Review Article

A survey of pulmonary nodule detection, segmentation and classification in computed tomography with deep learning techniques

Jianrong Wu, Tianyi Qian

Technology Committee of Tencent Healthcare, Shenzhen 518052, China

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study material or patients: None; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jianrong Wu. Technology Committee of Tencent Healthcare, Shenzhen 518052, China. Email: wjr0529@gmail.com.

Abstract: Lung cancer is the top cause for deaths by cancers whose 5-year survival rate is less than 20%. To improve the survival rate of patients with lung cancers, the early detection and early diagnosis is significant. Furthermore, early detection of pulmonary nodules is essential for the detection and diagnosis of lung cancer in early stage. The National Lung Screening Trial (NLST) showed annual screening by low-dose computed tomography (LDCT) could help to reduce the deaths caused by lung cancer of high-risk subjects by 20% comparing with screening by chest radiography. In past decade, there has been lots of works on computer-aided detection (CADe) and computer-aided diagnosis (CADx) for pulmonary nodules in computed tomography (CT) scans, whose target is to detect, segment the nodules and further classify them into benign and malignant efficiently and precisely. This survey reviews some recent works on detection, segmentation and classification for pulmonary nodule in CT scans with deep learning techniques.

Keywords: Pulmonary nodule detection; pulmonary nodule segmentation; pulmonary nodule classification; deep learning; convolutional neural networks (CNNs)

Received: 27 December 2018; Accepted: 29 March 2019; Published: 19 April 2019.
doi: 10.21037/jmai.2019.04.01
View this article at: http://dx.doi.org/10.21037/jmai.2019.04.01

Introduction

Lung carcinoma, well known as lung cancer, is the top cause of cancer related deaths all over the world. Research showed the incidence rates of women and men as 19.5% and 13% respectively (1,2). Since there is no symptom in its early stages, around 70% patients are diagnosed as lung cancer at advanced stage. In China, the 5-year survival rate of a patient at advanced stage is about 16% (3). However, the lung cancer could be cured effectively with great chance if it is diagnosed at an early stage rather than at an advanced stage, and the 5-year survival rate could rise to 70% (4). The early diagnosis of lung cancer, mainly rely on the detection of the lung nodule, also known as pulmonary nodule. To get a better method to detect nodules in early stage, the National Lung Screening Trial (NLST) research team did research on a large population of patients from different centers, whose results showed that high-risk smokers screened with low-dose computed tomography (LDCT) got less death by 20%, compared to those with thorax radiography (5). Furthermore, as recommended by Naidich *et al.* (6) and the American College of Radiology (ACR) (7), thin-slice computed tomography (CT) scans should be the first choice for the management of pulmonary nodules.

Therefore, CT scans as the most suitable modality to find the pulmonary nodules, are use more and more broadly by the radiologists nowadays. However, screening CT images manually is a very time-consuming job for radiologists, for there are hundreds of slices in one scan and there are less than 100 voxels in a single nodule usually. Advanced computer-aided detection systems (CADe) and computer-aided diagnosis systems (CADx) are likely to push such process on. However, this task is so complicated that, there are huge variation in shape, anatomical context, intensity and size between nodules. Besides, there are many non-nodules tissues similar to the appearance of nodules, such chest wall and vessels (8). Nodules can be categorized into three different kinds according to spatial locations, including juxta-pleural nodules, juxta-vascular nodules and solitary pulmonary nodules, On the other hand, based on their spatial patterns, they can be categorized into solid pulmonary nodules, part-solid pulmonary nodules and ground-glass pulmonary nodules. The CAD, including CADe and CADx, can help radiologists to find potential abnormalities. The accuracy of pulmonary nodule diagnosis will be improved by improving the radiologist's efficiency.

Considerable research efforts have been devoted to pulmonary nodule detection on chest CT images, so as to segmentation and classification. It could be divided into non-learning based methods (9-12) and learning based methods (13). The non-learning based methods are usually experience driven and compute morphological features of the images that could well distinguish the target regions and background, i.e., nodule and vessels. The learning based methods are data driven and learn the optimal decision boundary in a high-dimensional feature space from a set of prepared labeled data to distinguish target and background. There are two obvious differences between these two methods: (I) need of data with annotation or not; (II) who define the distinguishing boundary or features. The non-learning based methods do not need annotated data, and the distinguishing features or boundary are totally manually defined. However, annotated data is necessary for learning based methods, which is regarded as training data. A typical learning based method composes of feature extraction and classifier, while the classifier need to be learned from the training data, and the features could be handcrafted or learned from the training data. The handcrafted features are defined manually, such as HOG (14), LBP (15), SIFT (16), that showed great advantage before deep learning. Deep learning techniques compose of many layers with different filers that learned from training data and used to transform the input to high-dimensional feature. Convolutional neural networks (CNNs) are the most successful type for image analysis.

The rest of paper is organized as follows. Available public datasets for pulmonary nodule related applications are shown in section 2. Section 3 presents a brief overview introduction of deep learning techniques. Section 4 presents the three main applications of pulmonary nodule, including detection, segmentation and classification.

Open dataset of pulmonary nodule

Publicly medical image databases for development and evaluation have been available for about two decades (17,18) which many methods were developed or evaluated based on (19-22).

For case of lung cancer screening, the first trial was the Early Lung Cancer Action Program (ELCAP) that made the ELCAP Public Lung Image Database available in 2003. The database consists of 50 LDCT scans for the evaluation of CAD systems.

