Journal of X-Ray Science and Technology 28 (2020) 591–617
DOI 10.3233/XST-200660
IOS Press

Deep learning-based CAD schemes for the detection and classification of lung nodules from CT images: A survey

Rekka Mastouri[a,*], Nawres Khlifa[a], Henda Neji[b,c] and Saoussen Hantous-Zannad[b,c]
[a] *University of Tunis el Manar, Higher Institute of Medical Technologies of Tunis, Research Laboratory of Biophysics and Medical Technologies, 1006 Tunis, Tunisia*
[b] *University of Tunis el Manar, Faculty of Medicine of Tunis, 1007 Tunis, Tunisia*
[c] *Department of Medical Imaging, Abderrahmen Mami Hospital, 2035 Ariana, Tunisia*

Received 31 January 2020
Revised 17 May 2020
Accepted 30 May 2020

Abstract.
BACKGROUND: Lung cancer is the most common cancer in the world. Computed tomography (CT) is the standard medical imaging modality for early lung nodule detection and diagnosis that improves patient's survival rate. Recently, deep learning algorithms, especially convolutional neural networks (CNNs), have become a preferred methodology for developing computer-aided detection and diagnosis (CAD) schemes of lung CT images.
OBJECTIVE: Several CNN-based research projects have been initiated to design robust and efficient CAD schemes for the detection and classification of lung nodules. This paper reviews the recent works in this area and gives an insight into technical progress.
METHODS: First, a brief overview of CNN models and their basic structures is presented in this investigation. Then, we provide an analytic comparison of the existing approaches to discover recent trend and upcoming challenges. We also introduce an objective description of both handcrafted and deep learning features, as well as the types of nodules, the medical imaging modalities, the widely used databases, and related works in the last three years. The articles presented in this work were selected from various databases. About 57% of reviewed articles published in the last year.
RESULTS: Our analysis reveals that several methods achieved promising performance with high sensitivity rates ranging from 66% to 100% under the false-positive rates ranging from 1 to 15 per CT scan. It can be noted that CNN models have contributed to the accurate detection and early diagnosis of lung nodules.
CONCLUSIONS: From the critical discussion and an outline for prospective directions, this survey provide researchers valuable information to master the deep learning concepts and to deepen their knowledge of the trend and latest techniques in developing CAD schemes of lung CT images.

Keywords: CAD of CT images, deep learning, lung cancer screening, lung nodule detection, lung nodule classification

1. Introduction

Today, lung cancer is considered the leading cause of cancer deaths worldwide. According to the American Cancer Society for 2020, about 228,820 new cases will be diagnosed and approximately

*Corresponding author: Rekka Mastouri, University of Tunis el Manar, Higher Institute of Medical Technologies of Tunis, Research Laboratory of Biophysics and Medical Technologies, 1006 Tunis, Tunisia. E-mail: rekka.mastouri@gmail.com.

135,720 deaths are expected [1]. The five-year survival rate could be increased from 18.6% to 56% when lung cancer is diagnosed and localized at an early stage [2]. However, more than 60% of patients are diagnosed at a metastatic stage [1].

A nodule is a small soft tissue growth (<30 mm) in the lung parenchyma [3]. It corresponds to an opaque lesion on the images. It is considered the main radiological feature of an early-developing lung cancer [4]. The malignancy of a nodule is related to its characteristics: diameter, density, contours and content.

Computed tomography (CT) is the imaging modality generally used to detect and diagnose lung nodules as well as to characterize them and to control their progress, thanks to its high resolution and its low cost. Computer-Aided Detection (CADe) system can overcome these deficiencies by extracting valuable and relevant information on nodules, and can assist radiologists to establish the correct diagnosis

In CAD systems, we distinguish Computer Aided Detection systems (CADe) that detect potential lesions and discriminates nodules and non-nodules such as consolidations and, blood vessels; and Computer Aided Diagnosis systems (CADx) that aim to characterize lesions (type, progression, stage) and to classify them as malignant and benign [5]. Some CAD systems proceed simultaneously as CADe and CADx by identifying suspicious nodules and quantitatively or qualitatively evaluating the selected lesions.

Actually, CAD systems have become essential for the overall assessment and early diagnosis of lung nodules while minimizing the time required for radiologists to interpret CT images. These systems have greatly improved the health care sector's efficiency and performance, specifically the lung cancer screening process. According to the global CAD market, the estimated compound annual growth rate for the year 2025 is 11.6% worldwide and 23% of CAD systems will be intended for lung cancer detection [6, 7].

Traditional CAD systems are based on handcrafted feature extraction engineering such as texture and shape analysis. Hand-crafted features come from direct visual experience and may not be useful because they are abstract and constructed from poor quality of medical image samples. Besides, they suffer from a deficiency of uniformity, normalization, universality, and require an excessive time-consuming.

Recently, researchers have paid a lot of attention to deep learning methods, especially to Convolutional Neural Networks (CNNs), because of their accurate results and time-saving compared to conventional CAD systems. CNN method has found a solution to various learning problems such as the extraction of relevant features, object recognition, accurate detection, and so on. It has sufficiently proven its performance in different fields, including medical image analysis.

This paper reviews the latest CNN methods used in the development of CAD systems for the segmentation, the detection and the classification of lung nodules in 2D/3D CT images. The state of the art comes mainly from recent articles, selected from various databases such as PubMed, IEE Explore and Science Direct from April 2016 to February 2020.

2. CNN overview

Convolutional Neural Network (CNN) is a basic category of deep neural networks which are denoted as a branch of machine learning techniques based on regularized multilayer networks and which are all subtypes of artificial intelligence systems. The CNN model is mainly characterized by its capacity to gather features from several ascertained data, especially from images. Furthermore, it uses a restricted direct supervision in order to maximize the classification, also CNN is already able to auto-define unknown characteristics [8]. The CNN architecture is built by a stack of distinct layers aiming to extract relevant information directly from the input data sets without resorting to hand-crafted engineering

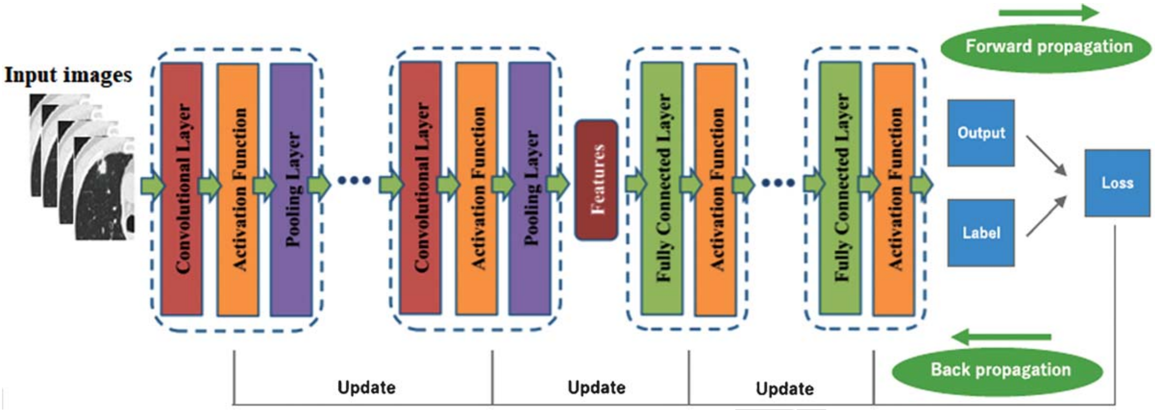

Fig. 1. An overview of a typical convolutional neural network (CNN) structure and the training process.

processes thanks to a differentiable function. As shown in Fig. 1, convolutional layer, pooling layer, fully connected and Softmax layers constitute the principal components of a typical CNN.

- Convolution layer represents the key component of the CNN architecture. Convolution is a mathematical operation where an element-wise product is achieved between a filter (kernel) and the input tensor of the image, thus generating its feature map [8, 9]. It represents the fundamental process for automatically extracting the deep features of images. Feature maps resulting from several convolution operations will be then processed through the pooling layer.
- Pooling layer is a non-linear form of subsampling that reduces the spatial dimension of the input while retaining relevant features [8, 9]. It is therefore, recommended to insert a pooling layer between two consecutive convolutional layers of a CNN architecture in order to reduce over-learning and numbers of parameters in the network. Max pooling is the most commonly used type of pooling formulation in CNN that retains the maximum value of each feature pattern [9, 10]. It is important to note that throughout the convolution and pooling phases, it is possible for pixels to have negative values. An activation function to reset all negative pixels to zero is then required. Rectified Linear Unit (ReLU) is the most commonly used activation function in the literature [10].
- Fully connected layers, consisting of multiple hidden layers, are mainly used at the end of the CNN structure for classification. Final output feature maps are typically flattened, meaning transformed to a vector array, and connected to fully-connected layer(s). Each neuron of a hidden layer is linked to all neurons of another layer while adjusting its weights during the execution of network training [9]. The last fully connected layer is often followed by a non-linear function, Softmax function, which normalizes the actual output values to the target class likelihoods.

CNN's type essentially depends on its convolution kernels. It can be either a 2D-CNN if its convolution kernels are two-dimensional or a 3D-CNN if the kernels are three-dimensional [11]. Furthermore, it is important to mention that a small amount of relevant information included in the input data can considerably decrease the performance of a CNN, and inversely.

