78% and AUC of 0.949.

Tu et al. [61] presented a CADx system to automatically categorize solid, part-solid and non-solid pulmonary nodules in CT images using a CNN. The implementation of the CNN model was based on 2D regions of interest (ROIs) due to the high variation of slice thickness (1.25-3 mm) in the LIDC database. A total of 570 nodules with diameters greater than 3 mm from the LIDC database were used, and a 10-times 10-fold cross validation was performed in the experiment. In both classification and regression tasks, the proposed models achieved promising results with the all-slice selection strategy.

Sun et al. [62] mainly aimed to compare the performance of three multichannel ROI-based deep learning algorithms (LeCun's model-based CNN [104, 105], DBN and SDAE) for automatic lung cancer diagnosis in CT images. For better comparison purposes, a traditional CADx system was also used based on hand-crafted features, including density features, texture features and morphological features. All the algorithms were tested on 134,668 samples from the LIDC-IDRI database, and a 10-fold cross-validation method based on cases was also applied to completely separate the training data and testing data. The results showed that CNN achieved the great AUC (0.899 ± 0.018), which was significantly higher than DBN (0.884 ± 0.022), SDAE (0.852 ± 0.025) and traditional CADx (0.848 ± 0.026).

Chen et al. [63] proposed a CANN [106] to learn unsupervised image features for pulmonary nodule classification. The main contribution was that the proposed learning model could be trained with a large amount of unlabeled data and a small

amount of labeled data. The method was tested on 4500 CT images, and it achieved a promising ACC of 95.00% and an AUC of 0.98. The proposed method achieved a classification rate of 93.1% (911 out of 979) and 93.9% (647 out of 689) in the LIDC-IDRI and ELCAP databases, respectively.

Ciompi et al. [65] proposed a deep learning system based on multi-stream multi-scale ConvNets for the automatic classification of six pulmonary nodule types (solid, non-solid, part-solid, calcifed, perifssural and spiculated). The proposed system consisted of nine streams of ConvNets, and all streams were grouped into three sets of three streams, one for each considered scale (namely, 10 mm, 20 mm and 40 mm for patch size). The proposed learning model was trained on 943 patients with 1352 nodules from the Multicentric Italian Lung Detection (MILD) trial and validated on 468 patients with 639 nodules from the Danish Lung Cancer Screening Trial (DLCST). For better comparison, SVM was trained based on the raw intensity features and features learned from raw data via an unsupervised learning approach [107]. The results showed that ConvNets achieved better performance than SVM, with an average ACC of 78.9%.

Liao et al. [68] designed a 3D deep neural network to automatically detect all suspicious pulmonary nodules and evaluate the whole pulmonary malignancy with CT images. In the detection stage, a 3D region proposal network (RPN) using a modified U-net [108] was employed as the backbone model. The block used in this network was a 3D residual block, and each block was composed of three residual units. For nodule classification, the system selected the top five nodules based on detection confidence and evaluated their malignancy with a leaky noisy-or gate [109]. A total of 832 nodules from the LUNA16 and Data Science Bowl 2017 (DSB) were used to train and validate the proposed model. The AUCs of the predicted cancer probability achieved on the training and test sets were 0.90 and 0.87.

Wang et al. [69] mainly proposed a hybrid learning model by integrating the traditional features and deep CNN-based features to improve the risk differentiation of benign and malignant pulmonary nodules in CT images. HOG [110] features were used to build a shape descriptor and LBP [111] features were used to enhance the statistical characteristics of nodules. Finally, all the features were fed into a CNN to build a hybrid CNN network model. To evaluate the performance of the proposed model, 2910 samples expanded by image inversion from the LIDC-IDRI database were used. Experimental results showed that the proposed hybrid fusion CNN model achieved the highest AUC of 0.9702 and ACC of 91.75%.

Zhao et al. [70] proposed a new Agile CNN framework to distinguish benign from malignant pulmonary nodules using CT images. The proposed framework was composed of the layer settings of LeNet [105] and the parameter settings of AlexNet [112]. The CNN was evaluated on a set of 743 nodules with diameters greater than 3 mm from the LIDC-IDRI database using tenfold cross-validation. The experimental results showed that the Agile CNN achieved an ACC of 82.23% and an AUC of 0.877 when the kernel size was set to 7×7, the learning rate was 0.005, the batch size was 32, and dropout and Gaussian initialization were used.

Nibali et al. [74] proposed a modified ResNet [94] with curriculum learning [113] and transfer learning for pulmonary nodule malignancy classification. The modified ResNet accepted three 2D planar view inputs instead of the full 3D volume, and it was modified to be "fully convolutional" without any fully connected layers. The modified ResNet was validated on a set of 831 pulmonary nodules (diameters between 3 and 42 mm) from the LIDC-IDRI database, and it achieved 89.90% ACC, 0.9459 AUC, 91.07% sensitivity and 88.64% specificity.

Xie et al. [53] proposed a method for pulmonary nodule classification based on texture, shape and deep model-learned information (Fuse-TSD). A deep convolutional neural network (DCNN) was employed to extract deep features, a GLCM was used to extract texture features, and a Fourier descriptor of the nodule boundary was used to characterize the heterogeneity of the nodule shape. Finally, the AdaBoosted BPNN algorithm was adopted to build an ensemble classifier. The proposed algorithm was evaluated on the three LIDC-IDRI datasets with 2669 nodules using 10-fold cross validation, and it achieved the best AUC of $(96.65 \pm 0.01)\%$, ACC of $(89.53 \pm 0.09)\%$, sensitivity of $(84.19 \pm 0.09)\%$ and specificity of $(92.02 \pm 0.01)\%$.

Liu et al. [80] proposed a 3D CNN to classify pulmonary nodule malignancy in low-dose chest CT scans. In their study, two 3D CNN architectures, CNN1 and CNN2, were designed. Meanwhile, several ensembles of the 3D CNN and traditional models were designed to find the potential performance using hybrid features. For the validation of the proposed method, a total of 326 pulmonary nodules (sizes between 5 and 14 mm) from the National Lung Cancer Screening Trial (NLST) and ELCAP were used. Two CNN models were trained and evaluated using 5-fold cross validation. The results showed that both the 3D CNN model and the ensemble models achieved better performance than only traditional models, with an AUC of 0.732 and 0.780, respectively.

