

# State-of-the-art DNN techniques for lung cancer diagnosis using chest CT scans: A review

Sakshiwala\* and Maheshwari Prasad Singh\*

\* *Department of Computer Science and Engineering, NIT PATNA, Ashok Rajpath, PATNA-800005, BIHAR, INDIA*

Received 20th of July, 2022; accepted 12th of May 2025

---

## Abstract

This paper presents an in-depth review on the early diagnosis of lung cancer using deep neural networks (DNN). First, the significance of lung cancer and the need of this review is discussed. The various architectures of the DNN, a critical discussion on the existing datasets, various methodologies, challenges in the diagnosis and different evaluation metrics are reviewed. Moreover, the performances of DNN-based techniques for segmentation, nodule detection, and nodule classification are also presented. Further, this review covers the malignancy classification along with the nodule detection tasks. A comparative analysis and limitations of the existing methods is also provided. Thus, this may provide necessary information to all the researchers to prepare a robust methodology for early detection of lung cancer and hence proper diagnosis.

**Keywords:** Convolution neural network, pulmonary nodule, nodule detection, nodule classification, 2D, 3D.

---

## 1 Introduction

Lung cancer is the irregular and unrestrained growth of cells in the lung tissues. Every year millions of people die because of lung cancer. The survival rate of patients suffering from lung cancer is lower than the colon, breast, and prostate cancers collectively [1]. In India, 63,475 people died because of lung cancer in the year 2018 [2]. According to the Indian Council of Medical Research, 27.1% of new cancer cases (3.7 lakhs) will be because of tobacco use [3]. If not treated at an early stage, lung cancer may spread to the other parts of the body [4, 5]. As per the report approximately 12.5% of patients come for treatment at their earlier stage [6]. The causes of lung cancer among patients may be long term tobacco smoking (85%), air pollution (10-15%), radioactive gas, polluted drinking water containing high levels of arsenic material, and other reasons [7]. General symptoms that may signify lung cancer are blood in the coughing, changes seen in the pattern of coughing, harsh coughing, pneumonia, chest pain, unexpected loss in weight, appetite loss, fatigue, hoarse voice and wheezing [8]. The broad classification of lung cancer is shown in Figure 1. There are some works on SCLC (Small cell lung cancer) [9, 10, 11], however, this review focus only on NSCLC (Non-small cell lung cancer).

Lung nodules or pulmonary nodules are masses of tissues with varied diameters ranging from 2 mm to 30 mm [12]. A pulmonary nodule can be further classified as benign or malignant. The benign nodules size is

---

Correspondence to: <sakshiwala.phd19.cs@nitp.ac.in>

Recommended for acceptance by Angel D. Sappa

<https://doi.org/10.5565/rev/elcvia.1597>

ELCVIA ISSN:15108-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

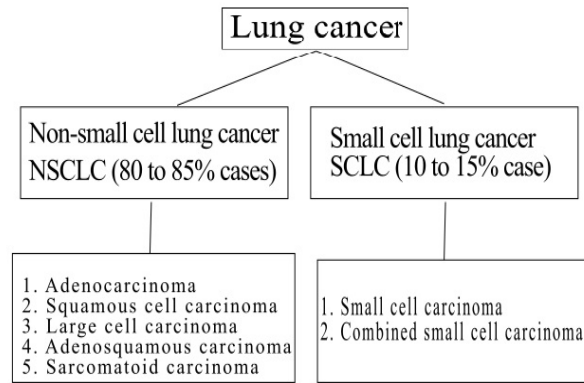


Figure 1: Types of Lung cancer [34]

generally smaller than 6 mm. It is in regular shapes, smoother edges and even color. It does not spread to the other parts of the body and it remains stable for at least two years. Malignant nodules change their size, shape, and appearance. It doubles its size every four months on average. It has an irregular shape, lobulated and spiculated margin. Table 1 lists the various attributes of a pulmonary nodule. The early detection of oddities

| Attributes         | Details                                                                                                                                                                       |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Subtlety           | Distinction difficulty due to its surroundings. Range of ratings from 1 to 5 means highly subtle to obvious                                                                   |
| Internal structure | The inner content of a nodule may be a soft tissue, fluid, fat or air                                                                                                         |
| Calcification      | Sometimes calcium accumulation may appear as nodules in CT images. It has ratings from 1 to 6 meaning popcorn, laminated, solid, non-central, central or absent respectively. |
| Sphericity         | Describes how convex a nodule is in three dimensions. Where rating of 1 means it is linear, 3 means it is ovoid and 5 means it is round.                                      |
| Margin             | States the clarity in the margin. It may be poorly defined rated as 1 or sharp margin rated as 5.                                                                             |
| Lobulation         | A nodule may have one or more lobules present in the nodule margin.                                                                                                           |
| Spiculation        | The presence of spikes shape is marked or not. If spikes are present in the nodule, then it has high chance of nodule to be a malignant nodule.                               |
| Texture            | A nodule can have solid, part-solid or non-solid texture.                                                                                                                     |
| Malignancy         | How likely it is to be a cancerous nodule. Range of ratings from 1 to 5 means highly unlikely to highly suspicious.                                                           |

Table 1: Attributes of a lung nodule

in lung tissues at a curable stage may increase chance of survival. Modern medical imaging modalities like X-ray imaging (chest radiography) [13], sputum cytology, Magnetic resonance imaging (MRI), and computed tomography (CT) [14, 15] help in assisting the diagnosis and treatment of lung cancer. In the lung cancer domain, most researchers work on CT scan [16]. CT scans contain a series of X-ray images taken from many angles. There are many visual challenges in detecting nodule malignancy from images due to the heterogeneity of nodules in texture, shape, and intensity [17]. Further, the complexity of the surrounding environment, a high degree of similarity between pulmonary nodules and surrounding tissues, nodules being similar to the lung walls in terms of intensity, and small-sized nodules having the similar intensity to the surrounding noise adds on to the existing challenges [17, 18, 19, 20]. A radiologist has to search for nodules in every image of a patient's CT scan and mark the malignancy details for each of the nodules based on its characteristic features. It is a tedious process and highly subjective. A study in [21] reported that the overall error rate is 7.6% for every 53 patients. A computer-aided diagnosis (CAD) system, can assist radiologists in diagnosing the pulmonary nodules and detecting nodules [22].