2.1. CNN training process

CNN training process is a supervised learning that consists of optimizing internal parameters by maximizing, as far as possible, the consistency between network predictions and ground truth annotations in the training dataset [12]. Back propagation is the most commonly used algorithm for CNN training, whose loss function and gradient descent optimization algorithm ensure self-adjustment of learning

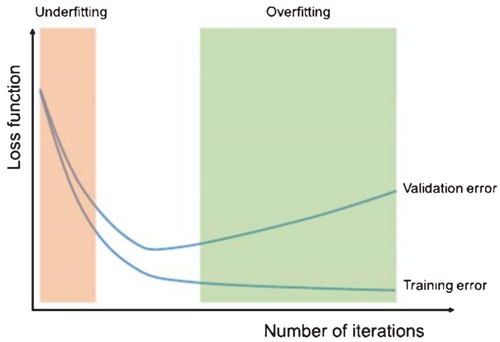

Fig. 2. A routine check of the Loss for recognizing overfitting during the training iteration [12].

parameters and updating of neural weights [12]. Cross entropy is the frequently used loss function while Adam, RMSprop and SGD [13, 14] are the commonly used gradient descent algorithms that represent hyperparameters to be determined according to a desired assignment.

The database is generally divided as follows: one large dataset for training the network (70%), a validation set to evaluate the network while training (20%) and a test set to evaluate the overall and final performance of the CNN (10%), although there are some alternatives, as a cross-validation process which used for training on a small amount of dataset [12].

For cross-validation technique, the training dataset is again randomly divided into equivalent k subsets. Only one sample is retained as a validation dataset and the remaining k-1 sets are used for training [15]. This process is repeated k times with different subset for validation. K-fold Cross-validation is one of the excellent methods commonly used to avoid over-adjustment and to select optimal hyperparameters in order to increase CNN's accuracy.

2.2. Over fitting

Over fitting denotes the situation on which the number of CNN parameters far exceeds the number of features in the input images [10], in a way that stores irrelevant noise rather than learning the signal (Fig. 2). As a result, the accuracy of CNN model decreases due to the high variance during the validation and test phases. Data augmentation and transfer learning are the most commonly used techniques to provide effective CNNs training on a restricted dataset and to solve the over fitting problem [10, 12].

Data augmentation is a powerful solution to overcome this discrepancy. It consists in applying various geometric transformations to the training dataset such as rotations, translations, zooming in and out, horizontal and vertical flipping, so that the CNN will consider them as new data [16]. In addition, Generative Adversarial Network (GAN), which is a type of trend in generative modeling, has become highly recommended as an interesting strategy for augmenting medical data [12, 16]. It consists of generating artificial images that approximate the original database while retaining similar characteristics.

Transfer learning is an effective paradigm and interesting strategy to learn a network in a small dataset [12]. It consists in training a source model on a big database as ImageNet [16] and transferring its knowledge on other task problems. The fundamental principle of transfer learning is to share generic characteristics learned from a large dataset between dissimilar datasets [16]. This portability of generic characteristics is an asset in favor of deep learning in the medical imaging field, which suffers from a lack of sufficient data. Today, there are several CNN models with different preformed structures on ImageNet challenge dataset that are available to the public of which the most used in the medical field are AlexNet [17] VGG [18] Resnet [19] and Inception [20]. In practice, the fine-tuning of these pre-trained architectures is the most common method in medical imaging analysis [12].

The fine-tuning process is a technique allowing preformed weights to be shared and used for the initialization of a new network [12, 16]. To do so, all convolution layers or last few layers, in addition to the fully connected layers, can be fine-tuned and re-trained with the new given data sets, while the first layers can be frozen to retain relevant generic features such as edges, textures, morphologies, intensities, etc. In this way, fine-tuned features will progressively grow more specific to the new dataset [12, 16].

It is worthy noted that there are various methods to avoid overfitting other than data augmentation and transfer learning like batch normalization and dropout, but these are not the subject of our review.

3. CAD system design

In most of the works selected in this literature review, a deep learning-based CAD system consists of five main stages: data collection, data pre-processing, image segmentation (lung parenchyma segmentation and/or nodule segmentation), candidate nodules detection and benign-malignant classification.

3.1. Data collection

Various public databases were found allowing worldwide researchers to have access to a wide variety of data sets for the development of efficient CAD systems. Most of the public databases used by researchers are:

– The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) includes 1018 thoracic CT scans, collected during lung cancer screening in the USA, having 7371 lung nodules with 2669 lesions as nodules $\geq$3 mm along with their associated descriptive XML files [21, 22]. All nodules were labeled by four skilled radiologists who provided a subjective assessment of calcification, speculation, texture, and internal structure to describe the likelihood of malignancy. In addition, all information regarding the diagnosis of each nodule, such as surgical resection, two-year radiological review, and biopsy was archived and placed in the database. It is worth pointing out that, this variety of annotated data forms the suitable basis of most recent publications relevant to the training of CNNs for the detection and classification of lung nodules.
– The LUNA16 dataset that was created by the 2016 LUng Nodule Analysis challenge. It is a reformatted version of the LIDC-IDRI database containing 888 CT scans in Meta Image format (mhd/raw) format with a resolution of 512×512 pixels and a slice thickness of less than 2.5 mm. The LUNA16 database includes all nodules larger than 3 mm annotated by at least 3 out of 4 radiologists, resulting in 1186 nodules. Non-nodules and nodules (<3 mm and > = 3 mm) annotated by less than 3 radiologists are treated as non-nodules [23].
– The Early Lung Cancer Action Program (ELCAP) database, available since 2003 and includes 50 datasets of Low-Dose CT (LDCT) scans with a 1.25 mm slice thickness from a lung cancer screening-related study. The locations of non-nodules and nodules between 2 mm and 5 mm in diameter, identified by expert radiologists were provided. [24].
– The 2009 Automated Nodule Detection Database (ANODE09) contains 55 CT scans with a resolution of 512×512 pixels and a slice thickness of 1.0 mm, of which only 5 scans are annotated by medical experts. In these scans, 39 nodules and 31 non-nodules were marked and can be used for the training of CNN or for the optimization of its parameters. The 50 unlabeled CT scans are only intended to test the algorithms [25].

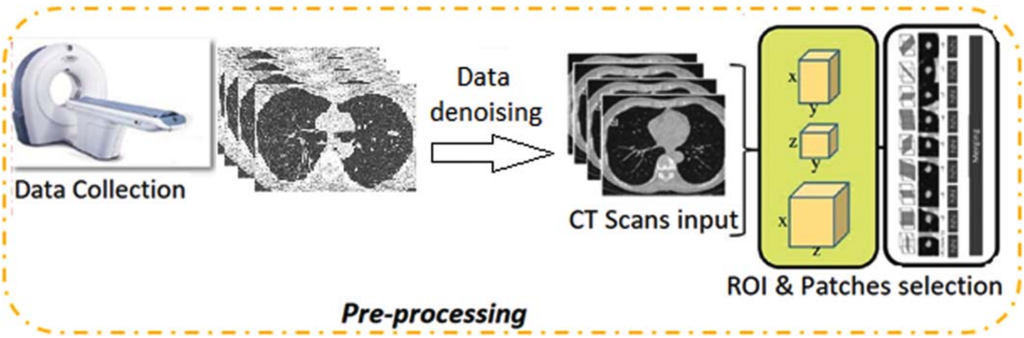

Fig. 3. Pre-processing steps for ROI nodules selection [26].

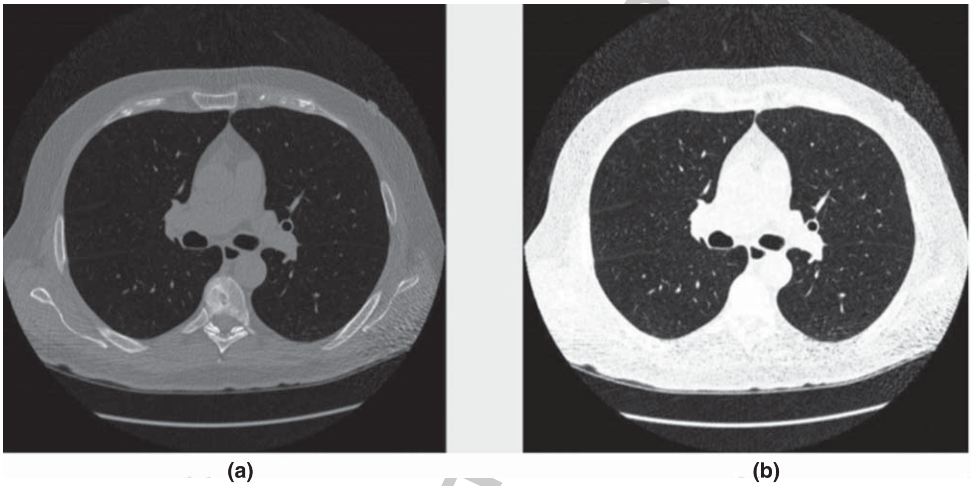

Fig. 4. Preprocessing of CT lung image. a) Original image. b) Image enhanced through Gabor filter [27].

3.2. Data pre-processing and region of interest (ROI) selection

It is sometimes necessary to proceed to some pre-processing steps before transmitting the benchmark of candidate nodules to CNNs (Fig. 3) [26].