Ferreira et al. [88] mainly aimed to classify malignant and benign pulmonary nodules based on texture and margin sharpness features in CT images. In this study, texture features were extracted using a cooccurrence matrix obtained from the nodule volume, and margin sharpness was extracted from perpendicular lines drawn over the borders on all nodule slices. Three different algorithms (statistical significance analysis, correlation-based filtering method, and a wrapper) were used for feature extraction, and seven traditional machine learning algorithms (KNN, SVM and NB, etc.) were used for classification. Classification was performed on the LIDC database with 1171 nodules with the diameters between 3 and 30 mm

using a stratified tenfold cross-validation. The experimental results showed that all extracted features combined with the RF classifier and an unbalanced dataset achieved the best performance, with 0.858 AUC, 80.0% ACC, 70.2% sensitivity and 85.6% specificity.

Yan et al. [77] developed a DSDAE framework based on extreme learning for nodule diagnosis in PET/CT images. In the framework, the extreme learning machine (ELM) was used as weight shaping method to optimize autoencoder framework. Highly discriminative features were learnt automatically from raw data, all features were fused using decision level fusion (DLF) and feature level fusion (FLF). In the experiment, 1600 pulmonary nodules from the private database were used with a 5-fold cross validation and the proposed method achieved an ACC of (92.81 ± 0.57) %, an AUC of 0.94, and a sensitivity of (91.75 ± 1.53) %.

In 2016, Firmino et al. [48] developed a new system for nodule detection and classification in CT images. To reduce false positives, a HOG was used to extract nodule features, and SVM was performed for classification. The likelihood of malignancy was divided into five degrees: highly unlikely, moderately unlikely, indeterminate, moderately suspicious and highly suspicious [114]. The system was evaluated on 420 patients from the LIDC-IDRI database with 1109 nodules. Compared to the FLD and Gaussian Naive Bayes (NB), the SVM showed competitive performance with an AUC of 0.91.

Dhara et al. [52]. developed a scheme for nodule classification by combining shape, margin and texture features based on CT images. First, pulmonary nodules were segmented using the method proposed in [115]. Then, a total of 57 features were extracted and 49 features were selected in the classification scheme. Finally, the proposed classification scheme was validated on a set of 891 nodules (diameters between 3 and 30 mm) from the LIDC-IDRI database and it achieved an AUC of 0.9505, 0.8822, and 0.8488 in configuration 1, configuration 2 and configuration 3, respectively.

Tajbakhsh et al. [56] compared the performance of MTANNs and CNNs in nodule detection and classification. A total of 489 (76 malignant and 413 benign) nodules with diameters between 3 and 29 mm were used. For nodule classification, the MTANNs yielded an AUC of 0.8806, which was significantly higher than the CNNs with an AUC of 0.7755.

Sun et al. [58] firstly reported the purely data driven approach to classify pulmonary nodules in CT images. They implemented and compared the performance of three different deep learning algorithms (CNN, DBNs and SDAE) with the traditional image features-based CADx system. The experiment was performed on 174,412 samples from the LIDC-IDRI database. The results showed that DBNs achieved the best ACC of 81.19%.

Cheng et al. [59] employed SDAE [116] for the classification of benign and malignant pulmonary nodules in CT images. A total of 1400 nodules were extracted from the LIDC database and 10 times of 10-fold cross validations were conducted to illustrate the performance of the SDAE-based CADx system. The experimental results showed that the proposed system achieved an ACC of (94.4 ± 3.2)%, an AUC of (98.4 ± 1.5)%, a sensitivity of (90.8 ± 5.3)%, and a specificity of (98.1 ± 2.2)%.

Liu et al. [75] developed a MV-CNN to classify pulmonary nodules in CT images. A total of 764 benign lesions and 3530 malignant lesions from the LIDC-IDRI database were used for binary classification experiment. The proposed network achieved 90.49% sensitivity, 99.91% specificity and 0.981 AUC.

Filho et al. [46] proposed a method for nodule diagnosis in CT images. After the segmentation of nodules, a descriptor based on shape features was employed. GA was used to select the best learning model. Finally, all the features were fed into SVM for classification. The proposed method was tested on a set of 1405 (394 malignant and 1011 benign) nodules from the LIDC–IDRI database and it achieved 93.19% ACC, 92.75% sensitivity and 93.33% specificity.

In 2015, Kumar et al. [41] proposed a CADx system to classify pulmonary nodules as either malignant or benign. After nodule extraction, a stacked autoencoder was used to learn deep features. All the deep features were fed into a binary decision tree for nodule classification. The proposed system was evaluated on a set of 4323 nodules from the LIDC database and it achieved an ACC of 75.01% and a sensitivity of 83.35%.

Orozco et al. [49] proposed a CADx system for pulmonary nodule classification. The GLCM [117, 118] was used to extract the texture information of the lung nodules and finally, 11 features were selected and fed into SVM for nodule classification. In the experiment, a total of 106 CT scans from the ELCAP and LIDC database were used. Among them, 61 CT scans (36 with malignant nodules and 25 without nodules) were used for the training and 45 CT scans (23 with malignant nodules and 22 without nodules) were used for testing. The proposed CADx system achieved 82% ACC, 90.90% sensitivity and 73.91% specificity.

Hua et al. [60] utilized a deep learning framework based on two techniques (CNN and DBN) for nodule classification in CT images. Additionally, SIFT [119, 120] and fractal analysis [121] were implemented for comparison. Classification of pulmonary nodules was tested on the LIDC database containing 2545 nodules. The results showed that the DBN achieved the best sensitivity of 73.4% and specificity of 82.2%.