The NLST randomized 26,724 subjects to the CT screening arm of its two-arm study. A total of 48,547 scans were selected from among the 75,133 LDCT scans acquired from 33 participating institutions were archived in the CT Image Library (CTIL). As the most commonly used and largest publicly available dataset for early diagnosis of lung cancer, the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) involved 1,010 patient records gathered from University of Iowa, University of Chicago, University of California at Los Angeles, Weill Cornell Medical College and University of Michigan. LUNA16 dataset is a subset of LIDC-IDRI with detection annotation only.

In this section, LIDC-IDRI and LUNA16 are presented in details.

LIDC-IDRI

After the initialization by National Cancer Institute (NCI), the LIDC-IDRI was further constructed by Foundation for the National Institutes of Health (FNIH) and Food and Drug Administration (FDA). There are 1,018 CT scans of 1,010 patients enrolled in this dataset, each of which includes a thoracic CT scan and a corresponding XML file that recorded the annotated results finished by four professional radiologists. The annotation process composed of two phases, which aimed to recall as entirely as possible all nodules in each CT scan. In the first phase, named as blinded-read phase, each radiologist reviewed scans independently and marked lesions, such as "nodule <3 mm", "nodule ≥3 mm", and "non-nodule ≥3 mm". In the subsequent phase, named as unblinded-read phase, each radiologist checked his own marks independently and made a final decision refer to anonymous marks of other radiologists. There are 7,371 lesions labeled as "nodule" by at least one radiologist inside the dataset; 2,669 of these lesions got at least one "nodule ≥3 mm" label from four radiologist, while 928 got four. Subjective nodule characteristic ratings and nodule outlines were involved in these 2,669 lesions.

The outline could describe the shape and size of nodule by localizing the outer border, which did not overlap pixels belonging to the nodule. If a nodule was marked as "nodule ≥3 mm" after two phases, its characteristics were assessed by radiologist, including internal structure, shape (sphericity), margin, solidity, lobulation, speculation, subtlety and likelihood of malignancy.

LUNA16

The LUNA16 dataset was created for the challenge of LUng Nodule Analysis 2016 including 888 CT scans, which were gathered from LIDC-IDRI with slice thickness less than 3mm. There are totally 36,378 annotations of nodules that were marked by more than one radiologists, while there are 2,290, 1,602, 1,186, and 777 nodules annotated by at least 1, 2, 3, or 4 radiologists, respectively. Nodules annotated by at least 3 radiologists are regarded as true nodules, whose annotations of diameters and positions are the average of annotation in LIDC-IDRI.

Overview of deep learning techniques

In this section, the deep learning technique is introduced, especially deep CNN. Generally, deep learning provides the potential to automate and merge the extraction of relevant features with the classification procedure (23-25). Recently, CNN has been shown to be very powerful in computer vision application and keep breaking the performance records in challenges (26,27). Specially, applications of medical image analysis have already leveraged CNN (28-30).

CNNs compose of many layers that transform their input with weights by convolutional operation. Research on CNNs has been started since the late seventies (31). The first successful real-world application of CNN is the LeNet (32) for recognition of hand-written digit. Krizhevsky *et al.* (26) proposed a CNN named AlexNet, and won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 by a large margin. In the following years, there was large progress using related and innovated deeper architectures (27,33). Nowadays in computer vision, CNNs has become the first choice, even in many clinical applications.

Deep learning in pulmonary nodule application

Generally, there are three main applications of pulmonary nodule: detection, segmentation and classification. The detection module aims to predict the exact position of nodule inside the lung. The segmentation module aims to predict which voxels are nodule voxels. And the classification module could predict the exact type of nodule, i.e., benign or malignant.

In this section, researches with deep learning on these three applications are presents, while some non-deep learning methods are presented for comparison.

Detection

Generally, nodule detection contains two main stages: (I) nodule candidate generation; (II) false positive reduction (FPR). Nodule candidate generation is a basis stage, while FPR is an advanced stage to make the detection more precise. In this part, detection of pulmonary nodule is presented in two parts differing from using deep learning or not.

Detection with Deep learning

The deep learning is furthermore to improve the ability to build the classification boundary from data, especially for big data. Ozdemir *et al.* (34) proposed a two-stage Bayesian CNN architecture, to leverage the segmentation predictions along with their uncertainties. First, segmentation networks operate on 2D axial CT slices. Then segmentation predictive mean and standard deviation maps are fused with the original image to form a 3-channel composite image, which is fed into a 3D Bayesian CNN for final nodule detection. Zhu *et al.* (19) integrated expectation-maximization (EM) to construct a new deep 3D CNN framework, named as DeepEM, whose aim was mining the weakly supervised labels in EMRs for pulmonary nodule detection. The nodule proposal generation was finished by Faster RCNN. Logistic regression and Half-Gaussian model are utilized for lobe location central slice respectively. In the E-step, all the slices, observations, and weak labels were employed to predict the nodule proposals, latent variable, by sampling or maximum a posteriori (MAP). In the M-step, parameters in logistic regression and Faster R-CNN were updated by using the estimated proposals. Xie (20) proposed a 3D U-Net leveraging residual and dense learning from ResNet and DenseNet (35) for pulmonary nodule detection. This method extended the region proposal network (RPN) in faster RCNN to 3D scheme with patch-based scanning. Dou *et al.* (36) integrated a set of 3D CNNs with different sizes of receptive field to involve multi-level contextual information around pulmonary nodules. Three different 3D CNNs are constructed for cropped cubes with different sizes, including 20×20×6, 30×30×10, 40×40×26. Three results are fused at last. Jiang *et al.* (37) propose a method for pulmonary nodule detection utilizing multi-group patches cut out from the lung images. First, Frangi-filters were used to eliminate the vessel-like structures, and a slope analysis method was designed to eliminate the nodule outside lung by repairing the juxta-pleural nodules. Finally, a CNN structure which utilized multi-crop (MC) pooling operation was designed to learn the knowledge of radiologists using original CT scans and its binarization.