3.2.1. Data pre-processing

Computed Tomography uses X-rays to explore the lungs. To prevent the side effects of radiation, radiologists should reduce the radiation dose, which decreases the image quality and generates noises. Thus, a preprocessing phase is essential to improve the quality of raw CT images and to reduce noises and artifacts to facilitate the nodule detection. Figure 4 illustrates an example of image preprocessing using Gabor filter [27].

3.2.2. Region of interest (ROI) selection

To generate ROI dataset for the training of a 2D CNN model, candidate nodules must be cropped around their (x, y, z) coordinates [28]. In addition, multiple 2D patches of different sizes can be centered on the location of each candidate and extracted from various plane orientations. As lung nodule is a small area in the whole CT image, it is easy to crop multiple views of a same nodule, thus providing rich information. Details and specific information about the lung nodules are provided through small

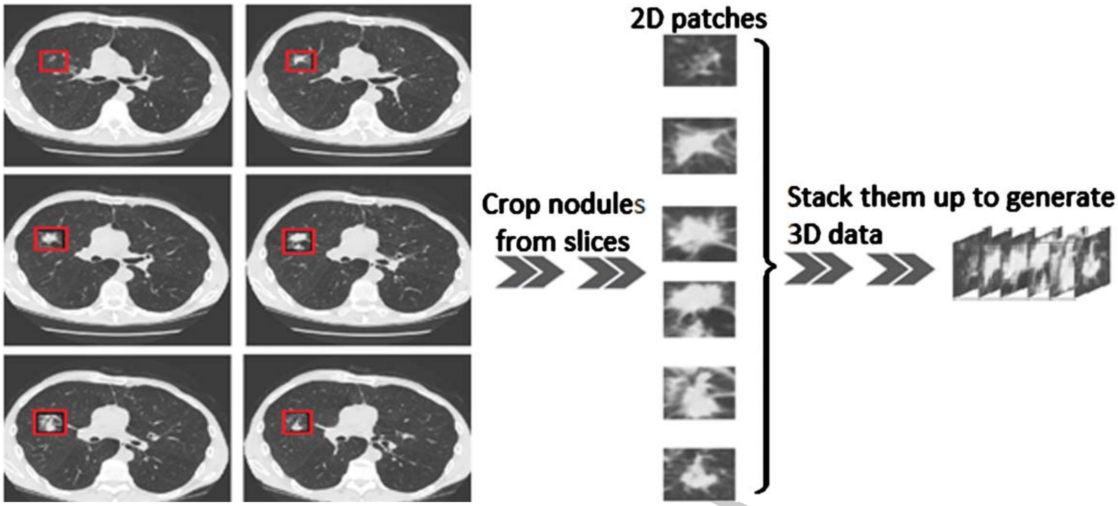

Fig. 5. 2D/3D dataset generation [28].

patches, while information around the nodal tissue is provided through large patches [28]. It is possible to generate 3D patches to train 3D CNNs by grouping all the 2D ROIs of the same nodule. Figure 5 illustrates an example of a 2D and 3D dataset selection.

3.3. Image segmentation

Segmentation of the lung parenchyma or nodules is a fundamental technique aiming to simplify and facilitate both the quantitative assessment of clinical parameters such as size, shape, location, density, texture and the CAD system [12, 29]. However, some lung nodules extraction is challenging because of their intensity, location or texture such as juxta-pleural (nodules directly attached to the pleura's surface), juxta-vascular (nodules connected to vessels) and ground glass nodules, as shown in Fig. 6.

Deep learning-based segmentation is considered as a pixel-by-pixel classification technique for calculating organ probability [12, 29]. This technique is consisting of two phases: the first one is the generation of the probability map using CNN and image patches, and the second ensures the refinement based on the general background of images and the probability map [12]. U-Net [30] is the most widely used architecture for medical image segmentation and it significantly improved the performance of this process. It consists of a contraction way to capture the anatomical structure and a symmetrical expansion way for an accurate localization [31]. Despite the difficulty of capturing both the global and local context, U-Net has allowed the segmentation process to integrate a spatial context at multiple scales. Consequently, it is able to be trained end-to-end from a limited amount of training data [31]. Table 1 presents a brief overview of several works based on deep learning for image segmentation [32–39].

3.4. Candidate nodule detection and false positive reduction

Lung nodule detection and false positive reduction are the two most important approaches to ensure an early diagnosis and reliable assessment of lung cancer [29]. The candidate nodule detection is the principal phase in a CADe system. It consists of extracting all suspicious lung nodules from the parenchyma by removing undesirable structures like alveoli, blood vessels, bronchi, ribs, etc. However, this task is difficult because of the various features (density, shape, texture . . .) of nodules. The sensitivity is generally used to evaluate the performance of the algorithm detection. An efficient

Table 1
Summary of deep learning architectures for Lung/Nodule segmentation

Ref	Database (Trainset/ Testset)	Organ	Architecture	Sen	DSC	PPV
Usman et al. [32]	LIDC-IDRI (356 nodules for training, 45 for validation and 492 for testing)	Nodule	3D Deep Residual U-Net	91.62%	87.55%	88.24%
Wang et al. [33]	LIDC-IDRI (450 nodules for training, 50 for validation and 393 for testing)	Nodule	Multi-View Convolutional Neural Networks (MV-CNN)	83.72%	77.67%	77.58%
Roy et al. [34]	LIDC-IDRI (220 CT images for training and 63 images for testing)	Nodule	encoder-decoder CNN + Level set	–	0.93	–
Huang et al. [35]	LUNA16 (5040 nodules for training, 1411 for validation and 1458 for testing)	Nodule	Fully Convolutional neural Network (FCN).	91.4%	0.793	–
Khosravan et al. [36]	LUNA16 (1186 nodules evaluated with 10-fold cross validation)	Nodule	3D deep multi-task CNN	98%	0.91	–
Mukherjee et al. [37]	LIDC (training sets were 15231 solid nodules and 9675 part-solid nodules, validation sets were 93 solid nodules and 35 part-solid nodules)	Nodule	U-Net FCN + Graph cut	–	Solid = 0.69 Part-solid = 0.65	–
Liu et al. [38]	LUNA16 (2134 images for training, 711 for validation and 711 for testing)	Nodule	Mask R-CNN	79.65%	–	–
Xu et al. [39]	Private (121728 image patches evaluated with 8-fold cross validation)	Lung parenchyma	k-means clustering + CNN	98.8%	0.968	99.5%

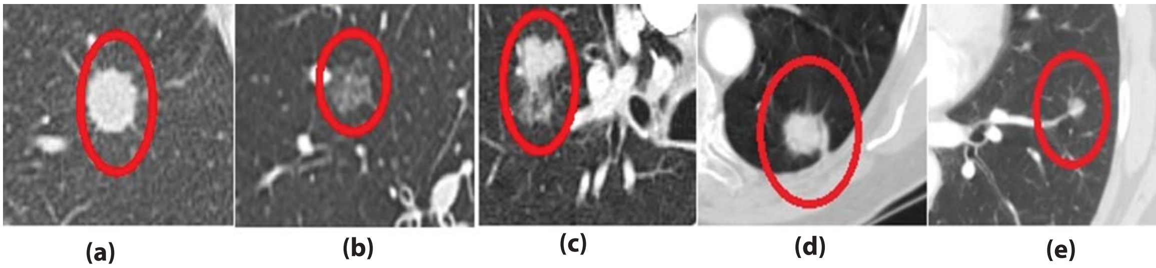

Fig. 6. Different categories of lung nodules: (a) Solid & isolated nodule, (b) Ground Glass Opacity (GGO) nodule, (c) Mixte nodule (d) Juxta-pleural nodule and (e) Juxta-vascular nodule.

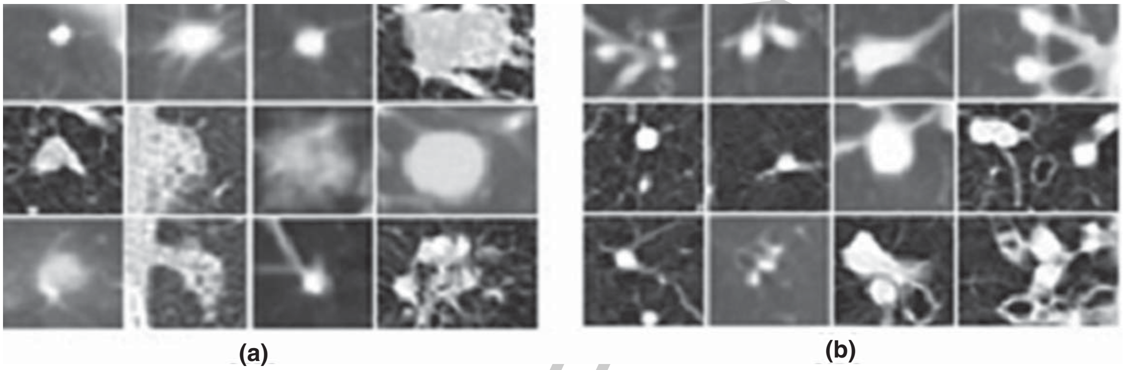

Fig. 7. Sample of candidate nodules: (a) True nodules, (b) False nodules, from the LUNA16 database [23].