The papers [23, 24, 25, 26], published in the year 2018, 2021, 2021 and 2021 respectively, focus on work that use both feature engineering and deep learning. Among the discussed approaches, only few approaches are based on deep learning. The papers [27, 28] published in the year 2019 and 2020 respectively, focusses explicitly only on the convolutional neural network-based methods. The paper [29, 18], published in the year 2018 and 2020 respectively, focus on all the computer-aided detection methods. The paper [30], published in the year 2019, surveys on many other lung diseases like Pulmonary embolism (PE), Pneumonia, Tuberculosis, Interstitial lung disease (ILD). The paper [31] discusses specifically on the segmentation of the lung cancer radiotherapy. The paper [32] discusses methods for nodule detection and classification but it lacks the correlative

studies of histology findings with CT findings. It also does not signify the importance of nodule attributes in lung cancer diagnosis. Also, the supervised and unsupervised approaches are not discussed explicitly. The paper [33] focus only on the 3D CNN based methods. While, this review paper presents a comprehensive review on different existing deep neural network-based methodologies to detect and classify the lung nodules.

The motivation behind this work is to provide a state-of-the-art review to researchers on deep learning-based techniques for lung cancer nodule diagnosis. Descriptive tables presented in Section 5 brings out the comparative view, on the grounds of main methods employed, diagnostic tasks, datasets used, and performance. Followed by discussion and open issues with future research directions. This study is useful for beginner researchers, providing the latest insights into this domain.

The scope of this review identifies following:

1. Which medical imaging modalities in lung cancer classification is useful?
2. What are the various medical CT imaging datasets used in lung cancer diagnosis?
3. What techniques are used by researchers to preprocess and augmentation of datasets?
4. How are the DNN based techniques applied to lung cancer nodule detection and diagnosis using CT images?
5. How is model performance evaluated?

Further, the paper's organization is as follows: Section 2 presents the articles selection methods. Section 3 describes the DNN architectures used to diagnose lung cancer. Section 4 provides evaluation metrics and Section 5 reviews the state-of-the-art DNN techniques for pre-processing, segmentation and data augmentation. Section 6 reviews lung nodule detection and classification approaches. Section 7 lays out the observations based on literature review to summarize the methods adopted by researchers. In the end, Section 8 concludes this review.

## 2 Methodological Considerations and Approaches

The number of articles using DNN for diagnosis of pulmonary nodules has massively increased in the past three years. Figure 2 shows the trend of published articles in this domain. In this study, 52 recent articles related to detection and diagnosis of pulmonary nodules in chest CT scans are reviewed.

The most recent articles published in the time-period 2018 to 2021, are selected from the following valid sources: (1) ACM, (2) ScienceDirect, (3) PubMed, (4) Springer, (5) Web of Science, (6) IEEE Xplore, and (7) Scopus. Following generic keywords combination are used to search the articles: 'neural network', 'deep neural network', 'convolution neural network', 'lung cancer', 'pulmonary nodules', 'medical image', 'computed tomography', 'lung cancer classification', 'detection', 'segmentation' and 'chest CT scan'.

Initially, the studies from high quality ranked journals according to the 'Scientific Journal Rankings' are selected. After reading their abstract the articles were further selected. In this review, the fifty-two selected articles are reviewed in depth and comprehensive summary of their work are provided.

## 3 Deep Neural Network

For decades, multidisciplinary researchers are working on automating the medical diagnosis systems [35, 36]. There are various architectures of deep neural networks that have recently evolved to meet need of automating the medical diagnosis systems [37, 38]. Below are the most common architectures that are further modified to adapt to the automated lung cancer diagnosis. This section briefly introduces the DNN architectures. These DNN architecture are further modified to get better performance.

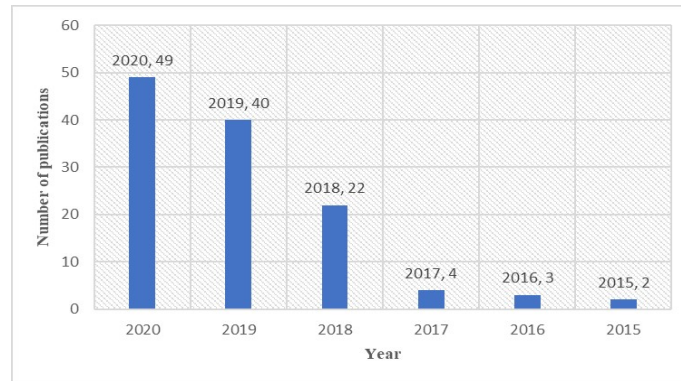


Figure 2: Number of papers published annually having following keywords “deep neural network” and “lung cancer” and “pulmonary nodule” from the year 2015 to 2020.

### 3.1 Convolutional Neural Network

Convolutional neural network (CNN), is a feedforward neural network. CNN uses 2D or 3D array filters to extract meaningful features from the input. The convolutional layer extracts meaningful features. The pooling layer reduces the dimension of feature maps. The normalization layer normalizes the feature maps. Fully-connected layers flatten all feature maps into a single 1D array to compute the output. An activation, for example, softmax or sigmoid, is used to give a final prediction. These models are capable of preserving the spatial and temporal information of the image. The development of CNN has improved the diagnosis of lung using medical images [39, 40, 41].

### 3.2 Faster-RCNN

Faster-RCNN is proposed for object detection tasks in natural images [43]. The architecture of faster-RCNN contains three subnetworks that are feature network, RPN (Region proposal network), and detection network. RPN is a network of convolutional layers that proposes different sized feature maps and a value that indicates object presence. Further, to compensate for the problem of size variation in feature maps, ROI pooling is used. The input to the detection network is the output from both the feature and RPN network. Its classification and the regression layer generate the output class and regressed bounding box.