FPR which could be regarded as classification of distinguishing nodule from non-nodule is a single topic that receives attention from researchers. Instead of using conventional convolutions, Winkels (38) proposed a FPR method from pulmonary nodule detection by using 3D roto-translation group convolutions (G-Convs). Its results showed that, it was more effective in sensitivity of malignant nodules, convergence speed than baseline architecture with similar number of parameters and regular convolutions. Golan *et al.* (39) proposed a sliding window based method to detect the nodules. A CNN for classification was trained with sub-volume of size 5×20×20 first. Then, it is applied to the whole CT scans using sliding window of shape as 5×20×20. The outputs of CNN in different sliding windows positions were averaged to compute the 3D voting grid, so as to predict the nodule's location. Anirudh *et al.* (40) also proposed a sliding window based method but employed 3D CNN to learn more discriminative features for nodule detection. The 3D CNN was trained to infer the categories of a single voxel, that inside a nodule or not. 3D Hessian was used as FPR method for dot enhancement to get the most like nodule candidates. Khosravan *et al.* (41) modeled lung nodule detection as a cell-wise classification problem, done simultaneously for all the cells, and used a single feed forward pass of a single network for detection, namely S4ND. The pipeline is constructed as a 3D dense CNN with end-to-end training. The input scan is divided by a 16×16×8 grid and is passed through the S4ND. The output is a probability map indicating the presence of a nodule in each cell.

Furthermore, some researches combined nodule generating and FPR in one pipeline.

Xie *et al.* (42) proposed a 2D nodule detection and FPR method. Firstly, the Faster RCNN was used to detect nodule candidates, with two RPN and a deconvolutional layer for the middle three slices of the nodule respectively, while extra neighboring two slices were used as input for each. Secondly for FPR, a boosting architecture by 2D CNN is proposed. To further boost the sensitivity of detection, the model is retrained with the misclassified samples. At last, the final classification results are made by fusion of results of those networks. Ding *et al.* (21) proposed a nodule detection method composing of 2D candidate generation and 3D FPR. The method detects candidates on axial slices by Faster RCNN with a deconvolutional structure, and the input is not only a single slice but with neighboring slices. The 3 slices in axial direction were rescaled to 600×600×3 for input. Then in FPR, a 3D DCNN is used. Compared with 2D CNN, the 3D CNN can capture the candidate's full range of contexts and extract more discriminative features. Dou *et al.* (22) proposed a nodule detection method including 3D CNN for candidate generation and FPR. To involve the size and location information, a hybrid-loss residual network was employed. In the candidate generation step, the 3D CNN was trained for classification of nodule and non-nodule with small 3D patches. To increase the ratio of hard samples, an online sample filtering method was used. To leverage both priori knowledge about lung nodules and machine learning, Huang *et al.* (43) used a local geometric-model-based filter to generate nodule candidates and further reduced the structure variability by estimating the local orientation. At last, a deep 3D CNN trained to classify nodule and non-nodule was used to classify the 3D cubes of nodule candidates. Zhang *et al.* (44) proposed a 3D progressive resolution-based densely dilated FCN, namely the progressive resolution network (PRN), to detect nodule candidates, and construct a densely dilated 3D CNN with hierarchical saliency, namely the hierarchical saliency network (HSN), for FPR. The PRN was constructed by a dual-path encoder and a decoder. In the encoder path, each input volumetric patch was enlarged to 64×64×64, then, feed the patch and its enlarged copy to two encoder paths, respectively. Each encoder path composes of convolutional layers, ReLU activation function and PRN blocks. The proposed HSN model consists of four subsequent HSN blocks. Each HSN block takes three volumetric patches as input, which are cropped on the chest CT scan according to the center of the detected nodule candidate. Tang *et al.* (8) proposed a nodule detection and FPR method. A 3D Faster RCNN inspired by Unet was proposed for nodule detection, which was trained with online hard negative mining. To extract detailed local information of the nodule, shortcuts from the end of each block to the last feature map were integrated to the 3D CNN classifier, which was used in FPR. The FPR classifier was trained on difficult examples produced during candidate generation. Sun *et al.* (45) used graph cuts segmentation to identify and segment the nodule candidates. And then, a CNN was utilized for FPR. Setio *et al.* (46) proposed a multi-view convolutional network, which comprises multiple streams of 2D CNN. Candidates by three detectors specifically designed for large nodules, solid and part-solid nodules were combined for input to the network. Each candidate is extracted 2D patches using nine views of a volumetric object, and go through corresponding stream. For the final classification, all outputs are combines using fusion method. Sakamoto *et al.* (47) proposed fusion classifier with the cascaded CNNs to classify nodule from non-nodules. Firstly, Nodule probabilities are calculated by CNN. Then the nodule probabilities were used to train the fusion classifier, which was also a deep CNN. The method operated as single-sided classifiers, filtering processes for obvious non-nodules. The input is a vector of the nodule probability of a nodule candidate generated from several CNN classifiers.

There is research focusing not only deep learning, but also leverage handcrafted features. Chen *et al.* (48) proposed three multi-task learning (MTL) schemes for the description of 9 semantic features of pulmonary nodule, which leveraged handcrafted HoG and Haar-like features, as well as features derived from deep learning models of CNN and stacked denoising autoencoder (SDAE).

Detection without deep learning

There are variable of research on pulmonary nodule detection before the rising trend of deep learning. Those non-deep learning methodologies mainly use morphological filers with thresholding method and traditional machine learning method compositing of handcrafted features and classifiers.