CADe system provides a high sensitivity rate, which can be expressed as:

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \tag{1}$$

Where TP indicates the rate of True Positives and FN indicates the rate of False Negatives. A TP indicates that the nodules are correctly identified and a FN indicates that the nodules are incorrectly rejected.

False Positive (FP) reduction is an important step that focuses on identifying the true nodules from all candidates and rejecting the pseudo nodules to increase the performance and efficiency of a CADe system (Fig. 7). The major challenge in the development of CADe is to minimize the FP rate as much as possible without affecting the efficiency of the system, while maintaining a high sensitivity rate. Thus, intensive research works based on deep learning have been developed in order to improve the early and automatic detection of pulmonary nodules [29]. In addition to the standard CNN models dedicated to various recognition tasks, there are numerous models designed specifically for a precise object detection that have demonstrated outstanding performance in the early and accurate detection of lung nodules (Table 2).

3.5. Benign-malignant classification (CADx)

The automatic and accurate distinction between benign and malignant nodules presents a great challenge that has been addressed in several research studies to support the decision-making [11, 40]. However, the classification of lung nodules is still quite difficult due to the diversity of statuses and the complexity of the characteristics of each nodule type [70]. In addition, it is important to know that not

Table 2
Overview of the selected works

Ref	Years	Method	Database	Datasets	Nodule types	FP/exam	CADe/ CADx	Type	Nodule size	Performance results
Usman et al. [32]	2020	3D Deep Residual U-Net	LIDC-IDRI	893 nodules (356 nodules for training, 45 for validation and 492 for testing)	NM	–	CADe	2D/3D	NM	Sensitivity = 91.62% Dice score = 87.55% PPV = 88.24%
Masood et al. [26]	2020	- Enhanced Faster RCNN for nodules detection - Position-sensitive score map for benign malignant classification	LIDC-IDRI	892 images evaluated with leave-one-out cross-validation	NM	2.19	CADe CADx	2D	3 - 30 mm	Sensitivity = 98.1% Accuracy = 97.35% AUC = 98.13%
Monkam et al. [41]	2019	Ensemble 3D CNNs + ELM classifier	LIDC-IDRI	13179 micro-nodules and 21315 non-nodules evaluated with 10-fold stratified cross-validation	Micro-nodules	NM	CADe	3D	? 3 mm	Sensitivity = 96.57% Accuracy = 97.35% F-score = 96.42% AUC = 0.98

Winkels et al. [36]	2019	3D Group Equivariant CNNs	NLST LIDC-IDRI	30000 sets for training/8889 for validation 8582 sets for testing	NM	NM	CADe	NM	NM	Sensitivity = 77% FROC = 88%
Roy et al. [34]	2019	SegNet (encoder-decoder CNN) + Level set	LIDC-IDRI	220 images for training / 63 for testing	Solid and Juxta-pleural nodules	NM	CADe	2D	NM	Dice Score$_{Solidnod}$ = 93% Dice Score$_{Pleural\,nod}$ = 90%
Xie et al. [44]	2019	- Faster R-CNN for candidate nodules detection - Boosting 2D CNN for FP reduction	LUNA16	888 CT scans evaluated with 10-fold cross validation	NM	1/8/4	CADe	2D	3 - 30 mm	Sensitivity = 86.42% AUC = 0.95 Sensitivity = 73.4% Sensitivity = 74.4% AUC = 0.954
Gong et al. [60]	2019	Squeeze-and-Excitation ResNet (SE-ResNet) for nodule detection + Deep CNN for FP reduction	LUNA16	1186 nodules evaluated with 10-fold cross validation	NM	41	CADe	3D	3 - 30 mm	Sensitivity = 93.6% Sensitivity = 95.7%

(*Continued*)

Table 2
(*Continued*)

Ref	Years	Method	Database	Datasets	Nodule types	FP/exam	CADe/CADx	Type	Nodule size	Performance results
Monkam et al. [45]	2018	Three CNNs with different depth	LIDC-IDRI	13179 micro-nodules and 21315 non-nodules evaluated with 5-fold cross-validation	Micro-nodules	NM	CADe	2D	? 3 mm	Sensitivity = 83.82% Accuracy = 88.28% F-score = 83.45% AUC = 0.87
Ali et al. [10]	2018	Reinforcement learning	LUNA16	888 scans, 70% for training (20% were cross-validated) and 30% for testing	NM	NM	CADe	2D	≥3 mm	Sensitivity = 58.9% Specificity = 55.3% Accuracy = 64.4%
Masood et al. [46]	2018	- Deep Fully CNN (DFCNet) for nodule and non-nodule classification - DFCNet for benign and malignant classification	LIDC-IDRI	2669 nodules	Ground glass nodule, mixed nodule and solid nodule	2.9	CADe/CADx	2D	3 - 70 mm	Sensitivity = 83.91% Specificity = 89.32% Dice score = 91.34% Accuracy = 86.02% Sensitivity = 80.91% Specificity = 83.22%

Kaya et al. [47]	2018	Handcrafted features with deep features extracted from AlexNet fed to a cascaded classifier (SVM, RF, Adaboost, KNN) with stacking method	LIDC	1402 nodules evaluated with leave-one-out cross-validation	NM	–	CADx	2D	NM	Sensitivity = 67.37% Specificity = 95.46% Accuracy = 84.70%
Zhu et al. [61]	2018	- 3D faster RCNN - 3D dual path CNN with GBM	LUNA16	888 CT scans evaluated with 10-fold cross validation	NM	8	CADe CADx	3D	$\geq$3 mm	Sensitivity = 85.8% Accuracy = 90.44%
Kido et al. [48]	2018	AlexNet + Multiclass SVM R-CNN + SVM	LIDC-IDRI	-163 nodules (77,080 patches with data augmentation) 30% for training / 70% for testing. -163 nodules evaluated with leave-one-out cross-validation	Ground glass nodule, mixed nodule and solid nodule	NM	CADx CADe	2D	NM	Accuracy = 99.4% Accuracy = 84.7%

(*Continued*)

Table 2
Continued

Ref	Years	Method	Database	Datasets	Nodule types	FP/exam	CADe/ CADx	Type	Nodule size	Performance results
Li et al. [50]	2018	Ensemble of Convolutional Neural Networks (E-CNNs)	JSRT	247 images evaluated with 5-fold cross validation	NM	25	CADe	2D	5 - 60 mm	Sensitivity = 94% Sensitivity = 84%
Silva et al. [51]	2018	Particle Swarm Optimization (PSO) + CNN	LIDC-IDRI	219522 CT slices (12,157 nodules for training, 2,000 nodules for validation and 2,000 nodules for testing)	NM	NM	CADe	2D	NM	Sensitivity = 92.20% Specificity = 98.64% Accuracy = 97.62% AUC = 0.955
Yuan et al. [49]	2018	Multi-view multi-scale CNN + Fisher vector encodings + Multi-class SVM	LIDC-IDRI ELCAP	744 CT scans (80% for training and validation, 20% for testing) 421 nodules for testing	Ground glass optic, Pleural-tail, Well-circumscribed, Juxta-pleural/vascular and Non-nodules.	–	CADx	2D	2 - 5 mm	Overall classification rate = 93.1% Overall classification rate = 93.9%
Wang et al. [33]	2017	Multi-View Convolutional Neural Networks (MV-CNN)	LIDC-IDRI	893 nodules (450 nodules for training, 50 for validation and 393 for testing)	Juxta-pleural, cavitary, and non-solid nodules,	NM	CADe	2D/3D	2.03 -38.12 mm	Sensitivity = 83.72% PPV = 77.58% Dice score= 77.67%

Tajbakhsh et al. [57]	2017	MTANNs sh-CNN LeNet rd-CNN AlexNet	LDCT	For nodule detection: 1057 images evaluated with 5-fold cross validation For nodule classification: 960 images evaluated with 5-fold cross validation	NM	2.7	CADe CADx CADx CADx CADx CADx	2D	NM	Sensitivity = 100% AUC = 0.8806 AUC = 0.7709 AUC = 0.7586 AUC = 0.7813 AUC = 0.7755
Shen et al. [56]	2017	Multi-crop CNN	LIDC-IDRI	2618 nodules evaluated with 5-fold cross validation	NM	–	CADx	2D	3-30 mm	Sensitivity = 77% Accuracy = 87.14% AUC = 0.93
Fu et al. [53]	2017	Handcrafted features + deep features from CNN fed to SVM	LIDC-IDRI	1366 nodules evaluated with 10-fold cross validation	Isolated, Juxta-pleural/vascular	14	CADe	2D/3D	3 - 30 mm	Sensitivity = 90.9% Sensitivity = 78.2%
Hamidian et al. [52]	2017	3D CNN for nodules detection + FCN for FP reduction	LIDC	534 CT scans (833 nodules for training and 104 nodules for testing	Isolated, Juxta-pleural/vascular	15.28	CADe	2D/3D	NM	Sensitivity = 80%

(*Continued*)

Table 2
(*Continued*)