### 3.3 U-Net

U-Net is the most popular architecture for biomedical image segmentation [18]. The task of U-Net is to classify every pixel in an image. The U-shaped symmetric network architecture has a contracting and expansive path with skip connections between the corresponding layers. The local information extracted in the contracting path maps with the global information in the expansive path. The convolution operations in the contracting path down sample the image while the transposed convolution operation in an expansive path up sizes the image. So, the localization and contextual information are preserved by concatenating the extracted features from previous layers.

### 3.4 ResNet

ResNet are residual network [44]. It is known that more layers in DNN may extract a higher number of features. Deeper networks, even 1000 layers, could be trained with ResNet. It improves the generalization ability of a deep neural network. It is the combination of concept of identity mapping and residual mapping between the network layers. Residual mapping is like adding a small error to the input to get to the final destination. In

identity mapping, an identity matrix transmits the input data forward to avoid the information loss because of vanishing gradients in deeper networks.

### 3.5 DenseNet

DenseNet is a densely connected convolution neural network [45]. The central concept behind DenseNet is the reuse of features extracted from previous layers to the subsequent layers. Channel-wise concatenation combines the input. Instead of using the summation operation in ResNet, DenseNet uses a concatenation operation to add new details to the network's standard information. DenseNet has a strong gradient flow. It has shown higher parameters and computational efficiency.

### 3.6 Autoencoder

Autoencoders are the self-supervised learning network [46]. The encoder part of the autoencoder encodes the input into compressed representation, and then the decoder will decode the compressed representation back to the reconstruction of the input. Training does not have any labels. As a result, they are lossy and turns out to be data-specific. The network architecture for both encoder and decoder can be any CNN, feedforward network, or LSTM (Long short term memory) networks, depending on the use-case. Their main applications are denoising of data, reduction of dimensionality in data, and anomaly detection.

### 3.7 GAN

GANs are generative adversarial networks [47]. The generator and the discriminator of GAN is neural networks. The generator generates fake images, and a discriminator network discriminates against the generator's fake image and the real image. First, the discriminator network trains while keeping the generator constant. When the discriminator has learned to find the generator's flaws, the generator trains by keeping the discriminator constant. The generator learns to fool the discriminator. The minimax loss function balances the training process.

## 4 Evaluation Criteria and Measures

Table 2 shows the parameters of performance metrics. Some of the standard metrics are listed below.

| Parameter           | Meaning                                              |
|---------------------|------------------------------------------------------|
| True positive (TP)  | correctly predicted malignant nodules                |
| True negative (TN)  | correctly predicted benign nodules                   |
| False positive (FP) | predicted malignant nodules that are actually benign |
| False negative (FN) | predicted benign nodules that are actually malignant |

Table 2: Parameters of performance metrics

*Accuracy* [48] is the proportion of correctly classified samples over all samples. The equation of the expression is

$$Accuracy(Acc) = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

*Sensitivity or true positive rate or recall* [48] is the proportion of correctly classified malignant nodules to all true malignant nodules. The equation of the expression is

$$Sensitivity(Sen) = \frac{TP}{TP + FN} \quad (2)$$

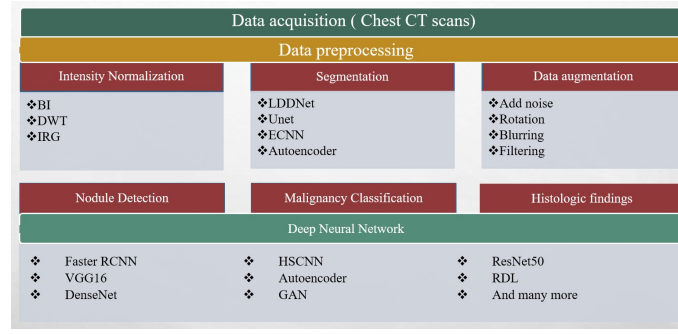


Figure 3: Methods in CAD-based diagnosis of lung cancer.

*Specificity* [48] is the proportion of benign nodules correctly classified to all benign nodules. The specificity equation is

$$Specificity(SP) = \frac{TN}{TN + FP} \quad (3)$$

Sensitivity and specificity metrics are independent of the total samples in the dataset. The higher value of accuracy, sensitivity, and specificity implies that a model has a low misdiagnosed rate, and its classification ability is high. *Precision* is the measure of correctly detected samples among all detected samples.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

*False positive rate (FPR)* is 1-specificity. It is expressed by equation

$$FPR = \frac{FP}{TN + FP} \quad (5)$$

*ROC or receiver operating characteristic curve* is a pixel-wise analysis of the model. It shows the performance of a binary classification model at various thresholds. It is a probability graph plotted between the TP rate in the y-axis and the FP rate in the x-axis. The model has the best performance when the curve is towards the upper left of the graph; that is when the TP rate is one, and the FP rate is 0. *AUC or Area under the ROC curve* [48] evaluates the average performance of the model. A higher AUC value means a model can better classify benign as benign and malignant as malignant. AUC can handle the data imbalance issues and make a better decision about which classifier is the best. *Competition performance metric (CPM)*: The average sensitivity on following FP rates 0.125, 0.25, 0.5, 1, 2, 4 and 8 is called CPM. *Dice coefficient* [49] measures the overlap. The highest overlap value is 1, and it means that the ground truth and the segmented image completely overlap each other. The lowest value, 0, means that they have no overlap. It can be expressed as

$$dice(A, B) = \frac{2 * TP}{2 * TP + FP + FN} \quad (6)$$

## 5 AI-Empowered Lung Cancer Detection: A Journey from Imaging to Diagnosis

Lung cancer diagnosis can be aided by computer programs. This process involves several key steps. Figure 3 shows various methods involved in CAD-based diagnosis of lung cancer. First, chest CT scans are performed to image the lungs, creating a dataset of images for analysis. Details about these image datasets can be found in Table 3. Figure 4 then illustrates the contributions of each dataset. Following image acquisition, the computer program performs various tasks as outlined in Table 4. These tasks are crucial for identifying potential signs of lung cancer.