By leveraging the power of morphological filters, Teramoto *et al.* (9) utilized active contour filter to enhance lung regions, after that the nodules were extracted by thresholding. John and Mini (11) used intermediate thresholding to locate nodules. Then, the binary image mask and micro-level thresholding were combined to obtain the initial nodule candidates. Liu *et al.* (12) proposed a model matching method to detect nodules, namely selective enhancement filter. Hidden conditional random field (HCRF) was used to detect nodules for 3D representation stacked by 2D CT scans.

Compared with the morphological filters, machine learning based method could build the optimal boundary from data itself. A supervised extraction of the region of interest was proposed by Orozco *et al.* (49) to eliminate the differences of shape between CT images. Features from each wavelet sub-band are combined in pairs as input to a SVM, which is used to classify CT images with cancerous nodules or without. Santos *et al.* (50) segmented the structures in lung by Gaussian mixture models firstly. And then, discriminate nodule from non-nodule by texture descriptor as Shannon's and Tsallis's Q entropy. Besides, the Hessian matrix was used to separate bronchi and blood vessels, and reduced false positives by SVM. Lu *et al.* (51) proposed a hybrid method for nodule detection, integrating local density maximum algorithm, fuzzy connectedness (FC) segmentation, geodesic distance map, dot-enhancement based on Hessian matrix, regression tree classification, and orphological operation. Farahani *et al.* (52) proposed a nodule detection method with ensemble of three classifiers including SVM, KNN and MLP. Specific features like compactness, circularity, roundness, eccentricity and ellipticity are calculated from the segmented 2D images and used for classification for each classifier. Finally, majority voting method is used to combine decisions for the diagnosis of nodule. Klik *et al.* (53) used the linear discriminate analysis (LDA) classifier to distinguish nodules from candidates generated by segmentation using gray level characteristics and optimal thresholding. The LDA classifier was regarded as a FPR classifier using geometrical and gray level characteristics. Forz *et al.* (54) integrated SVM and texture features together and proposed a methodology to classify lung nodule and non-nodule (FPR). To extract the texture features, three techniques were involved: artificial crawlers (ACs), rose diagram (RD) and a hybrid model combining the RD and texture measurements from ACs. Finally, the SVM was employed with a radial basis kernel. Comparison of different detection methods of nodules is shown in *Table 1*. Because of different sorting methods for LIDC-IDRI dataset, it is hard to compare different methods evaluated on LIDC-IDRI. However, the 3D input

Table 1 Illustration of results by referenced detection methods

Author	Dim	Trainset	Valset	Testset	Recall (%)	FP	Average recall (FROC)	F1-score
Ozdemir *et al.* [2017]*	2.5D	LUNA16 (80%)	10%	10%	92			
Zhu *et al.* [2018]*	3D	LUNA16 (10-fold cross validation), NCI NLST		Tianchi			0.764	
Xie [2017]*	3D	LUNA16 (5-fold cross validation)					0.9226 (best fold)	
Dou *et al.* [2016]*	3D	ISBI 2016's LUNA FPR track (10-fold cross validation)					0.827	
Jiang *et al.* [2017]*	2D	LIDC-IDRI (1,006 in all, 10-fold)			80.06	4.7		
Winkels *et al.* [2018]*	3D	NLST (30,000 samples)	NLST (8,889 samples)	LIDC (8,582 samples)			0.88 (best fold)	
Golan *et al.* [2016]*	2.5D	LIDC-IDRI (814 scans)	–	204 scans	71.2	10		
Anirudh *et al.* [2016]*	3D	SPIE-AAPM-LUNGx (20 scans)	–	47 scans	80	10		
Khosravan *et al.* [2018]*	3D	LUNA16 (10-fold cross validation)					0.897	
Xie *et al.* [2019]*	2D	LUNA16 (10-fold cross validation)					0.775	
Ding *et al.* [2017]*	2D	LUNA16 (10-fold cross validation)					0.891	
Ding *et al.* [2017]*	3D	LUNA16 (10-fold cross validation)					0.839	
Huang *et al.* [2017]*	3D	LIDC-IDRI (10-fold cross validation)			90	5		
Zhang *et al.* [2018]*	3D	LUNA16 (10-fold cross validation)					0.958	
Sun *et al.* [2017]*	2D	LIDC-IDRI (5-fold, 595 scans, 305 nodules)			87.7			0.8501
Setio *et al.* [2016]*	3D	LUNA16 (5-fold cross validation)					0.9226	
Sakamoto *et al.* [2018]*	2D	LUNA16 (10-fold cross validation)			94.4	4		
Santos *et al.* [2014]	2D	LIDC-IDRI (112 scans)	–	28 scans	90.6	1.17		
Lu *et al.* [2015]	2D	LIDC-IDRI (196 scans)	–	98 scans	85.2			
Kilk *et al.* [2006]	2D	284 nodules (10-fold cross validation)			65			
Froz *et al.* [2017]	3D	LIDC-IDRI (833 scans total, 6415 nodules)			91.86			

* indicates deep learning based methods. FP, false positive per case; Dim, dimension of input data.

data based methods achieved better performance than 2D and 2.5D based methods in general. Also, deep learning based methods outperform non-deep learning methods with larger training dataset and validation dataset. In the case of LUNA16, 3D CNN based method show obvious advantages compared with 2D or 2.5D based methods, while Zhang *et al.* achieved the best average recall of FROC with LUNA16 by a 3D CNN based method. The fact of better results by 3D based methods illustrates that the surrounding and texture of the nodules can be leveraged to facilitate the detection of nodule.

Segmentation

The segmentation of nodules in CT scans allows quantitative analysis of clinical parameters related to shape, volume and distribution of voxel values.