Discussion and future work

This review aims to report the recent techniques used in pulmonary nodule diagnosis with CT images since early diagnosis can improve the effectiveness of treatment and increase the patient survival rate. From Table 3, we find that the AUC of all the selected works ranges from 0.67 to 0.9975, the ACC

ranges from 75.01% to 99.95%, the sensitivity ranges from 67.37% to 100% and the specificity ranges from 72.73% to 99.9%. Some selected works achieved an AUC of more than 0.95, such as Li et al. [50] (0.96), Gong et al. [51] (0.99), Wang et al. [73] (0.961), Xie et al. [53] (0.9665), Chen et al. [63] (0.98), Wang et al. [69] (0.9702), Akram et al. [43] (0.9975), Dhara et al. [52] (0.9505), Cheng et al. [59] (0.984), Liu et al. [75] (0.981) and Akram et al. [44] (0.9967), and several others [43, 44, 51] achieved an promising AUC of more than 99%. Several works [41, 42, 45–47, 49, 58, 65, 67, 71, 81, 85] did not mention an AUC. Costa et al. [40] (91.81%), Filho et al. [45] (93.47%), Wang et al. [73] (95.91%), Zhu et al. [67] (90.44%), Sweetlin et al. [47] (94.36%), Silva et al. [57] (94.78%), Chen et al. [63] (95.00%), Wang et al. [69] (91.75%), Yan et al. [77] (92.81%), Akram et al. [43] (99.95%), Filho et al. [46] (93.19%), Cheng et al. [59] (94.4%) and Akram et al. [44] (96.54%) achieved a good ACC of higher than 90%. Additionally, the sensitivity of [39, 40, 43–47, 49, 50, 57, 59, 73, 75, 77, 85] was greater than 90%, and the specificity of [39, 40, 43–47, 53, 57, 59, 71, 73, 75, 76] was greater than 90%. [44, 85, 89] were validated with less pulmonary nodules.

Based on the literature research, a comparison of AUC, ACC, sensitivity and specificity was made from 2016 to 2018, as shown in Fig. 6. The number of samples in 2015 was too small, so we removed them. The data obtained from private databases were also eliminated because the authors mentioned that using private databases was less persuasive and may even prevent the replication of the results. From Fig. 6, we note that a significant improvement was presented in AUC, and an increasing trend was shown in ACC. One of the main reasons is the development of deep learning techniques.

The major advantage of deep learning is the ability to learn highly discriminative features from various sources of verified data. Initially, 2D CNN, DBN and MTANN were applied for lung nodule classification. These networks were relatively shallow, and the classification accuracies achieved were not promising. With the development of deep learning techniques, an increasing number of networks have been optimized based on CNNs, such as 2D MV-CNN [75] and 3D MC-CNN [55]. The multi-view strategy used in MV-CNN can help the model learn discriminative features to different degrees. The MC-CNN introduced the multi-crop pooling strategy to capture nodule salient information. In addition, the input data were transformed from 2D to 3D, which brought a much greater computational cost. However, compared with the 2D networks, 3D networks can encode richer spatial information from CT images to learn more distinguishable features, so they are more appropriate for the lung nodule classification task. For better classification performance, the network trend is to go deeper since deeper CNNs are thought to have increased representational power. To address this issue, some competitive networks, such as ResNet and DenseNet, were created for the nodule classification task. The skip connection adopted in 2D ResNet (depth = 18) in [74] encourages feature reuse, but it is not efficient in exploring new features. The 3D DenseBTNet (depth = 50) proposed in [72] inherited the properties of DenseNet and improved the classification performance with the center-crop operation, but it might suffer from feature redundancy due to its densely connected mechanism. To integrate the advantages of DenseNet for exploring new features and ResNet for feature reuse, 3D DPN (depth = 92) was designed to classify lung nodules in [67]. It should be noted that the classification accuracy achieved by DPN (90.44%) is higher than both ResNet (89.90%) and DenseBTNet (89.50%). In summary, we find that the developing trends of deep learning techniques used for nodule classification use 3D input data to encode richer spatial information of nodules and develop deeper or wider networks to characterize nodules.

The success of deep learning techniques in natural scene image classification stimulates the study of adopting them in medical image processing, such as ResNet and DenseNet. The

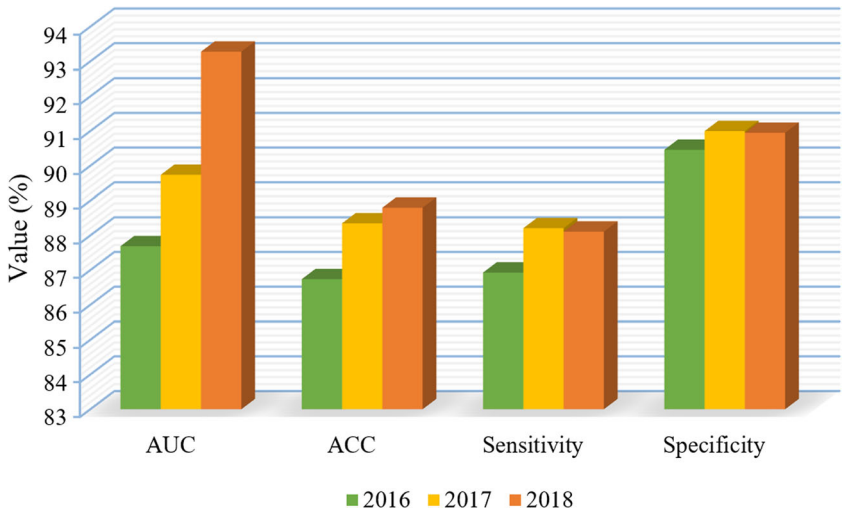

Fig. 6 The comparison of AUC, ACC, sensitivity and specificity from 2016 to 2018

squeeze-and-excitation network (SENet) [123], which encourages feature recalibration by explicitly modeling interdependencies between channels, was created for image classification in the ImageNet Large Scale Visual Recognition Competition (ILSVRC) and won first place. By combining the aggregated residual transformations (ResNeXt) with the advantage of feature reuse, it can improve the classification performance considerably. To the best of our knowledge, the effectiveness of SENet on lung nodule classification has not been extensively explored.