Ref	Years	Method	Database	Datasets	Nodule types	FP/exam	CADe/ CADx	Type	Nodule size	Performance results
Wang et al. [62]	2017	- Multi-branch CNN (gray-scale + LBP+ HOG) - Multi-channel feature fusion CNN	LIDC-IDRI	5820 images, training set (60%), validation set (15%) and test set (25%)	NM	–	CADx	2D/3D	3 - 30 mm	Accuracy = 90.92% AUC = 0.962 Accuracy = 91.75% AUC = 0.97
Dou et al. [59]	2017	Fusion of three 3D CNN with different levels	LUNA16	888 scans / 1186 nodules evaluated with 10-fold cross validation	NM	841	CADe	3D	≥3 mm	Sensitivity = 84.8% Sensitivity = 90.7% Sensitivity = 92.2%
Ding et al. [63]	2017	- Faster R-CNN for nodule detection - 3D DCNN for FP reducctio	LUNA16	888 scans / 1186 nodules	NM	41	CADe	2D/3D	≥3 mm	Sensitivity = 92.2% Sensitivity = 94.4% FROC-score = 0.891

Teramoto et al.[58]	2016	CNN + SVM	Private	104 slices containing 183 nodules evaluated with 5-fold cross validation	Solitary and GGO	4.9	CADe	2D/3D	4-30 mm	Sensitivity = 90.10%
Setio et al. [64]	2016	Multi-view CNN	LUNA16	888 scans / 1186 nodules evaluated with 5-fold cross-validation	Solid, Subsolid and large solide nodules	41	CADe	2D	≥3 mm	Sensitivity = 85.4% Sensitivity = 90.1% AUC = 0.996
Anirudh et al. [65]	2016	3D Multiscale CNN + Super-pixels algorithm	SPIE-AAPM-LUNGx	70 scans (20 for training and 47 for testing)	NM	10	CADe	3D	3–20 mm	Sensitivity = 80%

NM: Not Mentioned. AUC: Area under the ROC curve.

all lung nodules are malignant tumors and cannot all become cancers [11]. Thus, given the numerous images to be daily analyzed by the radiologist and the critical nature of the task, and because of its self-learning and reliable features generation capabilities, many deep learning-based CADx systems have been proposed in the literature for the automatic classification and early diagnosis of pulmonary nodules (Table 2).

4. Selected works

The articles presented in this work were selected from the following databases: Science Direct, IEE EXplore and PubMed. The most relevant keywords used during our searches are: Nodule detection; Lung cancer; Lung segmentation; CAD for lung nodule detection; CT image analysis, CNN for lung nodules classification... Initially, we selected 108 articles. Then, we filtered them according to their pertinence to the subject and based on statistical measures such as sensitivity; accuracy; FP rate; precision and time-consuming. This gave us 73 articles up to February 2020. About 60% of preselected articles are from the last two years.

In 2020, Usman et al. [32] proposed a new approach for the 3D segmentation of lung nodules on CT scans. The proposed approach had two stages; the first was based on a Deep Residual U-Net model allowing the binary segmentation of the lung nodule. Then, the segmented mask was processed by the adaptive ROI algorithm which enables to locate the ROI on the following slice. All the 2D segmented masks of the nodule were concatenated, giving thus a volumetric segmentation result. On second stage, two other 2D patch-wise segmentations of the nodule were achieved on coronal and sagittal planes using two extra Residual U-Nets architectures; and as in the first stage, all the 2D segmented masks of the nodule were concatenated, producing VOIs on both the coronal and sagittal axes. Finally, a consensus model was used to process VOIs of the nodule provided by the three planes and to generate the final 3D segmented nodule. The proposed method was evaluated on 12821 images containing 893 nodules from LIDC-IDRI database and achieved a Dice score, a sensitivity and a Positive Predictive Value (PPV) of 87.55%, 91.62% and 88.24%, respectively.

Masood et al. [26] proposed a new CAD system for the early detection and classification of lung nodules on CT scans based on an enhanced multidimensional Region-based Fully Convolutional Network (mRFCN). First, Authors used a VGG16 architecture ameliorated with a deconvolutional layer as a Faster RCNN model for candidate nodules detection. The accurate ROI selection of lung nodules was provided through a new multi-Layer fusion Region Proposal Network that successfully selected nodules of various positions, shapes and locations. Finally, a position-sensitive score map was used for the malignant and benign nodule classification step. The proposed system was trained using leave-one-out cross validation method and evaluated on 892 images from the public LIDC-IDRI database. Performance of the proposed CADe system achieved a sensitivity of 98.1% with a FP rate of 2.19 per scan, while the proposed CADx system achieved an accuracy rate of 97.91% with an AUC of 98.13%.

In 2019, Monkam et al. [41] developed an ensemble learning of five 3D CNNs with different path sizes to discriminate lung micro-nodules (with diameter <3 mm) and non-nodules from CT images. The authors implemented various fusion strategies (ELM, majority voting, and operator, averaging and auto-encoder) to integrate the five 3D-CNN outputs and construct ensemble models. The Extreme Learning Machine (ELM) strategy presented the best performance results for nodule classification. The proposed system was evaluated on 34494 3D images containing 13 179 micro-nodules and 21 315 non-nodules from the LIDC-IDRI database. It achieved a sensitivity of 96.57%, an accuracy of 97.35% and F-score of 96.42%.

Winkels et al. [42] proposed a new CNN architecture using 3D roto-translation group convolutions (3D G-CNNs) to improve CNN's traditional performance in reducing false lung nodules. The proposed architecture, which based on progressive data augmentation, gave performances that resemble those of pre-trained CNNs with 10× more data without additional setting, providing outstanding efficiency. It was more sensitive to malignant nodules. Performance system was evaluated using about 30000 datasets and 8889 datasets for training and validation, respectively, from NLST database and about 8582 datasets from LIDC-IDRI database for testing. The system achieved a sensitivity to malignant nodules of 77% and a FROC rate of 88% with only 3 epochs that gained training time.

Roy et al. [34] proposed to combine a Deep Learning architecture with level set algorithm to segment lung nodules from 2D CT scans. The authors implemented the method proposed by Shen et al. [43] to extract lung parenchyma and they integrated them into an encoder-decoder CNN to perform a coarse segmentation of nodules. Then, level set algorithm was used to reduce false positives. From the output of CNN, a centroid of each detected object was calculated and a bounding box around the centroid were initialized the level sets. Performance system was evaluated on 220 CT images for training and 63 images for testing from LIDC-IDRI database. It achieved dice scores of 93% for solid nodules and of 90% for juxtapleural nodules with an error of 11% and of 15% respectively. The proposed system proved its efficiency to segment lung nodules compared to literature methods.

Xie et al. [44] developed a novel and automated method based on 2D CNN to detect lung nodule and to reduce false positives. An improved faster R-CNN which integrated a deconvolution layer, to extend the feature map, with two networks for region proposal to fuse the relevant information from the bottom layer, were used. Three sub-networks models (feature extraction network; region proposal network; ROI classifier) were trained and their results were fused to obtain the candidate nodules. Thereafter, a candidate augmentation was used and a boosting CNN model was trained to minimize false nodules in which 3 models are sequentially trained, and each handle harder was mimed than last model. Finally, the outputs of these networks were merged, and the result of the classification was voted out. The algorithm was trained and tested on 888 CT scan from the public LUNA16 database. As results, the sensitivity of 86.42% for nodules detection, and 73.4% and 74.4% for false positive reduction at 1/8 and 1/4 FPs/scan, respectively, with 10-fold cross validation were achieved.

In 2018, Monkam et al. [45] developed three CNN architectures with 1, 2 or 4 convolutional layers to discriminate lung micro-nodules (with diameter <3 mm) and non-nodules from CT images with 3 different patch sizes (16×16, 32×32 and 64×64). The proposed CNNs with appropriate depths were evaluated on the LIDC-IDRI database and proved their effectiveness to distinguish between lung micro-nodules and non-nodules. The model with 2 convolutional layers, in case of 32×32 patches size, presented the best performance. The results, by fivefold cross-validation, showed a sensitivity, an accuracy, an AUC and an F-score of 83.82%, 88.28%, 0.87 and 83.45%, respectively.

Ali et al. [10] developed and validated a Reinforcement Learning (RL) model based on CNN for early lung nodules detection. The authors normalized the datasets by calculating the Z score for each image to reduce the effects of artifacts and different contrast values between CT images. The Data normalization was helpful to fine-tune the input information fed into RL model. Thus, a data augmentation of 888 CT scan from publicly LUNA16 database was made. As a result, a sensitivity of 99.2%, a specificity 99.1% and a Positive Predictive Value (PPV) of 99.1% were achieved for training step. For test step, the model achieved a sensitivity of 58.9%, a specificity of 55.3% and a PPV of 54.2%.

Masood et al. [46] developed a system based on Deep Fully Convolutional Neural Network (DFCNet) to detect and classify the pulmonary nodules in thoracic CT images. An initial nodules classification into 2 classes (nodules vs non-nodules) was done. Then, the authors used a data augmentation for the nodule class to improve the training of the DFCNet model. A second classification into 4 classes was made to show the stage of nodules. The proposed CAD was evaluated on various publicly databases (LIDIC-IDRI, LUNA16, RIDER and Lung CT-Diagnosis) and the results yielded a sensitivity of 83.91%, a specificity of 89.32% and an average FP rate of 2.9 for 2669 nodules from LIDIC-IDRI database.

Kaya et al. [47] proposed a cascaded classification method for predicting lung nodules malignancy. The authors used a CNN model (AlexNet architecture) to automatically extract features from nodules. Then, a hand-crafted feature extraction was made and a cascaded classifier was used. Handcrafted and Deep features were combined to train various classifiers (SVM, kNN, Adaboost, and Random Forest). The proposed method was validated on LIDC-IDRI database and the best performing cascaded classifier was showed an accuracy of 84.70%, a sensitivity of 67.37% and a specificity of 95.46%.