Segmentation with deep learning

For the segmentation with deep learning, some research regarded the segmentation mission as a classification mission voxel by voxel. Wang *et al.* (55) developed a multi-view convolutional neural networks (MV-CNN) for nodule segmentation to distinguish voxel by voxel if it is belonging to nodule or not. The proposed MV-CNN composed of three CNN branches, each of which takes multi-scale nodule patches from sagittal, coronal and axial views as input. A fully connected layer was applied to integrate the three CNN branches, to infer whether the voxel in patch center is nodule. Besides MV-CNN, Wang *et al.* (56) proposed another model, namely the central focused convolutional neural networks (CF-CNN), to segment lung nodules as classification voxels by voxels. Both multi-scale 2D features and 3D features were leveraged in a two-branch CNN. The branch of 2D-patch learns multi-scale features from multiple 2D patches, and the branch of 3D-patch learns multi-view features from multiple CT slices. To facilitate the training of model, a weighted sampling was employed to select training samples refer to the difficulty of segmentation. Trial of combining deep learning and non-deep learning has been made. Mukherjee *et al.* (57) hypothesize that combining both model based strategies and data driven in one hybrid approach could be more suitable for the complicate task of pulmonary nodules segmentation. An energy minimization based segmentation framework deep learned prior based graph cut (DLGC) was proposed, which combines a domain specific cost function using low level image features and a deep learned object localization. The combination of machine learned prior and unsupervised image based component makes it robust to initialization errors with higher flexibility.

Segmentation without deep learning

For segmentation of pulmonary nodules without deep learning, there are several specific methodologies, such as active contour, fuzzy related methods. Nithila and Kumar (58) developed a fuzzy c-means (FCM) and region based active contour model technique for segmentation of lung nodules. Reconstruction of lung parenchyma is operated by Gaussian filtering and selective binary with new signed pressure force function (SBGF-new SPF). And nodule segmentation was finished by clustering technique. Aresta *et al.* (10) used three methods including refined morphological, active contours and region-growing to segment nodules. Badura and Pietka (59) proposed a multilevel method to segment the pulmonary nodules in CT scans. The FC is first used to generate the masks, especially for the nodules connected to the vessels or pleura. Then evolutionary computation is utilized to improve the FC analysis and its accuracy. Gonçalves *et al.* (60) proposed a nodule segmentation approach using Hessian-based strategies. It is a multiscale process, which uses a Hessian-based strategy and the central medialness adaptive principle.

Table 2 shows the comparison of different segmentation methods, whose methods are evaluated on LIDC-IDRI. There are different sorting methods of dataset and evaluation parameters, and we refer to the Dice coefficient as standard here. The 2.5D based methods by Wang *et al.* achieved the best dice, which used a novel method called central pooling in CNN. The pooling process by central pooling can focus on the nodule and make the feature map after pooling more effective. Otherwise, the non-deep learning based methods do not use Dice coefficient for evaluation, it is hard to tell which is better in related methods.

Classification

Generally, to classify a nodule in a CADx system may comprise three components: (I) nodule detection; (II) nodule segmentation; (III) feature extraction from nodule candidates; and (IV) classification of the candidate as benign or malignant based on its extracted features.

The main mission in nodule classification is the distinguishing of benign and malignancy.

Classification with deep learning

Cheng *et al.* (61) proposed a stacked denoising auto-encoder (SDAE) to discriminate benign nodules from malignant nodules. Buty *et al.* (62) proposed to combine features of nodule 3D shape and appearance to predict the malignancy.

Table 2 Illustration of results by referenced segmentation methods

Author	Dim	Trainset	Valset	Testset	Acc (%)	RMSE (mm)	Sens (%)	Prec (%)	ASD (mm)	Dice (%)
Wang *et al.* [2017]*	2.5D	LIDC-IDRI (450 nod)	50 nod	393 nod			83.72	77.58	0.24	77.67
Wang *et al.* [2017]*	2.5D	LIDC-IDRI (450 nod)	50 nod	393 nod			92.75±12.83	75.84±13.14	0.17±0.23	82.15±10.76
	2.5D	GHGH	–	74 nod			83.19±15.22	79.30±12.09	0.35±0.34	80.02±11.09
Mukherjee *et al.* [2017]*	2D	LIDC-IDRI	LIDC	128 nod						0.69±0.14
Nithila *et al.* [2016]	2D	LIDC-IDRI	–	LIDC	98.9	0.1				
Badura *et al.* [2014]	3D	LIDC-IDRI	–	551 nod			95.5±7.86			

Dim, dimension of input data; Acc, accuracy; RMSE, root mean square; Sens, sensitivity; Prec, precision; ASD, average surface distance.