Nine selected works showed potential in the diagnosis of pulmonary nodules with promising performance. Three works [40, 43, 44] adopted traditional features and traditional classifiers for pulmonary nodule diagnosis. Wector et al. [40] applied texture-based features for nodule classification and obtained a promising AUC of 0.94 and an ACC of 91.81%. An important factor that led to good results was using different phylogenetic tree architectures for feature extraction. Akram et al. [43] combined shape (geometry)-based and intensity-based features to build a CADx system and achieved 0.9975 AUC and 99.95% ACC. Akram et al. [44] used intensity-based features for nodule classification and showed great performance, with a high AUC (0.9967) and ACC (96.54%). The classifiers used in those three works were SVM. Four works [57, 59, 63, 77] adopted deep features and deep learning classifiers to differentiate malignant nodules from benign ones. Silva et al. [57] developed an evolutionary CNN for feature extraction and nodule classification. Yan et al. [77] extracted the high-level discriminative features of nodules from the DSDAE framework and acquired good results (0.94 AUC and 92.81% ACC). Cheng et al. [59] achieved a great AUC of 0.984 and an ACC of 94.4%. The satisfactory results mainly benefited from using SDAE for nodule classification. Chen et al. [63] used deep features to characterize pulmonary nodules and achieved a high AUC of 0.98 and an ACC of 95.00%. However, private databases are not allowed to be used to compare different approaches. Two works [69, 73] used hybrid features and deep learning classifiers to classify malignant and benign nodules. Wang et al. [69] achieved promising results with a high AUC of 0.9702 and an ACC of 91.75%. The satisfactory results benefited from fusing deep features extracted from the CNN and texture-based and shape (geometry)-based features to characterize malignant and benign nodules. Wang et al. [73] achieved an AUC of 0.961 and an ACC of 95.91%. One reason for the very high AUC and ACC is the fact that the nodules were represented by deep features and texture-based and morphology-based features, and SS-ELM was employed for classification.

Although some traditional CADx systems showed satisfyingperformance, extracting and choosing a useful feature set is still time-consuming and complex. In addition, the traditional feature extraction depends greatly on lung segmentation, which leads to a low level of automation. The segmentation accuracy might also affect the classification performance. Even if the finalized feature set is able to achieve good classification accuracy, the performance is still uncertain when using other testing databases because the traditional systems are sensitive to small variations of nodule types and extracted feature types. Deep learning techniques are a biologically inspired trainable architecture that can learn multi-level hierarchies of features, which might have the potential to maintain reliable performance on different databases. From all the selected works, we found that the average AUC achieved by the deep learning-based method is 0.904, which is higher than the AUC achieved by the traditional method, with an average of 0.889. The deep learning method also has higher ACC than the traditional method (average of 88.31% vs. 87.58%). Deep learning-based methods have emerged as promising approaches for pulmonary nodule diagnosis, and we believe that deep learning techniques will achieve a breakthrough based on its advantages in lung cancer diagnosis.

In our study, a total of 20 papers employed the LIDC-IDRI as their database for training and testing with the maximum usage rate. In feature extraction, deep features showed satisfactory performance in nodule characterization. Texture-based, shape (geometry)-based and intensity-based features also achieved good results in nodule representation. For nodule classification, SVM was the most popular traditional classifier. SDAE achieved the best AUC (0.984) among deep learning classifiers. Several CNN-based methods (MC-CNN, ResNet and MV-CNN, etc.) also showed high AUC and ACC.

Designing an efficient CADx system for lung cancer has great significance because patients can receive timely and effective treatment at the early stage, thereby increasing the survival rate. In this review, some of the selected works achieved promising performance in the diagnosis of pulmonary nodules. However, many limitations still exist such as low AUC, low ACC, low sensitivity and specificity, high cost, less image data and low calculation efficiency. To further develop CADx systems' contribution to lung cancer treatment, the systems should be improved as follows:

(i) Develop new deep learning techniques to diagnose nodules for lung cancer, such as by building deeper and wider networks based on SENet to extract more rich deep features, adaptively recalibrating channel-wise feature responses by explicitly modeling interdependencies between channels, enhancing useful features and selectively suppressing less useful features.
(ii) Optimize existing techniques to enhance the ability of nodule diagnosis, such as by enhancing the ability to segment different types of nodules and building efficient feature sets (fusing traditional features and deep features) for better nodule classification.
(iii) Design an efficient CADe system to detect pulmonary nodules accurately, improve the ability to detect

different types of nodules such as small nodules (sizes less than 5 mm), irregularly shaped nodules, isolated nodules, and juxta-pleural or juxta-vascular nodules.

(iv) Develop large and high-quality labeled databases with malignant and benign pulmonary nodules for training and testing.

(v) Strengthen and promote communication between medical organizations and academic institutions; combine the knowledge of professional radiologists and computer analysis to enhance the accuracy of CADx systems.

Conclusion

In this paper, we aim to provide a systematic survey of nodule diagnosis for lung cancer with CT images. This review introduces a detailed report of the five major stages in the CADx systems: data acquisition, nodule segmentation, feature extraction, feature selection and nodule classification. A detailed report of the selected works is organized, and a comprehensive comparison between selected works is performed. Moreover, this review summarizes the better techniques used in nodule diagnosis. Further research and challenges are also reported in the discussion of future work. Therefore, this review is very useful for researchers to further learn about the latest and advanced techniques of pulmonary nodule diagnosis.