Kido et al. [48] developed a CAD system for the detection and classification of lung nodules based on CNN and R-CNN. In the primary phase, the authors used a pre-trained Alexnet architecture to extract features of 1304 nodules, obtained with data augmentation. Then, these features were used to train a multiclass SVM that classified malignant and benign nodules. The classification results achieved an accuracy of 95.2% without data augmentation and 99.4% with data augmentation. In the second phase, the authors applied an R-CNN framework to detect lung nodules from CT images. R-CNN is a famous object detector based on selective search algorithm that produce bounding boxes and region proposals. Every proposal passed through to a CNN model which contain on its final layer an SVM classifier. The SVM indicated that if the region was a lung nodule and marked its corresponding bounding box. The proposed method proved its efficiency to correctly detect various types of nodules especially the juxtapleural nodules.

Yuan et al. [49] used hybrid descriptors to classify and distinguish different nodule types through a multi-class SVM. The authors extracted statistical and geometric features from multi-view multi-scale CNNs and Fisher Vector (FV) encodings, respectively, to consist the hybrid descriptors. They selected 744 CT scans (1738 nodules and 1000 non-nodules) from LIDC-IDRI database, and 421 nodules from ELCAP database (used for the test) to evaluate the proposed system. The latter achieved an overall classification rate of 93.1% and 93.9% for LIDC-IDRI and ELCAP databases respectively.

Li et al. [50] developed a new CNN-based CAD to detect lung nodules. The authors started by applying a fuzzy mask to enhance images. Then, they proposed 3 different CNNs with different input sizes to detect candidate nodules and an Ensemble of CNNs (E-CNNs) for FPs reduction. To evaluate the proposed system, authors used a 5-fold cross-validation with a free-response receiver-operating characteristic for the test of 154 scans from the publicly available JSRT database. The system showed a sensitivity of 94% and 84% with a FP/image of 5 and 2, respectively.

Silva et al. [51] developed a CNN architecture ameliorate by a Particle Swarm Optimization method for nodule classification. The PSO algorithm was very helpful for optimizing the CNN hyper-parameters. A total of 16157 nodule candidates were selected from the public LIDC-IDRI database to evaluate the performance of the proposed CNN scheme. The experiment showed a sensitivity of 92.20%, a specificity of 98.64%, an accuracy of 97.62% and an AUC of 0.955.

In 2017, Hamidian et al. [52] proposed to train a 3D CNN with selected Volumes Of Interest (VOI) to automatically detect lung nodules from CT scans. The authors converted the 3D CNN into a 3D Fully Convolutional Network (FCN) that efficiently generated the score map of all volume. In addition, the FCN scheme was performed to generate negative examples for the training of a new discriminant CNN. The proposed method was validated on 534 CT scans from publicly available LIDC database and the system reached a sensitivity and a FP rate of 80% and of 15.28/scan, respectively.

Fu et al. [53] proposed to combine thresholding followed by morphological operations with a growing region method, both for lung segmentation and for candidate nodule detection. The authors trained 3 CNNs using 3 different data sets and hand-crafted features were assembled to train an SVM classifier in order to reduce FPs. Performance system was evaluated on 1366 nodules from public LIDC database and reached a precision of 90.90% with a PF rate of 4 per scan.

Wang et al. [33] developed a Multi-View Convolutional Neural Networks (MV-CNN) for nodule segmentation in CT scans. The proposed architecture had three offshoots including six convolution layers to extract deep features of transversal, frontal and sagittal representation of the image. At the end, the 3 CNNs were fused in a fully connected layer to predict the nodule area. The proposed method proved its efficiency to segment several nodule types such as juxta-pleural, cavitary, and non-solid nodules. The system was validated on 893 nodules form publicly LIDC-IDRI database and it achieved a sensitivity of 83.72%, a PPV of 77.58% with an average Dice Similarity Coefficient (DSC) and Average Surface Distance (ASD) of 77.67% and 0.24 mm respectively.

The previous articles [5, 8, 18, 54, 55] reviewed and discussed other relevant works published before 2016. A comparative analysis is provided through Table 2.

5. Discussion

As CT scan is the gold standard for lung parenchyma exploration, most researchers are interested in developing CAD systems for chest CT scans. In the last years, deep learning has gained a lot of attention and several CNN architectures were progressively implemented in CAD schemes for lung nodule diagnosis. The CNN procedure is essentially characterized by its capacity to acquire from several ascertained data. Furthermore, it uses a restricted direct supervision in order to maximize the classification, also CNN is already able to auto-define unknown characteristics. Some methods show convincing results with either high sensitivity rates or low FP rates, but never both at the same time. However, the maintenance of a high sensitivity rate associated with a low FP level is necessary to build powerful CADe systems and to prove their importance in the discrimination of pulmonary nodules. From the chosen works' capacity to correctly identify pulmonary nodules using CNN, the sensitivity varies from 66% to 100% and the false positives range from 1 to 15 per scan. Yuan et al. [49], Wang et al. [33] and Shen et al. [56] applied a novel deep learning architecture named Multi-Crop CNN, which allowed to crop nodule regions with salient information from convolutional feature maps and to apply several max pooling. On the other hand, Tajbakhshet al. [57] realized a general comparison between MTANNs and different CNN models. MTANNs proved their efficiency, as appropriate methods, to detect and classify nodules with less-level semantic features. Some researchers used more than one CNN architecture to improve the performance of the system. Li et al. [50] used an ensemble of CNNs to reduce false positives as much as possible. Monkam et al. [45], Xie et al. [44] and Monkam et al. [41] implemented three CNNs with various patch sizes for candidate nodule detection. Hamidian et al. [52] and Fu et al. [53] applied CNN models for feature extraction and nodule classification, while Yuan et al. [49] and Teramoto et al. [58] applied CNNs just for feature extraction and they used SVM classifier to distinguish between nodules and non-nodules.

As the SVM is considered a binary classifier, it was one of the most widely used machine learning models for the lung nodule classification task [40]. From the selected works, it was found that in some cases, the combination of handcrafted features with deep features achieved a greater performance than using one of them with SVM alone or CNN alone. Consequently, some researchers have combined CNN architecture with SVM classifier to take advantages of both. CNN models were typically used for automatic feature extraction. Kaya et al. [47] extracted deep features using AlexNet architecture and combined them with hand-crafted features to train a cascaded classifier. Yuan et al. [49] extracted statistical features from multi-view multi-scale CNNs and combined them with geometrical features from FV encodings based on SIFT to train a multi-class SVM. In some hybrid CAD systems, classic methods were used to segment the lung parenchyma and detect candidate nodules, while the CNN model was implemented to reduce FPs and discriminate the nodules malignancy [53].

Since CT scans are 3D in nature, 3D CNNs were more appropriate for detecting and classifying pulmonary nodules than 2D CNNs. Moreover, 3D CNN presented a more adaptable ability to model spatial information through 3D convolution and 3D pooling operations than 2D CNN. However, 3D CNNs were little used in the fields of medical image analysis compared to widely used 2D CNN architectures. In this review, about 91% of selected deep learning-based CAD works [45, 46] used 2D CNN architectures.

Only few works of selected papers used 3D CNN for lung nodules detection. Dou et al. [59] applied three multi-level contextual 3D CNNs for the detection of candidate nodules. Then, they merged the networks to minimize FPs.

Meanwhile, Ding et al. [63] started with introducing a deconvolutional layer into Faster R-CNN model to detect candidate nodules. Secondly, they used a 3D CNN to reduce false nodules. Likewise, Masood et al. [26] introduced an enhanced multidimensional Faster R-CNN with a deconvolutional layer for candidate nodules detection. As for Anirudh et al. [65], they trained a 3D CNN using a weakly tagged data from expanded voxel locations. We found that the works discussed above did not attach importance to the detection of small nodules, which is a difficult and a challenging task. In addition, it has been proven that CNNs surpass limits of other supervised learning methods, which improves the performance of lung nodule detection method. We estimate that CNN will reach a great progress thanks to its advantages.

Compared to conventional CAD methods and other machine-learning techniques, deep learning-based CAD system has denoted an important advancement in several applications. It was found that over the past three years, more than 70% of researchers have followed a strong trend towards the development of CAD systems based on deep learning, especially the CNN. An obvious advantage of deep learning over conventional machine learning methods is its less reliance on feature extraction techniques since it can automatically generate relevant features through its convolution and pooling layers [8, 9]. In addition, it enables to directly learn from input data by means of the back-propagation process which allows to automatically adjust the weights of layers according to the supervised learning parameters provided through validated data [8]. In the other hand, the implementation of a traditional CAD system often requires regular human interventions and consistent engineering to ensure its smooth operation. In comparison, the CNN has the capability to greatly improve the global detection rate of lung nodules whenever an output is reported as wrong using iterative self-learning technique [8]. It is worth noting that the more the iterations increase during the learning stage, the greater the sensitivity of the system. Additionally, as CNN learns directly from images, there is no need for a nodule segmentation step, which ensures that most of the information is preserved and nothing is lost [66]. On contrast, in a traditional CAD system, the segmentation of nodules is a necessary step for the extraction of relevant features. Thus, important information can be lost especially when an inaccurate segmentation arises [66].