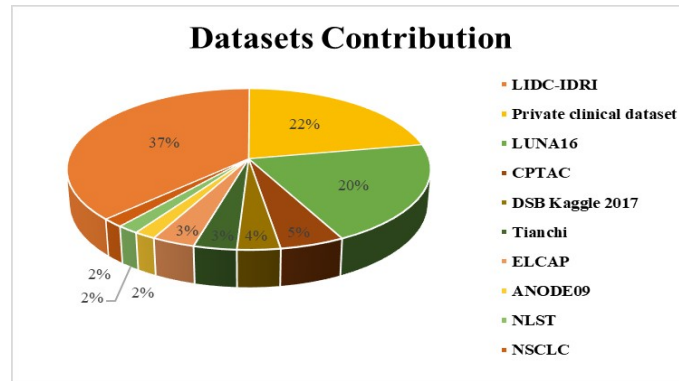


Figure 4: Different contributions for each of the dataset

## 5.1 Imaging the Lungs for Cancer Diagnosis

Computed Tomography or CT imaging modality [50] was first proposed in 1972 by Allan M. Cormack and Godfrey N. Hounsfield. In CT scans, the X-ray tube rotates at different angles. It induces the X-ray photons into patients which flows into their body. These X-ray photons are captured by the detectors, which uses various algorithms to reconstruct the image digitally. Slice is generated by slowly moving the patient's bed forward in the tube. Depending on the kind of CT machine and tissue thickness, the generated image slice varies from 1-10 mm. The above process repeats several times to generate a 3D image. Images are captured using various reconstruction algorithms.

The advancement of technologies in CT scanners has reduce the dose of radiation induced in patients by 70-80%. Modern CT scanners take only one or two seconds or a maximum of thirty seconds to complete the scanning process. Compared to other modalities, CT is a fast, painless, and cost-effective imaging tool. Also, no radiation remains in the body of patients after the scan. CT scanners have helped radiologists visualize small-sized nodules that were earlier impossible to see with a film X-ray. Images acquired by a CT scanner exhibit many differences in terms of kVp, mAs, reconstructed slice thickness, or reconstructed interval [44]. If the difference is less than 0, between reconstruction interval and slice thickness then the reconstructed CT slices may overlap. There are various factors of CT images that are indirectly responsible for the performance of a CAD system. Some of them are section thickness, dose, nodule location, and nodule size [51, 52].

| Dataset         | Number of CT scans | Access type | Available work                                           | Web link                                                                                                                                            |
|-----------------|--------------------|-------------|----------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| LIDC-IDRI       | 1018               | Public      | [82, 46, 21, 83, 84, 57, 69, 85, 86, 53, 88, 74, 89, 63] | <a href="https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI">https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI</a>         |
| LUNA16          | 888                | Public      | [43, 67, 41, 18, 57, 69, 45, 87, 55, 71, 72, 65]         | <a href="https://luna16.grand-challenge.org/data/">https://luna16.grand-challenge.org/data/</a>                                                     |
| DSB Kaggle 2017 | 2101               | Restricted  | [70, 55]                                                 | <a href="https://www.kaggle.com/c/data-science-bowl-2017/data">https://www.kaggle.com/c/data-science-bowl-2017/data</a>                             |
| ANODE09         | 55                 | Public      | [57]                                                     | <a href="https://anode09.grand-challenge.org/Home/">https://anode09.grand-challenge.org/Home/</a>                                                   |
| Tianchi         | 3000               | Public      | [64, 71]                                                 | <a href="https://tianchi.aliyun.com/competition/entrance/231601/information">https://tianchi.aliyun.com/competition/entrance/231601/information</a> |
| CPTAC           | 5043               | Public      | [78, 58, 82]                                             | <a href="https://wiki.cancerimagingarchive.net/display/Public/CPTAC-LSCC">https://wiki.cancerimagingarchive.net/display/Public/CPTAC-LSCC</a>       |
| ELCAP           | 50                 | Public      | [77, 75]                                                 | <a href="https://veet.via.cornell.edu/lungdb.html">https://veet.via.cornell.edu/lungdb.html</a>                                                     |
| NLST            | 75000              | Restricted  | [21]                                                     | <a href="https://cdas.cancer.gov/datasets/nlst/">https://cdas.cancer.gov/datasets/nlst/</a>                                                         |

LIDC-IDRI: Lung Image Database Consortium Image Collection, LUNA16: Lung Nodule Analysis 2016, DSB: Data Science Bowl, ANODE09: Automatic nodule detection 2009, CPTAC: Clinical Proteomic Tumor Analysis Consortium, ELCAP: Early lung cancer action program

Table 3: Available CT scans datasets used in the selected articles

## 5.2 Data Wrangling: Preparing, Segmenting and Augmenting the Data

Lung CT scans contain many incoherent details such as tissues, blood, trachea, chest walls which are irrelevant information for the further processing. Also, the images captured might be of a low quality, which may severely

| Diagnostic task       | Details                                                                                                                                                                                                                                                                    |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Lung segmentation     | The process segments out the lung parenchyma from the unnecessary background of CT scan images. It is an essential step.                                                                                                                                                   |
| Lung nodule detection | The nodule detection task is about detecting a nodule candidate in the Region of Interest (ROI). An ROI may have more than one nodule candidate. All of them need to be detected. Thus, it is a challenging task.                                                          |
| FP reduction          | The detection of FP nodules from all the detected nodules. The false-positive reduction is about decreasing the number of falsely detected nodules that may be blood vessels or lung parenchyma or some other lung region. So that the accuracy of the model may increase. |
| Nodule classification | For fully automated diagnosis, the model may further perform the classification of the detected nodules. The detected nodules are classified into various categories by classifiers.                                                                                       |

Table 4: Lung cancer diagnosis task

influence the performance. So, to boost the potential and reduce the complexity of the algorithm, there is a need for data pre-processing.