Acknowledgements This study was funded by the National Natural Science Foundation of China (grant number 81871457), National Natural Science Foundation of China (grant number 51811530310), National Natural Science Foundation of China (grant number 51775368) and Science and Technology Planning Project of Guangdong Province, China (grant number 2017B020210004) and the Science and Technology Project of Tianjin (grant number 18YFZCSY01300).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Siegel, R. L., Miller, K. D., and Jemal, A., Cancer statistics, 2018. Ca-Cancer J Clin 60(5):277–300, 2018.
2. Siegel, R., Naishadham, D., and Jemal, A., Cancer statistics, 2013. Ca-Cancer J Clin, 2013.
3. McGuire, S., World cancer report 2014. World Health Organization 7(2):418–419, 2015.
4. Henschke, C. I., Mccauley, D. I., Yankelevitz, D. F., Naidich, D. P., Mcguinness, G., Miettinen, O. S., Libby, D. M., Pasmantier, M. W., Koizumi, J., and Altorki, N. K., Early Lung Cancer Action Project: overall design and findings from baseline screening. Cancer-Am Cancer Soc 354(9173):2474–2482, 1999.
5. Gibaldi, A., Barone, D., Gavelli, G., Malavasi, S., and Bevilacqua, A., Effects of Guided Random Sampling of TCCs on Blood Flow Values in CT Perfusion Studies of Lung Tumors. Acad. Radiol. 22(1):58–69, 2015.
6. Ng, Q. S., and Goh, V., Angiogenesis in non-small cell lung cancer: imaging with perfusion computed tomography. J Thorac Inag 25(2):142, 2010.
7. Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., Fagerstrom, R. M., Gareen, I. F., Gatsonis, C., Marcus, P. M., and Sicks, J. D., Reduced lung-cancer mortality with low-dose computed tomographic screening. New Engl J Med 365(5):395–409, 2011.
8. Gould, M. K., Maclean, C. C., Kuschner, W. G., Rydzak, C. E., and Owens, D. K., Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. Jama-J Am Med Assoc 285(7):914–924, 2001.
9. Doi, K., Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Comput. Med. Imaging Graph. 31(4):198–211, 2007.
10. Ma, Z., Tavares, J. M. R. S., and Jorge, R. M. N., A review on the current segmentation algorithms for medical images. In: 1st International Conference on Imaging Theory and Applications (IMAGAPP), INSTICC Press, Lisbon, 135–140, 2015.
11. Ma, Z., Tavares, J. M. R. S., Jorge, R. N., and Mascarenhas, T., A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. Comput Method Biomec 13(2):235–246, 2010.
12. Zhang, J. J., Xia, Y., Cui, H. F., and Zhang, Y. N., Pulmonary nodule detection in medical images: A survey. Biomed Signal Proces 43:138–147, 2018.
13. Zhang, G. B., Jiang, S., Yang, Z. Y., Gong, L., Ma, X. D., Zhou, Z. Y., Bao, C., and Liu, Q., Automatic nodule detection for lung cancer in CT images: A review. Comput. Biol. Med. 103:287–300, 2018.
14. El-Regaily, S. A., Salem, M. A., Aziz, M. H. A., and Roushdy, M. I., Survey of Computer Aided Detection Systems for Lung Cancer in Computed Tomography. Curr Med Imaging Rev 14(1):3–18, 2018.
15. Rehman, M. Z. U., Javaid, M., Shah, S. I. A., Gilani, S. O., Jamil, M., and Butt, S. I., An appraisal of nodules detection techniques for lung cancer in CT images. Biomed Signal Proces 41:140–151, 2018.
16. Naqi, S. M., and Sharif, M., Recent Developments in Computer Aided Diagnosis for Lung Nodule Detection from CT images: A Review. Curr Med Imaging Rev 13(1):3–19, 2017.
17. Valente, I. R. S., Cortez, P. C., Neto, E. C., Soares, J. M., Albuquerque, V. H. C. D., and Tavares, J. M. R. S., Automatic 3D pulmonary nodule detection in CT images: A survey. Comput. Methods Prog. Biomed. 124(C):91–107, 2016.
18. Lee, S. L. A., Kouzani, A. Z., and Hu, E. J., Automated detection of lung nodules in computed tomography images: a review. Mach. Vis. Appl. 23(1):151–163, 2012.
19. El-Baz, A., Elnakib, A., Abou, E.-G. M., Gimel'Farb, G., Falk, R., and Farag, A., Automatic Detection of 2D and 3D Lung Nodules in Chest Spiral CT Scans. International Journal of Biomedical Imaging 2013(1):517632, 2013.