Nodule shape was modeled from radiologist's binary nodule segmentations and parameterized using spherical harmonics (SH). The proposal combined three orthogonal patches of each nodule as one image input for the DCNN, and extracted the appearance features from the first fully-connected layer. A random forest classifier was used to predict the nodule malignancy based on the combining of both sets of features. Hussein *et al.* (63) propose a multi-view CNN for nodule characterization of malignancy. Three 2D patches corresponding to each dimension were concatenated to form a 3D tensor. The features of the input image were extracted by the network and then a Gaussian process (GP) regression was used for prediction of malignancy score. Dey *et al.* (64) proposed four two-pathway CNN to predict the malignancy of nodule, including a 3D DenseNet, a novel multi-output network, a basic 3D CNN and an augmented 3D DenseNet with multi-outputs. Each network has two pathways with 3D inputs of different scales. Based on the evaluation of LIDC-IDRI dataset, the 3D multi-output DenseNet achieves better classification accuracy. Wu *et al.* (65) proposed an interpretable and MTL CNN to predict the malignancy of pulmonary nodules, namely joint learning for pulmonary nodule segmentation attributes and malignancy prediction (PN-SAMP). Besides prediction of malignancy, the areas of detected nodules and the semantic high-level attributes could be provided. The image patches of two different window centers and window widths are stacked together as the input to the CNN. Shen *et al.* (66) presented a 3D interpretable deep hierarchical semantic CNN (HSCNN) to predict the malignancy of a nodule. There are two levels of output: (I) low-level radiologist semantic features; and (II) a high-level malignancy prediction score. The low-level task predicts five semantic diagnostic features: margin, texture, sphericity, subtlety, and calcification. The high-level task incorporates information from both the generalizable image features and the low-level tasks to produce an overall prediction of malignancy. Shen *et al.* (67) built an end-to-end machine-learning framework for malignancy classification of pulmonary nodule, named as multi-crop convolutional neural network (MC-CNN). To automatically extract nodule salient information, a MC pooling strategy cropping different regions from convolutional feature maps was employed and max-pooling was applied different times for each other. After the last fully connected layers, the features were concatenated together for classification. Hussein *et al.* (68) proposed a MTL to perform joint learning of tasks, such as malignancy, sphericity, spiculation, lobulation, texture, margin and calcification. Seven 3D CNN was fine-tuned using labels of those 7 attributes respectively. The features by the first fully connected layer in each model were utilized as feature representation. Graph regularized sparse least square optimization function was used to fuse features from different CNNs and got the coefficient vectors corresponding to each task. In the testing phase, feature representation of the testing image with the coefficient vector was multiplied to obtain the malignancy score.

It is easy to say that, a fully automatic CADx system could not rely on single classification, and detection is necessary beforehand. Zhu *et al.* (69) proposed a method combining detection and classification (benign or malignant) named as DeepLung. Two deep 3D dual path networks (DPN) are designed for nodule detection and classification respectively. For nodule detection, a 3D faster RCNN with a U-net-

Table 3 Illustration of results by classification methods

Author	Dim	Trainset	Valset	Testset	Acc (%)	AUC	Sens (%)	Prec (%)
Cheng *et al.* [2016]*	2D	LIDC-IDRI (1,360 nodules, 10 times 10-fold cross validation)		140 nod	94.4±3.2	98.4±1.5	90.8±5.3	91.6±4.4
Buty *et al.* [2016]*	2.5D	LIDC –IDRI (1,432 nod, 10-fold cross validation)			82.4			
Hussein *et al.* [2017]*	2.5D	LIDC –IDRI (1,145 nod, 10-fold cross validation)			82.47		62	
Dey *et al.* [2018]*	3D	LIDC-IDRI (147 scans, 5-fold cross validation)			90.4	95.48%	90.47	0.9055
Wu *et al.* [2018]*	3D	LIDC-IDRI (1,404 nod, 5-fold cross validation)			97.58±1.32			
Shen *et al.* [2018]*	3D	LIDC-IDRI (897 scans, 4-fold cross validation)			84.2±2.5	0.856±0.026	70.5±4.5	
Shen *et al.* [2017]*	3D	LIDC-IDRI (825 nod, 5-fold cross validation)	275 nod	275 nod	87.14	0.93	77	
Hussein *et al.* [2017]*	3D	LIDC-IDRI (1,340 nod, 10-fold cross validation)			91.26			
Zhu *et al.* [2018]*	3D	LIDC-IDRI (1,004 nod, 5-fold cross validation)			90.44			
Liao *et al.* [2017]*	3D	DSB2017 (754 nod)	–	78 nod	81.42	0.87		

Dim, dimension of input data; Acc, accuracy; Sens, sensitivity; Prec, precision.

like encoder-decoder structure and 3D dual path blocks was designed. For nodule classification, gradient boosting machine (GBM) with 3D DPN features is proposed. Liao *et al.* (70) proposed a method for detection and classification (benign or malignant), which won the Data Science Bowl 2017 (DSB2017). For the nodule detection, a 3D RPN with backbone of modified U-net was proposed. And for the classification, the top five nodules were selected based on the detection confidence, evaluates their cancer probabilities and combines them with a leaky noisy-or gate to obtain the probability of lung cancer for the subject, which share the same backbone as detection

Besides the malignancy, there is effort on types of nodule. Zhang *et al.* (71) proposed a semi-supervised classification method for four types of lung nodules, including pleural tail, vascularized, well-circumscribed and juxta-pleura. There were two steps in the proposal: (I) construction of bipartite graph, which showed the similar relationship between unlabeled and labeled images; (II) calculation of ranking score, which calculated the probability of unlabeled images for each type.

Classification without deep learning

Shen *et al.* (72) proposed a hierarchical learning framework for nodule classification of benign or malignancy, named as MC-CNN. The method extracted features from alternatingly stacked layers, so as to capture the nodule heterogeneity. Klik *et al.* (53) proposed a non-deep learning method to distinguish perifissural opacities (PFOs) from potentially malignant nodules. The proposal leveraged two characteristic properties of PFOs for classification, including its attachment to plate-like structures in the direct neighborhood of the nodule (the lung fissures) and the typical flattened surface.

Classification methods shown in *Table 3* are all deep learning based methods, with training dataset mainly from LIDC-IDRI, except one from DSB2017. Methods with 3D input outperform 2.5D input in LIDC-IDRI dataset, and achieved the best accuracy as high as 97.58%. The accuracy of the 2D input methods by Cheng *et al.* is 94.4%, and outperforms most of other methods, except the 3D input methods by Wu *et al.* It also shows the surrounding and texture of the nodules can be leveraged to facilitate the classification of nodule.