Feature extraction engineering for the classification of lung nodules in CT images is the subject of several researches in the machine learning field and plays a predominant and important role in the construction of classifiers. CAD system based on handcrafted features is essentially reliant on the expert's professional knowledge and analysis, although this may be subjective [8, 66]. Besides, it suffers from a deficiency of uniformity, normalization and universality [33]. Hand-crafted features come from a straightforward visual experience and usually include a limited number of stringent hypotheses which are varied, and sometimes cannot be useful since they are abstract and constructed with poor medical image samples. However, deep features have an impressive representational capacity thanks to massive training data and a sophisticated structure allowing looking for high-level patterns directly from the lung nodule images [33]. Moreover, performing a classification with a pre-trained network is comparatively fast and time-saving [66]. In addition, since the classification process in the CNN is an end-to-end mechanism and feature calculation is unnecessary, all steps of ROI segmentation, feature calculation and relevant feature selection are useless, allowing for prompt and efficient development of the CAD system [66]. To highlight even more the differences in performance between conventional methods and deep learning methods, previous works are summarized in Table 3 and compared in terms of accuracy, sensitivity, false positive rate and AUC.

Table 3
Comparison between machine learning models and deep learning-based models for the detection and classification of lung nodules

Ref	Method	Database	CADe/ CADx	Accuracy	Sensitivity	FPs/ scan	AUC
Conventional machine learning-based systems							
Gupta et al. [67]	Three layered feed-forward	LIDC-IDRI (1390 nodules)	CADe	–	85.6%	8	0.957
	neural network	ELCAP (40 nodules)		–	70.4%	8	0.847
Gong et al. [60]	3D tensor filtering analysis with	LUNA16 (1186 nodules)	CADe	–	79.3%	4	–
	Random forest classifier	ANODE09 (39 nodules)		–	84.62%	2.8	–
Zhang et al. [68]	3D skeletonization feature with SVM classifier	LIDC-IDRI (168 nodules)	CADe	–	89.3%	2.1	–
Aresta et al. [69]	Rule-based classifier, SVM	LIDC-IDRI (510 nodules)	CADe	–	57.4%	4	–
Hussein et al. [70]	Clustering and SVM	LIDC-IDRI	CADx	78.06%	77.85%	–	–
Nishio et al. [71]	(1) 3D LBP + SVM + TPE	LUNGx challenge and	CADx	79.7%	–	–	0.850
	(2) 3D LBP + XGBoost + TPE	NSCLC	CADx	82%	–	–	0.896
Wei et al. [72]	Spectral clustering	LIDC-IDRI	CADx	85.4%	–	–	–
Deep learning-based systems							
Masood et al [26].	(1) Enhanced Faster RCNN	LIDC-IDRI (892 images)	CADe	–	98.1%	2.19	–
	(2) Position-sensitive score map		CADx	97.91%	–	–	0.981
Xie et al. [44]	2D CNN	LUNA16 (888 CT scans)	CADe	–	74.4%	1/4	0.954
Dou et al. [59]	Ensemble 3D deep CNN	LIDC	CADe	–	92.2%	8	–
Setio et al. [23]	3D deep CNNs	LIDC-IDRI	CADe	–	96.9%	1	–
					98.2%	4	
Shen et al. [54]	2D CNN	LIDC-IDRI	CADx	87.14%	–	–	0.93
Zhu et al. [61]	(1) 3D faster RCNN	LIDC-IDRI	CADe	–	95.8%	8	
	(2) 3D CNN with GBM		CADx	90.44%	–	–	–
Causey et al. [73]	3D CNN	LIDC-IDRI	CADx	95.2%	94.2%	–	0.993

Based on Table 3, we can conclude that the CNN-based CAD system with its advantages represents the most efficient model for the detection and classification of lung nodules in CT images. The significant achievement found when using CNN for nodules detection and classification motivates researchers to utilize it in the precocious analysis of cancers. Nevertheless, we should highlight that there are likewise some inconveniences in using deep learning in actual application:

(1) In order to attain higher results comparing to other techniques, Deep learning systems frequently need a wide number of training data.
(2) The deep learning technique body looks like a black box. We, therefore, have no information about the ideal methodology to entirely understand its system.
(3) These models are extremely complicated and computationally expensive to train even using the support of most powerful GPU hardware.

6. Conclusion and prospects

The detection and classification of pulmonary nodules is an open challenge. They are the subject of much research over the past decade. In this review, we critically described and analyzed the most common techniques used in CAD systems for the detection and classification of pulmonary nodules from CT scans. We reviewed the current trends in the detection and classification of this type of nodules. We have also tried to identify the future challenges and issues that need to be addressed to improve the performance of CADs. This paper is the result of research work analysis from various scientific databases published during the last three years, until February 2020.

A brief description of the structure of a CNN-based CAD system has been introduced in this paper: CAD for lung parenchyma and nodule segmentation, CAD for nodule detection and CAD for nodule classification. In addition, a general description for each selected work was established with statistical results to provide a systematic analysis. We have also noticed that LIDC-IDRI database are widely used by researchers compared to other databases.

While analyzing various research works related the lung nodules detection and classification, we found out that CAD systems based on deep learning achieve a great progression thanks to its advantages. The main advantage of CNN lies in its ability to directly learn from a variety of data sources. In addition, CNN itself can determine unknown characteristics, which maximizes classification with limited direct supervision.

Overall, some selected works have shown good capabilities in the detection and classification of pulmonary nodules in CT images. However, there are still many limitations, such as low sensitivity, high false-positive rate, high time-consuming, low database, low performance rates, etc.

We believe that the accuracy, reliability, and rapid evolution of CAD systems improve both the early detection of lung cancer and the survival rate of affected patients. Therefore, CAD optimization is needed for the early diagnosis of lung cancer. This review is useful for researchers and radiologists to deepen their knowledge of trends and latest techniques in computer-aided diagnosis systems.

In the future, researches on CNN-based CAD systems development for lung nodules detection and classification should mainly focus on:

- Developing new and more robust technique for nodule detection that increase sensitivity and maintain a low number of false positives in order to overcome the challenge of using automated system in daily medical practice.
- Developing a deep learning method able to detect different histological types of lung nodules (solid, non-solid, mixed) at different locations, (isolated, juxta-vascular or juxta-pleural) and with diameter ≤ 3 mm.
- Developing a new CAD system based on powerful feature map visualization techniques to better analyze CNN's decision and transfer it to radiologists.
- Fine-tuning a pre-trained CNN model instead of training it from scratch to increase its robustness and surpass the limitation of annotated medical data.

- The development of in-depth research on GAN models which, with their varieties and various advantages, can solve the problem of the lacking medical database and which can give promising results in the detection and classification of pulmonary nodules.
- Design new CAD systems including two or more of the CNN architectures such as Bilinear CNN [74], Unet-Vnet-Fast-R-CNN [11], AgileNet [75], to address the problem of overfitting that occurs during the training process; this will therefore contribute to better and early lung cancer diagnosis.

References

[1] Cancer Facts & Figures 2020. Atlanta: American Cancer Society. [Cited 2020 May 02]. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2020.html

[2] Lung Cancer Fact Sheet. American Lung Association. [Cited 2019 April 20]. Available from: http://www.lung.org/lung-health-and-diseases/lung-disease-lookup/lung-cancer/learn-about-lung-cancer/lung-cancer-fact-sheet.html.

[3] J.R. Mayo and S. Lam, Computed tomography and the secrets of lung nodules, *Canadian Association of Radiologists Journal* **66** (2015), 2–4.

[4] J. Yuan, et al., Hybrid-feature-guided lung nodule type classification on CT images, *Computers & Graphics* **70** (2018), 288–299.

[5] M.Z. Rehman, et al., An appraisal of nodules detection techniques for lung cancer in CT images, *Biomedical Signal Processing and Control* **41**(2018), 140–151.

[6] Global Computer-Aided Detection (CAD) Market US$ 2.7 Billion by 2025. [Cited 2020 May 5]. Available from: https://www.ihealthcareanalyst.com/pre-screening-diagnostic-technology-adoption-computer-aideddetection-market/

[7] X. Wang, et al., An Appraisal of Lung Nodules Automatic Classification Algorithms for CT Images, *Sensors* **19**(1) (2019), 194.

[8] A. Murphy, et al., The utilisation of convolutional neural networks in detecting pulmonary nodules: a review, *The British Journal of Radiology* **91** (2018), 20180028.

[9] A. Khan, et al., A survey of the recent architectures of deep convolutional neural networks, *arXiv:1901.06032* (2019).

[10] I. Ali, et al., Lung nodule detection via deep reinforcement learning, *Frontiers in Oncology* **8** (2018), 108.

[11] P. Monkam, et al., Detection and classification of pulmonary nodules using convolutional neural networks: a survey, *IEEE Access* **7** (2019), 78075–78091.

[12] R. Yamashita, et al., Convolutional neural networks: an overview and application in radiology, *Insights Into Imaging* **9**(4) (2018), 611–629.

[13] D.P. Kingma, et al., Adam: a method for stochastic optimization, *arXiv: 1412.6980* (2014).

[14] S. Ruder, An overview of gradient descent optimization algorithms, *arXiv:1609.04747* (2016).