### 5.2.1 Data Pre-processing

Generally, to minimize noise and subsequently enhance the quality of the raw CT images, -1000 HU (air) is taken as lower bound and +3000 HU (bone) as upper bound for images. Further, it may be converted from 16-bit to 8-bit (0-255) pixel values [43]. Moreover, in some papers, the pixel values in the CT scan image are transformed to Hounsfield unit (HU) scales and then converted to a range of (0-1) [53], this may decrease the memory usage. The authors [44], created many samples for false positives and abstracts the HU values between 0 to 1.

For identifying the change in the intensity values, there are various filters like Daubechies filters [54], Gaussian filters [55], Sato filter, vessel enhancing diffusion filter [56], Frangi filter [56], 3D bilateral smoothing filters [57], enhancement filter [57], average window filters [57], linear noise filters [57] and many more. For example, each gray pixel's value is replaced by the neighborhood pixels gray matter in the median filtering method. It is a non-linear smoothing filter. Some differential operator based filters like the Laplacian filter will enhance such areas where there is a sudden change in gray pixels. Also, at the same time, it will weaken those areas, which shows a continuous growth in gray pixel values. In [57] the enhancement filters enhance the nodule structure and suppress the tubular shapes. After this step various multi-scale, multi-angle, and multi-view ROI patches were extracted. An unsharp energy mask enhances the intensity value of edges [54]. Improved threshold probability map, Discrete wavelet transform [54] are used to denoise images.

The study on density, pixel intensity, and cumulative density in the image gives much information about pixel distribution. Further, based on the above study, the quality of the images may be improved [58]. The accuracy may be increased by pixel intensive testing process that efficiently interprets pixels in the picture. The undesired pixels, such as noise pixels, may be shifted. The image histogram filter, which is responsible for improving the image's quality, works best for different images [59]. For lung mask production, the optimal threshold computed with FODPSO (Fractional-order darwin PSO-based optimal threshold) [46] can suit all the CT slices. For accurate segmentation, there is a need to improve the lung mask structure done by a hole filling and the boundary smoothing process by applying morphological operations [46]. For the juxta-pleural nodules, the convex hull and dilation [55] pulmonary border extraction, and then mending [56] are useful. Since bones are the high luminance tissue similar to the calcified nodules, intensity normalization makes bones look like normal tissues [55]. The morphology-based operations and multilevel thresholding highlights the nodules positioned near the lungs' boundaries or are lying close to the vascular tissues [44].

Bilinear interpolation is commonly used to resize image patches [49]. However, this technique can introduce limitations when using fixed-size cubes, like 64x64x64 or 32x32x32, centered around nodules. These cubes may not capture the entire nodule volume, potentially reducing the information available to the deep learning model. Ideally, 3D voxel patches should encompass the whole nodule [53]. In this particular study [57], the authors addressed the challenge of working with two different datasets by resampling the CT scans using spline interpolation and a sharp kernel.

In medical imaging, CT scans are often stored in DICOM format. For deep neural network processing, DICOM data may need conversion to a more suitable format, such as PNG or JPG. The pre-processing steps applied to the data can significantly impact the performance of a DNN-based model. While some approaches

forego pre-processing altogether [60], it can be crucial for optimal model performance.

### 5.2.2 Segmentation Approaches

This work proposes a semi-automatic lung nodule segmentation method using a Dual-Branch Residual Network (DB-ResNet) architecture [61]. DB-ResNet leverages multi-scale and multi-view extraction techniques to capture both spatial and intensity information from 2D lung CT slices. For comparison, a Convolutional Neural Network (CNN)-based encoder-decoder architecture, LDDNet [4], is also explored. Unlike traditional U-Net models that employ summation at skip connections, LDDNet utilizes concatenation of feature maps from earlier layers. LDDNet incorporates preprocessing steps like Laplacian and median filtering, potentially improving segmentation accuracy.

Another approach leverages a hierarchical strategy. It first segments lung cancer regions using a K-means algorithm and a deep neural network trained with an explosion function [59]. Then, it extracts local and contextual information using central intensity pooling to achieve a Dice Similarity Coefficient (DSC) of 82.74% and PPV of 79.64% [59].

This paper proposes a segmentation method for 3D tumor areas using an active contour model [62]. This method allows contour points to move freely in all three dimensions, enabling the incorporation of both internal and external energy forces in 3D space. Additionally, a multi-scale Gaussian filter is employed to effectively capture tumor features. To further improve segmentation accuracy and reduce false positives, an enhanced convolutional neural network (E-CNN) architecture is utilized for tumor classification. The E-CNN architecture comprises three convolutional layers followed by two fully-connected layers, with each layer's output being normalized.

However, another approach utilizes a Faster R-CNN model to detect nodules directly from raw CT scans [43]. It reuses weights from a pre-trained VGG architecture and employs non-maximum suppression to reduce the number of candidate regions. The object detection network classifies these proposals, and overlapping candidates are merged based on a close Euclidean distance threshold. Finally, a CNN-based method and a modified fully-convolutional neural network (FCN) are used to further reduce false positives and perform pixel-wise segmentation, respectively. This approach achieves an average Dice coefficient of 0.793, a classification performance measure (CPM) of 0.880, an accuracy of 94.6% with 4 false positives (FPs), and a sensitivity of 95.2% with 19.8 FPs/scan. The Intersection over Union (IoU) is reported as  $70.24 \pm 12.04\%$  [43].