Conclusions

This paper presented a survey of automatic CADe and CADx for detection, segmentation and classification of pulmonary nodule on CT scans. The cited papers are mainly gathered from popular academic journals and conferences, which were published in recent years. As the fast-growing in deep learning techniques and break-through performance, variables of deep learning-based methodologies whose results show great advantage have been proposed for these three missions as well as non-deep learning ones. In most research, deep learning is leveraged to solve or improve a single mission. However, several multi-tasks work combining detection and classification are done. In some deep learning-based works, non-deep learning method is involved to boost its ability, and the results show solid evidence. There are efforts to expedite the progress of research on pulmonary nodules, such as more and more available open dataset and challenges. It is generous to make effort to build the open available dataset, because it is really a complicate hard job to collect the data and annotate it by doctors, and it can really help researcher to push on their research and validate their ideas. And well known challenges as Data Science Bowl 2017 and LUNA 2016 were held which provided a stage for knowledge sharing rather than competition. In some case, proposed methods have reached even exceed radiologists' level. However, there is a significant problem for all learning-based methods that is generalization ability. It is due to the diversity of training dataset and the methods itself. As a conclusion, deep learning has made great impact on the applications of pulmonary nodule, and we believe that it will be a more powerful methodology in the future.

Acknowledgments

It is so appreciated for support and help from workmates in the project of Miying in Tencent.
Funding: None.

Footnote

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jmai.2019.04.01). J Wu and T Qian are Tencent employees.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-86.
2. Bhavanishankar K, Sudhamani MV. Techniques for detection of solitary pulmonary nodules in human lung and their classifications-a survey. Int J Cybern 2015;4:27-40.
3. Zeng H, Zheng R, Guo Y, et al. Cancer survival in China, 2003–2005: A population-based study. Int J Cancer 2015;136:1921-30.
4. Baldwin DR. Prediction of risk of lung cancer in populations and in pulmonary nodules: significant progress to drive changes in paradigms. Lung Cancer 2015;89:1-3.
5. Camarlinghi N, Gori I, Retico A, et al. Combination of computer-aided detection algorithms for automatic lung nodule identification. Int J Comput Assist Radiol Surg 2012;7:455-64.
6. Naidich DP, Bankier AA, MacMahon H, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. Radiology 2013;266:304-17.
7. Kazerooni EA, Austin JHM, Black WC, et al. ACR–STR practice parameter for the performance and reporting of lung cancer screening thoracic computed tomography (CT): 2014 (Resolution 4). J Thorac Imaging 2014;29:310-6.
8. Tang H, Kim DR, Xie X. Automated pulmonary nodule detection using 3D deep convolutional neural networks. Proc ISBI 2018:523-6.
9. Teramoto A, Fujita H, Yamamuro O, et al. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional

neural network technique. Med Phys 2016;43:2821-7.
10. Aresta G, Cunha A, Campilho A. Detection of juxta-pleural lung nodules in computed tomography images. Med Imag 2017;10134:101343N.
11. John J, Mini MG. Multilevel thresholding based segmentation and feature extraction for pulmonary nodule detection. Proc Tech 2016;24:957-63.
12. Liu Y, Wang Z, Guo M, et al. Hidden conditional random field for lung nodule detection. Proc ICIP 2014;3518-21.
13. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60-88.
14. Dalal N, Triggs B. Histograms of oriented gradients for human detection. Proc CVPR 2005;1:886-93.
15. Rahim MA, Azam MS, Hossain N, et al. Face recognition using local binary patterns (LBP). Glob J Comp Sci Tech 2013;13.
16. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Com Vis 2004;60:91-110.
17. Armato III SG, 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 2011;38:915-31.
18. Setio AAA, 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 2017;42:1-13.
19. Zhu W, Wang Y, Huang Y, et al. DeepEM: deep 3D convnets with EM for weakly supervised pulmonary nodule detection. Proc MICCAI 2018;812-20.
20. Xie Z. 3D region proposal U-net with dense and residual learning for lung nodule detection. LUNA16. 2017. Available online: https://www.semanticscholar.org/paper/D-Region-Proposal-U-Net-with-Dense-and-Residual-for/e629bbeba17b8312fde05b0334760e6743d11a4d
21. Ding J, Li A, Hu Z, et al. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. Proc MICCAI 2017;559-67.
22. Dou Q, Chen H, Jin Y, et al. Automated pulmonary nodule detection via 3d convnets with online sample filtering and hybrid-loss residual learning. Proc MICCAI 2017;630-8.
23. Schmidhuber J. Deep learning in neural networks: An overview. Neural Netw 2015;61:85-117.
24. Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks. Pat Rec 2018;77:354-77.
25. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.
26. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neu Inf Pro Sys 2012;1097-105.
27. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. Pro CVPR 2016;770-8.
28. Chen H, Zheng Y, Park JH, et al. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. Pro MICCAI 2016;487-95.
29. Bychkov D, Turkki R, Haglund C, et al. Deep learning for tissue microarray image-based outcome prediction in patients with colorectal cancer. Med Ima 2016;9791:979115.
30. Ortiz A, Munilla J, Gorriz JM, et al. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. Int J Neural Syst 2016;26:1650025.
31. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 1980;36:193-202.
32. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Pro IEEE 1998;86:2278-324.
33. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. Int J Com Vis 2015;115:211-52.
34. Ozdemir O, Woodward B, Berlin AA. Propagating uncertainty in multi-stage Bayesian convolutional neural networks with application to pulmonary nodule detection. arXiv preprint 2017 arXiv:1712.00497.
35. Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. Proc CVPR 2017;4700-8.
36. Dou Q, Chen H, Yu L, et al. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. IEEE Trans Biomed Eng 2017;64:1558-67.
37. Jiang H, Ma H, Qian W, et al. An automatic detection system of lung nodule based on multigroup patch-based deep learning network. IEEE J Biomed Health Inform 2018;22:1227-37.
38. Winkels M, Cohen TS. 3D G-CNNs for pulmonary nodule detection. arXiv preprint 2018 arXiv:1804.04656.
39. Golan R, Jacob C, Denzinger J. Lung nodule detection in CT images using deep convolutional neural networks. Proc IJCNN 2016;243-50.
40. Anirudh R, Thiagarajan JJ, Bremer T, et al. Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data. Proc Med Ima 2016;9785:978532.