[15] N. Aboudi, et al., Multi-objectives optimization of features selection for the classification of thyroid nodules in ultrasound images, *IET Image Processing* (2020). DOI: 10.1049/iet-ipr.2019.1540

[16] C. Shorten, and T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, *Journal of Big Data* **6**(1) (2019), 60.

[17] A. Krizhevsky, et al., ImageNet classification with deep convolutional neural networks, *Adv Neural Inf Process Syst* **25** (2012).

[18] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv:1409.1556* (2015).

[19] K. He, et al., Deep residual learning for image recognition, *arXiv:1512.03385* (2015).

[20] C. Szegedy, et al., Going deeper with convolutions. *In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015). DOI:10.1109/CVPR.2015.7298594

[21] A.P. Reeves, et al., The Lung Image Database Consortium (LIDC): a comparison of different size metrics for pulmonary nodule measurements, *Acad Radiol* **14**(12) (2007), 1475–1485.

[22] S.G. Armato III, 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** (2011), 915–931.

[23] A.A.A. Setio, 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** (2017), 1–13.

[24] ELCAP Public Lung Image Database. ELCAP. Available from: http://www.via.cornell.edu/lungdb.html.

[25] B. van Ginneken, et al., Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study, *Med Image Anal* **14**(6) (2010), 707–722.

[26] A. Masood, et al., Automated decision support system for lung cancer detection and classification via enhanced RFCN with multilayer fusion RPN, *IEEE Transactions on Industrial Informatics* (2020). DOI:10.1109/TII.2020.2972918
[27] B. Abdillah, et al., Image processing based detection of lung cancer on CT scan images, *In Journal of Physics: IOP Publishing* **893** (2017), 012063.
[28] G. Kang, et al., 3D multi-view convolutional neural networks for lung nodule classification, *PloS One* **12** (2017), 0188290.
[29] A. Halder, et al., Lung nodule detection from feature engineering to deep learning in thoracic CT images: a comprehensive review, *Journal of Digital Imaging* **33** (2020), 655–677.
[30] O. Ronneberger, et al., U-net: convolutional networks for biomedical image segmentation, *In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer* (2015), 234–241.
[31] J. Ma, et al., Survey on deep learning for pulmonary medical imaging, *Frontiers of Medicine* (2019), 1–20. DOI: https://doi.org/10.1007/s11684-019-0726-4
[32] M. Usman, et al., Volumetric lung nodule segmentation using adaptive ROI with multi-view residual learning, *arXiv:1912.13335* (2019).
[33] S. Wang, et al., A multi-view deep convolutional neural networks for lung nodule segmentation, *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2017), 1752–1755.
[34] R. Roy, et al., A deep learning-shape driven level set synergism for pulmonary nodule segmentation, *Pattern Recognition Letters* **123** (2019), 31–38.
[35] X. Huang, et al., Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks, *Computerized Medical Imaging and Graphics* **74** (2019), 25–36.
[36] N. Khosravan and U. Bagc, Semi-supervised multi-task learning for lung cancer diagnosis, *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE* (2018), 710–713.
[37] S. Mukherjee, et al., Lung nodule segmentation using deep learned prior based graph cut, *14th International Symposium on Biomedical Imaging. IEEE* (2017), 1205-1208.
[38] M. Liu, et al., Segmentation of lung nodule in CT images based on mask R-CNN, *9th International Conference on Awareness Science and Technology (iCAST). IEEE* (2018), 1–6.
[39] M. Xu, et al., Segmentation of lung parenchyma in CT images using CNN trained with the clustering algorithm generated dataset, *Biomedical engineering online* **18** (2019), 2.
[40] B. Liu, et al., Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: three decades' development course and future prospect, *Journal of Cancer Research and Clinical Oncology* **146**(1) (2020), 153–185.
[41] P. Monkam, et al., Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images, *IEEE Access* **7** (2019), 5564–5576.
[42] M. Winkels and T.S. Cohen, Pulmonary nodule detection in CT scans with equivariant CNNs, *Medical Image Analysis* **55** (2019), 15–26.
[43] S. Shen, et al., An automated lung segmentation approach using bidirectional chain codes to improve nodule detection accuracy. *Computers in Biology and Medicine* **57** (2015), 139–149.
[44] H. Xie, et al., Automated pulmonary nodule detection in CT images using deep convolutional neural networks, *Pattern Recognition* **85** (2019), 109–119.
[45] P. Monkam, et al., CNN models discriminating between pulmonary micro-nodules and non-nodules from CT images, *Bio Medical Engineering On Line* **17** (2018), 1941.
[46] A. Masood, et al., Computer-assisted decision support system in pulmonary cancer detection and stage classification on CT images, *Journal of Biomedical Informatics* **79** (2018), 117–128.
[47] A. Kaya, Cascaded classifiers and stacking methods for classification of pulmonary nodule characteristics, *Computer Methods and Programs in Biomedicine* **166** (2018), 77–89.
[48] S. Kido, Y. Hirano, N. Hashimoto, Detection and classification of lung abnormalities by use of convolutional neural network (CNN) and regions with CNN features (R-CNN), *International Workshop on Advanced Image Technology (IWAIT)* (2018), 1–4. DOI: 10.1109/IWAIT.2018.8369798.
[49] J. Yuan, et al., Hybrid-feature-guided lung nodule type classification on CT images, *Computers & Graphics* **70** (2018), 288–299.
[50] C. Li, et al., False-positive reduction on lung nodules detection in chest radiographs by ensemble of convolutional neural networks, *IEEE Access* **6** (2018), 16060–16067.
[51] G. Silva, et al., Convolutional neural network-based PSO for lung nodule false positive reduction on CT images, *Computer Methods & Programs in Biomedicine* **162** (2018) 109–118.
[52] S. Hamidian, et al., 3D convolutional neural network for automatic detection of lung nodules in chest CT, *Proc. SPIE* **10134** (2017), 1013409.

[53] L. Fu, et al., Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features, *Proc SPIE* **10134** (2017), 101340A.

[54] I.R.S. Valente, et al., Automatic 3D pulmonary nodule detection in CT images: a survey. *Computer methods and programs in biomedicine* **124** (2016), 91–107.

[55] G. Zhang, et al., Automatic nodule detection for lung cancer in CT images: A review, *Computers in Biology and Medicine* **103** (2018), 287–300.

[56] W. Shen, et al., Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification, *Pattern Recognition* **61** (2017), 663–673.

[57] N. Tajbakhsh and K. Suzuki, Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition* **63** (2017), 476–486.

[58] A. Teramoto, et al., Automated detection of pulmonary nodules in PET/CT images: ensemble false-positive reduction using a convolutional neural network technique, *Med Phys* **43** (2016), 2821–2827.

[59] D. Qi, et al., Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection, *IEEE Transactions on Biomedical Engineering* **64** (2017), 1558–1567.

[60] J. Gong, et al., Automatic detection of pulmonary nodules in CT images by incorporating 3D tensor filtering with local image feature analysis, *Physica Medica* **46** (2018), 124–133.

[61] W. Zhu, et al., Deep Lung: Deep 3D dual path nets for automated pulmonary nodule detection and classification, *in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)* (2018), 673–681.

[62] H. Wang, et al., A hybrid CNN feature model for pulmonary nodule malignancy risk differentiation, *Journal of X-Ray Science and Technology* **26**(2) (2018), 171–187.

[63] J. Ding, et al., Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks, *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2017), 559–567.

[64] A.A.A. Setio, et al., Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks, *IEEE Trans. Med. Imaging* **35** (5) (2016), 1160–1169.

[65] R. Anirudh, et al., Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data, *Proc SPIE* **9785** (2016), 978532.

[66] K. Suzuki, Overview of deep learning in medical imaging, *Radiological Physics and Technology* **10** (2017), 257–273.

[67] A. Gupta, et al., Automatic detection of multi-size pulmonary nodules in CT images: Large-scale validation of the false-positive reduction step, *Medical Physics* **45**(3) (2018), 1135–1149.

[68] Y. Zhang, et al., 3D skeletonization feature based computer-aided detection system for pulmonary nodules in CT datasets, *Computers in Biology and Medicine* **92** (2018), 64–72.

[69] G. Aresta, et al., Detection of juxtapleural lung nodules in computed tomography images, *Proc SPIE* **10134** (2017), 101343N.

[70] S. Hussein, et al., Lung nodule characterization using multi-view convolutional neural network with gaussian process, *In 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI)* (2017), 1007-1010.

[71] M. Nishio, et al., Computer-aided diagnosis of lung nodule using gradient tree boosting and Bayesian optimization, *PLoS One* **13** (2018), e0195875.

[72] G. Wei, et al., Lung nodule classification using local kernel regression models with out-of-sample extension, *Biomedical Signal Processing and Control* **40** (2018) 1–9.

[73] J.L. Causey, et al., Highly accurate model for prediction of lung nodule malignancy with CT scans, *Scientific Reports* **8** (2018), 9286.

[74] T.Y. Lin, et al., Bilinear CNN models for fine-grained visual recognition, *Proceedings of the IEEE International Conference on Computer Vision* (2015), 1449–1457.

[75] X. Zhao, et al., Deep CNN models for pulmonary nodule classification: model modification, model integration, and transfer learning, *Journal of X-ray Science and Technology* **27** (2019), 615–629.