20. Eadie, L. H., Paul, T., and Gibson, A. P., A systematic review of computer-assisted diagnosis in diagnostic cancer imaging. Eur. J. Radiol. 81(1):e70–e76, 2012.
21. Firmino, M., Morais, A. H., Mendoça, R. M., Dantas, M. R., Hekis, H. R., and Valentim, R., Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects. Biomed. Eng. Online 13(1):41, 2014.
22. Yang, Y. X., Feng, X. Y., Chi, W. H., Li, Z. Y., Duan, W. Z., Liu, H. P., Liang, W. H., Wang, W., Chen, P., He, J. X., and Liu, B., Deep learning aided decision support for pulmonary nodules diagnosing: a review. J Thorac Dis 10:S867–S875, 2018.
23. Hu, Z., Tang, J., Wang, Z., Kai, Z., Lin, Z., and Sun, Q., Deep Learning for Image-based Cancer Detection and Diagnosis — A Survey. Pattern Recogn. 83:134–149, 2018.
24. Paulraj, T., and Chellliah, K. S. V., Computer-Aided Diagnosis of lung cancer in Computed Tomography scans: A Review. Curr Med Imaging Rev 14(3):374–388, 2018.
25. Dean, J. C., and Ilvento, C. C., Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers. Breast Diseases A Year Book Quarterly 187(1):20–28, 2006.
26. Singh, S., Maxwell, J., Baker, J. A., and Nicholas, J. L., Computer -aided classification of breast masses: performance and interobserver variability of expert radiologists versus residents. International Journal of Medical Radiology 258(1):73–80, 2011.
27. Berkman, S., Heang-Ping, C., Roubidoux, M. A., Hadjiiski, L. M., Helvie, M. A., Chintana, P., Janet, B., Nees, A. V., and Caroline, B., Malignant and benign breast masses on 3D US volumetric images: effect of computer-aided diagnosis on radiologist accuracy. Radiology 242(3):716–724, 2007.
28. Giger, M. L., Karssemeijer, N., and Schnabel, J. A., Breast Image Analysis for Risk Assessment, Detection, Diagnosis, and Treatment of Cancer. Annu. Rev. Biomed. Eng. 15(1):327–357, 2013.
29. Reeves, A. P., Biancardi, A. M., Apanasovich, T. V. et al., The Lung Image Database Consortium (LIDC): A comparison of different size metrics for pulmonary nodule measurements. Acad. Radiol. 14(12):1475–1485, 2007.
30. Armato, S. G., Geoffrey, M. L., Mcnitt-Gray, M. F., Meyer, C. R., David, Y., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., and Heber, M. M., Lung image database consortium: developing a resource for the medical imaging research community. Radiology 232(3):739–748, 2004.
31. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., and Pringle, M., The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. J. Digit. Imaging 26(6):1045–1057, 2013.
32. Messay, T., Hardie, R. C., and Tuinstra, T. R., Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the Lung Image Database Consortium and Image Database Resource Initiative dataset. Med. Image Anal. 22(1):48–62, 2015.
33. Setio, A. A. A., Traverso, A., de Bel, T. et al., Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Med. Image Anal. 42:1–13, 2017.
34. Henschke, C. I., Yankelevitz, D. F., Libby, D. M., Pasmantier, M. W., Smith, J. P., and Miettinen, O. S., Survival of patients with stage I lung cancer detected on CT screening. New Engl J Med 355(17):1763–1771, 2006.
35. Rowena, Y., Henschke, C. I., Yankelevitz, D. F., and Smith, J. P., CT screening for lung cancer: alternative definitions of positive test result based on the national lung screening trial and international early lung cancer action program databases. Radiology 273(2):591–596, 2014.
36. Carrilloa, M. C., and Katz, R. G., Maximizing the Alzheimer's Disease Neuroimaging Initiative II. Alzheimers & Dementia the Journal of the Alzheimers Association 5(3):271–275, 2009.
37. Weisheng, W., Jiawei, L., Xuedong, Y., and Hongli, L., Data analysis of the Lung Imaging Database Consortium and Image Database Resource Initiative. Acad. Radiol. 22(4):488–495, 2015.
38. Armato, S. G., McLennan, G., Bidaut, L. 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.
39. Filho, A. O. D. C., Silva, A. C., Paiva, A. C. D., Nunes, R. A., and Gattass, M., Computer-Aided Diagnosis of Lung Nodules in Computed Tomography by Using Phylogenetic Diversity, Genetic Algorithm, and SVM. J. Digit. Imaging 30(6):812–822, 2017.
40. Costa, R. W. D. S., Silva, G. L. F. D., Filho, A. O. D. C., Silva, A. C., Paiva, A. C. D., and Gattass, M., Classification of malignant and benign lung nodules using taxonomic diversity index and phylogenetic distance. Med. Biol. Eng. Comput. 56(11):2125–2136, 2018.
41. Kumar, D., Wong, A., and Clausi, D. A., Lung Nodule Classification Using Deep Features in CT Images. Conference on Computer & Robot Vision:133–138, 2015.
42. Wu, P., Xia, K., and Yu, H., Relevance Vector Machine Based Pulmonary Nodule Classification. J Med Imag Health In 6(1): 163–169, 2016.
43. Akram, S., Javed, Y., Akram, U., Qamar, U., and Hassan, A., Pulmonary Nodules Detection and Classification Using Hybrid Features from Computerized Tomographic Images. J Med Imag Health In 6(1):252–259, 2016.
44. Akram, S., Javed, M. Y., Hussain, A., Riaz, F., and Akram, M. U., Intensity-based statistical features for classification of lungs CT scan nodules using artificial intelligence techniques. J Exp Theor Artif In 27(6):737–751, 2015.
45. Filho, A. O. D. C., Silva, A. C. E., Paiva, A. C. D., Nunes, R. A., and Gattass, M., Classification of patterns of benignity and malignancy based on CT using topology-based phylogenetic diversity index and convolutional neural network. Pattern Recogn. 81:200–212, 2018.
46. Filho, A. O. D. C., Silva, A. C., Paiva, A. C. D., Nunes, R. A., and Gattass, M., Computer-aided diagnosis system for lung nodules based on computed tomography using shape analysis, a genetic algorithm, and SVM. Med. Biol. Eng. Comput. 55(8):1129–1146, 2017.
47. Sweetlin, J. D., Nehemiah, H. K., and Kannan, A., Computer aided diagnosis of pulmonary hamartoma from CT scan images using ant colony optimization based feature selection. Alexandria Engineering Journal 57(3):1557–1567, 2017.
48. Firmino, M., Angelo, G., Morais, H., Dantas, M. R., and Valentim, R., Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. Biomed. Eng. Online 15(1):2, 2016.
49. Orozco, H. M., Villegas, O. O. V., Sánchez, V. G. C., Domínguez, H. D. J. O., and Alfaro, M. D. J. N., Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. Biomed. Eng. Online 14(1):9, 2015.
50. Li, X. X., Li, B., Tian, L. F., and Zhang, L., Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm. IET Image Process. 12(7): 1253–1264, 2018.
51. Gong, J., Liu, J. Y., Sun, X. W., Zheng, B., and Nie, S. D., Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules. Phys. Med. Biol. 63(3):035036, 2018.
52. Dhara, A. K., Mukhopadhyay, S., Dutta, A., Garg, M., and Khandelwal, N., A Combination of Shape and Texture Features