41. Khosravan N, Bagci U. S4ND: single-shot single-scale lung nodule detection. Proc MICCAI 2018;794-802.
42. Xie H, Yang D, Sun N, et al. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. Pat Rec 2019;85:109-19.
43. Huang X, Shan J, Vaidya V. Lung nodule detection in CT using 3D convolutional neural networks. Proc ISBI 2017:379-83.
44. Zhang J, Xia Y, Zhang Y. A Pulmonary Nodule Detection Model Based on Progressive Resolution and Hierarchical Saliency. arXiv preprint 2018;arXiv:1807.00598.
45. Sun W, Huang X, Tseng TLB, et al. Automatic lung nodule graph cuts segmentation with deep learning false positive reduction. Proc Med Ima 2017;10134:101343M.
46. Setio AAA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016;35:1160-9.
47. Sakamoto M, Nakano H, Zhao K, et al. Lung nodule classification by the combination of fusion classifier and cascaded convolutional neural networks. Proc ISBI 2018;822-5.
48. Sihong Chen, Jing Qin, Xing Ji, et al. Automatic scoring of multiple semantic attributes with multi-task feature leverage: A study on pulmonary nodules in CT images. EEE Trans Med Imaging 2017;36:802-14.
49. Madero Orozco H, Villegas OOV, Sánchez VG C, et al. Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. Biomed Eng Online 2015;14:9.
50. Santos AM, de Carvalho Filho AO, Silva AC, et al. Automatic detection of small lung nodules in 3D CT data using Gaussian mixture models, Tsallis entropy and SVM. Eng App Art Int 2014;36:27-39.
51. Lu L, Tan Y, Schwartz LH, et al. Hybrid detection of lung nodules on CT scan images. Med Phys 2015;42:5042-54.
52. Farahani FV, Ahmadi A, Zarandi MHF. Lung nodule diagnosis from CT images based on ensemble learning. Proc CIBCB 2015;1-7.
53. Klik MAJ, v Rikxoort EM, Peters JF, et al. Improved classification of pulmonary nodules by automated detection of benign subpleural lymph nodes. Proc ISBI 2006;494-7.
54. Froz BR, de Carvalho Filho AO, Silva AC, et al. Lung nodule classification using artificial crawlers, directional texture and support vector machine. Exp Sys App 2017;69:176-88.
55. Wang S, Zhou M, Gevaert O, et al. A multi-view deep convolutional neural networks for lung nodule segmentation. Proc EMBC 2017;1752-5.
56. Wang S, Zhou M, Liu Z, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Med Image Anal 2017;40:172-83.
57. Mukherjee S, Huang X, Bhagalia RR. Lung nodule segmentation using deep learned prior based graph cut. Proc ISBI 2017;1205-8.
58. Nithila EE, Kumar SS. Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering. Alex Eng J 2016;55:2583-8.
59. Badura P, Pietka E. Soft computing approach to 3D lung nodule segmentation in CT. Comput Biol Med 2014;53:230-43.
60. Gonçalves L, Novo J, Campilho A. Hessian based approaches for 3D lung nodule segmentation. Exp Sys App 2016;61:1-15.
61. Cheng JZ, Ni D, Chou YH, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep 2016;6:24454.
62. Buty M, Xu Z, Gao M, et al. Characterization of lung nodule malignancy using hybrid shape and appearance features. Proc MICCAI 2016;662-670.
63. Hussein S, Gillies R, Cao K, et al. Tumornet: lung nodule characterization using mult-view convolutional neural network with Gaussian process. IEEE International Symposium on Biomedical Imaging 2017;arXiv:1703.00645.
64. Dey R, Lu Z, Hong Y. Diagnostic classification of lung nodules using 3D neural networks. Proc ISBI 2018;774-8.
65. Wu B, Zhou Z, Wang J, et al. Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction. Proc ISBI 2018;1109-13.
66. Shen S, Han SX, Aberle DR, et al. An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. Exp Sys App 2019;arXiv:1806.00712.
67. Shen W, Zhou M, Yang F, et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pat Rec 2017;61:663-73.
68. Hussein S, Cao K, Song Q, et al. Risk stratification of lung nodules using 3d cnn-based multi-task learning. Proc IPMI 2017;249-60.
69. Zhu W, Liu C, Fan W, et al. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. Proc WACV 2018;673-81.
70. Liao F, Liang M, Li Z, et al. Evaluate the Malignancy

of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-or Network. IEEE Trans Neu Net Lea Sys 2019;arXiv:1711.08324.
71. Zhang F, Song Y, Cai W, et al. A ranking-based lung nodule image classification method using unlabeled image knowledge. Proc ISBI 2014;1356-9.
72. Shen W, Zhou M, Yang F, et al. Multi-scale convolutional neural networks for lung nodule classification. Proc IPMI 2015:588-99.

doi: 10.21037/jmai.2019.04.01
Cite this article as: Wu J, Qian T. A survey of pulmonary nodule detection, segmentation and classification in computed tomography with deep learning techniques. J Med Artif Intell 2019;2:8.