One approach to lung nodule segmentation extracts a comprehensive set of features [46]. It analyzes both 2D and 3D image data, generating thirty-two texture descriptors and twelve geometric properties. These features are then combined using a two-level union operation, creating a single vector called a hybrid geometric texture feature (HGTF). To reduce the dimensionality of this complex feature vector, a two-stage stacked autoencoder is employed. The first autoencoder acts as an encoder, transforming the HGTF vector into a lower-dimensional representation. This encoded representation is then fed into a second autoencoder for further compression. Finally, the resulting low-dimensional feature vector is used by a SoftMax classifier to differentiate between nodules and background tissue. This method achieves promising performance metrics, with a Volume of Interest (VoI) of 1.97, Probabilistic Rand Index (PRI) of 0.92, and Jaccard similarity coefficient of 0.90.

### 5.2.3 Data Augmentation

To address class imbalance and improve model robustness, data augmentation techniques artificially expand the training dataset. This is achieved by applying various transformations to existing images [44]. Common techniques include adding noise (e.g., Salt and Pepper [63], Gaussian, Perlin) to enhance model generalizability. Geometric transformations like random rotation, cropping, flipping (both 2D and 3D), scaling [64], and translation further enrich the dataset [47, 49, 54, 57]. Additionally, techniques like blurring, swapping image patches, and random voxel translations in 3D data can be employed. Figure 5 showcases examples of pulmonary nodule detection using Computer-Aided Detection (CAD) after applying data augmentation. Notably, Regions of Interest (ROIs) can also be randomly shifted along the x, y, and z axes to simulate real-world variations [65].

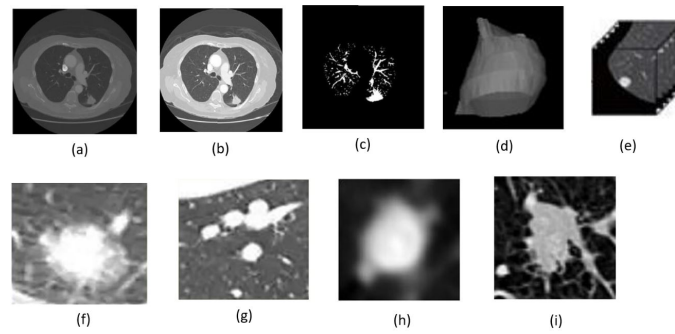


Figure 5: Lung CT images; (a) CT scan slice of lung (b) After pre-processing (c) presence of blood vessels and other artifacts in CT image (d) 3D image of a nodule [78] (e) 3D voxel [65] (f) a TP nodule (g) a FP nodule (h) a benign nodule (i) a malignant nodule

## 6 DNNs: Revolutionizing lung cancer detection and diagnosis

Further, the next step is to predict the lung cancer. Early-stage lung cancer is often asymptomatic, making detection challenging. DNNs hold immense promise for identifying early signs of the disease, leading to better patient outcomes and improved survival rates. In the following sub-section, the models are classified broadly in two categories based on the nodule detection and malignancy classification tasks.

### 6.1 Deep Learning for Lung Nodule Detection

Traditionally, radiologists have relied on their expertise to identify nodules during CT scan analysis. However, this process can be time-consuming and prone to human error, especially with the large number of images often involved. Deep learning is revolutionizing the way lung nodules are detected in CT scans. These algorithms, like super-powered pattern recognizers, can analyze vast amounts of image data to find these small shadows that might be cancerous. This not only helps radiologists by highlighting areas of concern but can also improve accuracy by identifying subtle abnormalities invisible to the human eye. This subsection categorizes various research approaches into two main groups: those that rely on 2D input data and those that utilize 3D input data. The Table 5 show the performance summary of the deep learning based lung nodule detection approaches.

#### 6.1.1 2D deep learning approaches

A method [66] detects lung nodules using a Euclidean distance-based depth map. In this lungs are segmented using region growing, thresholding, and morphological operations. Further, a rule-based classifier reduces the false-positive nodules that exclude non-nodule candidates using the following features major axis length, minor axis length, volume, area, and spherical disproportion. And three 2D CNN networks each for the coronal, axial, and sagittal view whose outputs are fused with the logical-OR operation to classify nodules.

A DNN-based Faster R-CNN is proposed to detect nodules [67]. Different anchor sizes are designed: 12\*12, 18\*18, 27\*27, 36\*36, 51\*51, 75\*75 and 120\*120 to propose nodule regions. VGG16 is the feature extraction model. A score is assigned to every proposed region and then classified. A boosting architecture of 2D CNN reduces false positives. Using three weak classifiers, and then majority voting to give the final classification result. The proposed model has a CPM of 0.790, which is comparatively less.

Author develops a method to segment the lung parenchyma and thus eliminated the vessels [56]. The vascular structures are eliminated using multi-scale Farangi filter. They normalized the dataset using object centralization concept. A four-channel CNN network, each taking different sizes of patches as input, was then used to detect the nodules. Model has sensitivity of 94% at 15.1 FPs/scan and 90.13% at 11.3 FPs/scan.