for Classification of Pulmonary Nodules in Lung CT Images. J. Digit. Imaging 29(4):466–475, 2016.
53. Xie, Y. T., Zhang, J. P., Xia, Y., Fulham, M., and Zhang, Y. N., Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. Inform Fusion 42:102–110, 2018.
54. Cataldo, S. D., Bottino, A., Islam, I. U., Vieira, T. F., and Ficarra, E., Subclass Discriminant Analysis of morphological and textural features for HEp-2 staining pattern classification. Pattern Recogn. 47(7):2389–2399, 2014.
55. Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., Zang, Y., and Tian, J., Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification. Pattern Recogn. 61(61):663–673, 2017.
56. Tajbakhsh, N., and Suzuki, K., Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification MTANNs vs. CNNs. Pattern Recogn. 63:476–486, 2017.
57. Silva, G. L. F. D., Neto, O. P. D. S., Silva, A. C., Paiva, A. C. D., and Gattass, M., Lung nodules diagnosis based on evolutionary convolutional neural network. Multimed. Tools Appl. (2):1–17, 2017.
58. Sun, W., Zheng, B., and Wei, Q., Computer aided lung cancer diagnosis with deep learning algorithms. Medical Imaging: Computer-Aided Diagnosis, 2015.
59. Cheng, J. Z., Ni, D., Chou, Y. H., Qin, J., Tiu, C. M., Chang, Y. C., Huang, C. S., Shen, D., and Chen, C. M., Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. Sci Rep-UK 6:24454, 2016.
60. Hua, K. L., Hsu, C. H., Hidayati, S. C., Cheng, W. H., and Chen, Y. J., Computer-aided classification of lung nodules on computed tomography images via deep learning technique. Oncotargets Ther 8:2015, 2015-2022.
61. Tu, X., Xie, M., Gao, J., Ma, Z., Chen, D., Wang, Q., Finlayson, S. G., Ou, Y., and Cheng, J. Z., Automatic Categorization and Scoring of Solid, Part-Solid and Non-Solid Pulmonary Nodules in CT Images with. Convolutional Neural Network. Sci Rep-UK 7(1):8533, 2017.
62. Sun, W., Zheng, B., and Qian, W., Automatic Feature Learning Using Multichannel ROI Based on Deep Structured Algorithms for Computerized Lung Cancer Diagnosis. Comput. Biol. Med. 89:530–539, 2017.
63. Chen, M., Shi, X. B., Zhang, Y., Wu, D., and Mohsen, G., Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network. IEEE Transactions on Big Data PP(99):1, 2017.
64. Yuan, J. J., Liu, X. L., Hou, F., Qin, H., and Hao, A. M., Hybrid-feature-guided lung nodule type classification on CT images. Comput. Graph. 70:288–299, 2017.
65. Ciompi, F., Chung, K., Riel, S. J. V., Setio, A. A. A., Gerke, P. K., Jacobs, C., Scholten, E. T., Schaeferprokop, C., Wille, M. M. W., and Marchianò, A., Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci Rep-UK 7:46479, 2016.
66. Zhu, W., DeepLung: 3D Deep Convolutional Nets for Automated Pulmonary Nodule Detection and Classification, 2017.
67. Zhu, W., Liu, C., Wei, F., and Xie, X., DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification. In: IEEE Winter Conf. on Applications of Computer Vision (WACV2018). pp 673–681, 2018.
68. Liao, F., Ming, L., Zhe, L., Hu, X., and Song, S., Evaluate the Malignancy of Pulmonary Nodules Using the 3D Deep Leaky Noisy-or Network, 2017.
69. Wang, H. F., Zhao, T. T., Li, L. C., Pan, H. X., Liu, W. Q., Gao, H. Q., Han, F. F., Wang, Y. H., Qi, Y. F., and Liang, Z. R., A hybrid CNN feature model for pulmonary nodule malignancy risk differentiation. J X-Ray Sci Technol 26(2):171–187, 2018.
70. Zhao, X., Liu, L., Qi, S., Teng, Y., Li, J., and Wei, Q., Agile convolutional neural network for pulmonary nodule classification using CT images. Int J Comput Ass Rad 13(4):585–595, 2018.
71. Kaya, A., Cascaded Classifiers and Stacking Methods for Classification of Pulmonary Nodule Characteristics. Comput. Methods Prog. Biomed. 166:77–89, 2018.
72. Liu, Y., Hao, P., Zhang, P., Xu, X., Wu, J., and Chen, W., Dense Convolutional Binary-Tree Networks for Lung Nodule Classification. IEEE Access 6:49080–49088, 2018.
73. Wang, Z., Xin, J., Sun, P., Lin, Z., Yao, Y., and Gao, X., Improved Lung Nodule Diagnosis Accuracy Using Lung CT Images With Uncertain Class. Comput. Methods Prog. Biomed. 162:197–209, 2018.
74. Nibali, A., He, Z., and Wollersheim, D., Pulmonary nodule classification with deep residual networks. Int J Comput Ass Rad 12(10):1799–1808, 2017.
75. Liu, K., and Kang, G., 3D multi-view convolutional neural networks for lung nodule classification. PLoS One 12(1):12–22, 2017.
76. Zhang, G., 3D Spatial Pyramid Dilated Network for Pulmonary Nodule Classification. Symmetry-Basel. 10(9), 2018.
77. Yan, Q., Lei, G., Xin, Z., Zhang, X., and Tang, X., Pulmonary nodule diagnosis using dual-modal supervised autoencoder based on extreme learning machine. Expert. Syst. 34(6):e12224, 2017.
78. Jung, H., Kim, B., Lee, I., Lee, J., and Kang, J., Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method. BMC Med. Imaging 18, 2018.
79. Nishio, M., Sugiyama, O., Yakami, M., Ueno, S., Kubo, T., Kuroda, T., and Togashi, K., Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. PLoS One 13(7):e0200721, 2018.
80. Liu, S., Xie, Y., Jirapatnakul, A., and Reeves, A. P., Pulmonary nodule classification in lung cancer screening with three-dimensional convolutional neural networks. Journal of Medical Imaging 4(4):041308, 2017.
81. Wei, G. H., Ma, H., Qian, W., Han, F. F., Jiang, H. Y., Qi, S. L., and Qiu, M., Lung nodule classification using local kernel regression models with out-of-sample extension. Biomed Signal Proces 40: 1–9, 2018.
82. Farag, A. A., Ali, A., Elshazly, S., and Farag, A. A., Feature fusion for lung nodule classification. Int J Comput Ass Rad 12(10):1809–1818, 2017.
83. Dilger, S. K., Judisch, A., Uthoff, J., Hammond, E., Newell, J. D., and Sieren, J. C., Improved pulmonary nodule classification utilizing lung parenchyma texture features. Medical Imaging: Computer-Aided Diagnosis, 2015.
84. Kaya, A., and Can, A. B., A weighted rule based method for predicting malignancy of pulmonary nodules by nodule characteristics. J. Biomed. Inform. 56(C):69–79, 2015.
85. Chen, C. H., Chang, C. K., Tu, C. Y., Liao, W. C., Wu, B. R., Chou, K. T., Chiou, Y. R., Yang, S. N., Zhang, G., and Huang, T. C., Radiomic features analysis in computed tomography images of lung nodule classification. PLoS One 13(2):e0192002, 2018.
86. Reeves, A. P., Xie, Y., and Jirapatnakul, A., Automated pulmonary nodule CT image characterization in lung cancer screening. Int J Comput Ass Rad 11(1):1–16, 2015.
87. Sun, W., Xia, H., Tseng, T. L., Zhang, J., and Wei, Q., Computerized lung cancer malignancy level analysis using 3D texture features. Medical Imaging: Computer-aided Diagnosis, 2016.

88. Ferreira, J. R., and Oliveira, M. C., Characterization of Pulmonary Nodules Based on Features of Margin Sharpness and Texture. J. Digit. Imaging:1–13, 2017.
89. Nishio, M., Nishizawa, M., Sugiyama, O., Kojima, R., Yakami, M., Kuroda, T., and Togashi, K., Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization. PLoS One 13(4):e0195875, 2017.
90. Hancock, M. C., and Magnan, J. F., Lung nodule malignancy classification using only radiologist-quantified image features as inputs to statistical learning algorithms: probing the Lung Image Database Consortium dataset with two statistical learning methods. J Med Imaging 3(4):044504, 2016.
91. Mao, K. M., and Deng, Z. F., Lung Nodule Image Classification Based on Local Difference Pattern and Combined Classifier. Comput Math Method M, 2016.
92. Hawkins, S. H., Korecki, J. N., Balagurunathan, Y., Gu, Y., Kumar, V., Basu, S., Hall, L. O., Goldgof, D. B., Gatenby, R. A., and Gillies, R. J., Predicting Outcomes of Nonsmall Cell Lung Cancer Using CT Image Features. IEEE Access 2:1418–1426, 2014.
93. Li, Q., and Doi, K., Reduction of bias and variance for evaluation of computer-aided diagnostic schemes. Med. Phys. 33(4):868–875, 2006.
94. He, K. M., Zhang, X. Y., Ren S. Q., Sun, J., Deep residual learning for image recognition. In CVPR: 770–778, 2016.
95. Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q., Densely connected convolutional networks. In CVPR 2261–2269, 2017.
96. Krizhevsky, A., Sutskever, I., and Hinton, G. E., ImageNet Classification with Deep Convolutional Neural Networks. International Conference on Neural Information Processing Systems, 2012.
97. Padma, A., and Giridharan, N., Performance comparison of texture feature analysis methods using PNN classifier for segmentation and classification of brain CT images: Performance Comparison of Texture Feature Analysis Methods Using PNN Classifier for Segmentation and Classification of B. Int. J. Imaging Syst. Technol. 26(2):97–105, 2016.
98. Amato, F., Mazzocca, N., Moscato, F., and Vivenzio, E., Multilayer Perceptron: An Intelligent Model for Classification and Intrusion Detection. International Conference on Advanced Information Networking & Applications Workshops, 2017.
99. Sindhumol, S., Kumar, A., and Balakrishnan, K., Spectral clustering independent component analysis for tissue classification from brain MRI. Biomed Signal Proces 8(6):667–674, 2013.
100. Feiping, N., Zinan, Z., Tsang, I. W., Dong, X., and Changshui, Z., Spectral embedded clustering: a framework for in-sample and out-of-sample spectral clustering. IEEE Trans. Neural Netw. 22(11):1796–1808, 2011.
101. de Carvalho, A. O., de Sampaio, W. B., Silva, A. C., de Paivaa, A. C., Nunes, R. A., and Gattass, M., Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index. Artif. Intell. Med. 60(3):165–177, 2014.
102. Ohtsu, N., A Threshold Selection Method from Gray-Level Histograms. IEEE T Syst Man Cy-S 9(1):62–66, 2007.
103. Eberhart R, Kennedy J, A new optimizer using particle swarm theory. In: Proceedings of the sixth international symposium on micro machine and human science, vol. 1, pp. 39–43. New York, 1995. https://doi.org/10.1109/MHS.1995.494215
104. Lecun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D., Backpropagation Applied to Handwritten Zip Code Recognition. Neural Comput. 1(4):541–551, 2014.
105. Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P., Gradient-based learning applied to document recognition. P IEEE 86(11):2278–2324, 1998.
106. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., and Montreal, U., Greedy layer-wise training of deep networks. Adv. Neural Inf. Proces. Syst. 19:153–160, 2007.
107. Coates A, Lee H, and Ng, A. Y., An analysis of single-layer networks in unsupervised feature learning. In Aistats, 2011.
108. Ronneberger, O., Fischer, P., and Brox, T., U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention 9351:234–241, 2015.
109. Heckerman, D., A tractable inference algorithm for diagnosing multiple diseases. Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence. North-Holland Publishing Co:163–172, 1990.
110. Dalal, N., and Triggs, B., Histograms of oriented gradients for human detection 886–893, 2005.
111. Ojala, T., Pietikainen, M., and Harwood, D., A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. 29(1):51–59, 1996.
112. Krizhevsky, A., Sutskever, I., and Hinton, G. E., ImageNet classification with deep convolutional neural networks. In: NIPS: 1097–1105, 2012.
113. Bengio, Y., Louradour, J., Collobert, R., and Weston, J., Curriculum learning. Procintconfon Machine Learning 60(60):6, 2009.
114. Mcnittgray, M. F., Armato, S. G., Iii, M. C. R., Reeves, A. P., Mclennan, G., Pais, R., Freymann, J., Brown, M. S., Engelmann, R. M., and Bland, P. H., The Lung Image Database Consortium (LIDC) Data Collection Process for Nodule Detection and Annotation. Acad. Radiol. 14(12):1464–1474, 2007.
115. Dhara, A. K., Mukhopadhyay, S., Gupta, R. D., Garg, M., and Khandelwal, N., Erratum to: A Segmentation Framework of Pulmonary Nodules in Lung CT Images. J. Digit. Imaging 29(1):86–103, 2016.
116. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P. A., Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. J. Mach. Learn. Res. 11(12):3371–3408, 2010.
117. Jing-Jing, W., Hai-Feng, W., Tao, S., Xia, L., Wei, W., Li-Xin, T., Da, H., Ping-Xin, L., Wen, H., and Xiu-Hua, G., Prediction models for solitary pulmonary nodules based on curvelet textural features and clinical parameters. Asian Pacific Journal of Cancer Prevention Apjcp 14(10):6019–6023, 2013.
118. Haifeng, W., Tao, S., Jingjing, W., Xia, L., Wei, W., Da, H., Pingxin, L., Wen, H., Keyang, W., and Xiuhua, G., Combination of radiological and gray level co-occurrence matrix textural features used to distinguish solitary pulmonary nodules by computed tomography. J. Digit. Imaging 26(4):797–802, 2013.
119. Farag, A. A., Farag, A. A., Falk, R., Ali, A. M., Graham, J., Elshazly, S., Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest, 2011.
120. Farag, A., Elhabian, S., Graham, J., Farag, A., and Falk, R., Toward Precise Pulmonary Nodule Descriptors for Nodule Type Classification. International Conference on Medical Image Computing & Computer-assisted Intervention, 2010.
121. Lin, P. L., Huang, P. W., Lee, C. H., and Wu, M. T., Automatic classification for solitary pulmonary nodule in CT image by fractal analysis based on fractional Brownian motion model. Pattern Recogn. 46(12):3279–3287, 2013.