| Ref ID | Datasets                              | Sample size                               | Methods                                                                         | Performance (%)                   | Nodule category                       | Limitations & Research Gaps                                                                                                                                                                                                                      |
|--------|---------------------------------------|-------------------------------------------|---------------------------------------------------------------------------------|-----------------------------------|---------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [66]   | LIDC                                  | 650 scans or 829 nodules                  | rule-based classifier, multi-view CNN, 2D                                       | Acc 89.895, Sen 85.256, SP 90.658 | nodule, non-nodule                    | instead of multi-view 3D images of nodule can capture more features.                                                                                                                                                                             |
| [67]   | LUNA16                                | 888 scans                                 | Faster-RCNN, VGG16, Boosting architecture of 2D CNN, 2D                         | Sen 86.42                         | nodule or non-nodule                  | fusing more context information about the nodules, such as the connections with the surrounding blood vessels, and the information about the patient such as the medical history report can also be analyzed in the automated diagnosis systems. |
| [56]   | LIDC-IDRI                             | 1006 cases                                | four channel CNN model, Frangi filter, 2D                                       | Sen 80.06                         | nodule, non-nodule                    | 2D processing was used, 3D processing would reflect whole information of the nodule.                                                                                                                                                             |
| [47]   | LIDC-IDRI                             | 1375 benign 930 malignant, 1380 uncertain | GAN, Encoder, Threshold probability map, 2D                                     | Acc 95.32, Sen 94.15, SP 90.78    | anomalous and normal                  | uncertain nodules are rejected. The method requires estimating malignancy scoring threshold. The model is trained only on the benign nodules.                                                                                                    |
| [8]    | lung cancer, breast cancer, leukaemia | 32+34+699 records                         | recurrent neural network, Levenberg Marquardt, glow-worm swarm optimization, 2D | Acc 98, Sen 96, SP 95             | NI                                    | insufficient data                                                                                                                                                                                                                                |
| [68]   | private                               | 500 CT                                    | Wavelet dynamic analysis, K-means clustering, genetic algorithm, 2D             |                                   | NI                                    | model ability to detect all types of nodules is limited.                                                                                                                                                                                         |
| [49]   | LIDC-IDRI                             | 1250 nodules                              | 50-layer ResNet, Siamese network, CNN, 2D                                       | Acc 93.5, Sen 93, SP 89.4         | NI                                    | the correlations between benign-malignancy with different attributes solely or different combinations of attributes.                                                                                                                             |
| [44]   | LIDC                                  | NI                                        | hybrid ensemble CNN, 2D                                                         | Acc 98, Sen 93.5                  | suspicious nodule, FP                 | robustness to diverse lung nodule size is missing.                                                                                                                                                                                               |
| [18]   | LUNA16                                | 1186 nodules                              | U-Net, Proposed dual pooling, ensemble learning, 3D                             | Sen $\geq 90$                     | true, false nodule                    | detects the presence of nodule. The proposed methodology has low CPM score of 0.911. Malignancy attributes could have helped in better detection of nodules.                                                                                     |
| [69]   | LIDC-IDRI                             | 1186 nodules                              | U-Net, VGG-Net, MIP images (1, 5, 10, 15)mm, 3D                                 | Sen 92.7                          | nodule, non-nodule                    | MIP images may not accurately depict the 3D relationships. Calcifications in vessel walls are problem in MIP. MIP images cannot be trusted alone for evaluation.                                                                                 |
| [70]   | LUNA16, Kaggle DSB                    | 2101 patients                             | FCNN in V-Net, 3D                                                               | Sen 92                            | Nodule, non-nodule, benign, malignant | lack of large nodules during training. If a malignant lesion is missed by the CADe model, the subsequent CADx model has no way of classifying that scan as malignant, resulting in the biggest source of false negatives.                        |
| [71]   | LUNA16, Ali Tianchi                   | 1,888 CT scans or 326570 slices           | CNN, ensemble learning, RT-ReLU, 3D                                             | Sen 88.3                          | true nodule, false positive nodule    | high false positive rate makes the model inefficient.                                                                                                                                                                                            |
| [72]   | LUNA16                                | 1186 nodule                               | 3D CNN, FCN                                                                     | Sen 98                            | NI                                    | robustness to diverse nodule types missing.                                                                                                                                                                                                      |
| [74]   | LIDC-IDRI                             | 219522 slices                             | PSO-based CNN, 3D                                                               | Acc 97.62, Sen 92.2, SP 98.64     | nodules, non-nodules                  | evaluated on only single dataset                                                                                                                                                                                                                 |
| [65]   | LUNA16                                | 1186 nodules                              | CNN, cube clustering, 3D                                                        | Sen 87.94                         | nodule, other tissue                  | pure ground glass nodules are missed                                                                                                                                                                                                             |

FCNN: Fully Convolutional Neural Network, RT-ReLU: Random translation Rectified Linear Unit, PSO: Particle swarm optimization, NI: Not informed.

Table 5: Performance of nodule detection models

A study [68] shows that a CNN optimized by genetic algorithm performs better than the traditional CNN to detect and classify pulmonary nodule images. In [8] optimized the parameters of recurrent neural network (with Levenberg Marquardt model) by glowworm swarm optimization. The RMSE value of their proposed RNN-GSO is 0.36.

In the clinical diagnosis, the radiologists come across more benign cases compared to malignant. So, a model trained only with benign cases and validated on malignant cases may be useful at the clinical diagnosis level. Thus, an encoder based unsupervised multi-discriminator GAN trains the network to learn features of anomalous benign images [47]. Further tests the network using non-anomalous malignant images. However, the uncertain nodules are neglected in this work and the method requires estimating the malignancy scoring threshold.

Further, many researchers in their dataset sampling do not consider the nodules having label '3' malignancy because they are difficult to classify. Authors [49] addressed this difficulty. To correctly classify the marginal nodules, considers two weight-shared networks (Siamese): regression and classification module. The regression module outputs nodule attributes scores and classification module predicts malignancy labels. While training both the networks reflect the loss.

For addressing the problem of variations in CT acquisition, an ensemble of VGG, DenseNet, and ResNet is used to reduce false-positive nodules [44]. In this a physics-based data augmentation technique (PBDA) on two factors: slice thickness and patients' dose during CT scans is introduced. In PBDA, the average of the group of that continuous slice is taken instead of taking each slice. And the number of photons was set to match the

noise intensity, and Bowtie filter thickness to calculate per-patient attenuation.

2D approaches struggle to capture depth and spatial relationships inherent in 3D data. It can be difficult to understand why a model makes a particular prediction, hindering interpretability and debugging.

### 6.1.2 3D Deep Learning Approaches

To prevail over the weakness of 2D image constraints, various 3D image based novel algorithms has been evolved.

A scheme suggested a two-stage detection method [18]. In the first stage, preliminary nodules are detected by a new sampling strategy with an improved U-Net based segmentation approach. The second stage proposes a random mask approach for data augmentation and dual pooling. The dual pooling merged the central pooling to preserve features from the center and central cropping to get multi-scale features from the patch. The 3D dual pooling SeResNet, DenseNet, and IncepNet architectures individually predict malignancy probability of a nodule. The final prediction result is the average of the outputs. The proposed methodology has a CPM score of 0.911.

Following method [69] resizes each image to a 1mm slice spacing in the z-direction using maximum intensity projection (MIP). They merged outputs from four U-Net architectures using contour retrieval mode and approximation method. An ensemble of two 3D CNN VGGNet, took 16 and 32 pixels cube as input. A weighted linear combination, fused the output probability. However, MIP images may not accurately depict the 3D relationships because calcification in vessel walls may appear as nodules. Also, MIP images cannot be trusted alone for evaluation.

Method proposed in [70] can both detect and diagnose the components. CADe has a 3D segmentation network that labels every voxel with nodule probability. A 3D scoring network estimates nodule probability. CADx ranks each nodule according to its malignancy and provide a probabilistic malignancy score for each CT scan. However, suppose the CADe model misses a malignant lesion. In that case, the subsequent CADx model has no way of classifying that scan as malignant, resulting in the most significant source of false negatives.

Radiologists vary in their perspectives during nodule detection. So, authors [71] suggested an amalgamated CNN where they used three CNN architectures, which take different scales (32\*32\*32 with five layers, 64\*64\*64-with seven layers, 96\*96\*96-with 9 layers) of segmented images as input. Further random translation-ReLU and AdaBoost is used to fuse the results. Then multiple different CNNs learn more non-nodular structures. However, they did not consider nodules less than 5mm for training. With 8 FPs/scan, sensitivity is 91.4%.

Moreover [72] detected lung nodules by a 3D fully convolutional neural (FCN) network. To reduce the false positives, the voxels output from FCN applied a 3D non-maximum suppression operation, thresholding, and finally averaged the score of classifiers.

There are many hyperparameters used in CNN training [73], and it may directly affect the model accuracy. A CNN-based PSO [74] optimizes the following hyperparameters: number of filters in the convolutional layers, number of neurons in the hidden layer, size of trainable filters, type of pooling, number of batches in training, and probabilities of dropout regularization in convolutional and fully connected layers. However, they have just optimized the CNN parameters, and nothing specific is in the model that would characterize a lung nodule, for example, attributes malignancy.

Moreover a method detects nodules by extracting multi-scale cubes and then cluster it with density-based spatial clustering of applications with noise (DBSCAN) [65]. The advantage is that the model can choose the best scale combination from multi-scale according to the need. The CPM score of the model is 0.7967. However, a particular type of nodule, pure ground-glass nodules, are missed by this approach because of no solid components present in it.

## 6.2 Discerning Danger: Classifying Benign and Malignant Nodules

Not all nodules are created equal! Doctors face the challenge of discerning between benign and malignant nodules. This involves a careful analysis of factors like size, shape, and internal characteristics. Through imaging tests and potentially biopsies, they can classify the nodule. This distinction is crucial, as benign nodules are typically harmless, while malignant ones may signal cancer. Early and accurate classification paves the way for appropriate treatment, making a world of difference in patient outcomes. This subsection examines how different malignancy classification techniques are chosen depending on the amount of information available from the patient (2-dimensional vs. 3-dimensional). The Table 6 show the performance summary of the deep learning based malignancy classification approaches.

### 6.2.1 2D Classification Techniques

There are some works that consider handcrafted features with in-depth features. The study in [54] comparatively studies the traditional methods and CNN techniques. In traditional methods, Daubechies filters extracts texture, density, and morphological features. SVM (Support Vector Machines) classifier classifies the nodules. It achieved a low accuracy of 77.5%. Using GANs, and the eight layered CNN architecture shows better accuracy 93.9%, and ROC of 0.934. In [75], VGG-Net-f CNN architecture, has only one ReLU layer, dense layer, max-pooling layer, and classification layer. It used externally extracted features: solidity, circularity, discrete fourier transformation of radial length function, histogram of oriented gradients and second-order moments of the image. The model achieved a precision of 95.26%.

The optimization algorithms are well-known to ease the learning task. In [58], the authors used the VGG19 network for image segmentation. Then using a hybrid intelligent spiral optimization-based generalized rough set approach, reduced fifty features to fourteen useful features. Then using boosting ensemble classifiers and IDNN (Improved deep neural network), classifies the cancerous features. The model shows a precision of 0.913 and F-measure of 0.921.

With the ideology to develop a model that could show good performance for small datasets, authors in [48], designed two different architectural models, the Multi-Channels-Multi-Slices-2D-CNN (MCMS-2D-CNN) and Voxel-level-1D CNN (V-1D CNN) model. Trains the network on only 18 benign and 49 malignant nodules. In MCMS-2D-CNN, there is one specific channel that takes a raw image as input. Then more channels with LBP features, gradient features and HOG features are added depending on the criteria that it helps to improve accuracy. The V-1D CNN has two convolutions, two pooling, and one global average pooling layer. Then a voting algorithm is used to classify the nodules as benign or malignant. But because of the unbalanced dataset the V-1D CNN shows Acc of only 0.78, Sen of 0.8, SP of 0.53 and AUC of 0.71. The model shows an insufficient learning. The performance of MCMS-2D-CNN is shown in Table 6.

A different scheme uses meta-learning method, where the output of the primary learners is input dataset to the secondary learners [76]. The eight variants of CNN, namely AgileNet, AlexNet, CifarNet, Transfer AlexNet, Transfer CifarNet, GoogLeNet, ResNet and Inception ResNet extract features and predicts the malignancy value for each patch. The predicted values for each patch are then used as a new dataset for the following ten ensemble learners VOT (majority voting), AVE (averaging), KNN (K-nearest neighbour), SVM, NB (Naive Bayes), DT (Decision Trees), MLP (Multi-layer perceptron), RF (Random Forest), GBRT (Gradient Boosting Regression Trees), and AdaBoost (Adaptive Boosting). With the above approach following limitations exists: (i) training time might be longer because of ensemble learners, (ii) multi-scale feature may have better performance than multiple learners, (iii) The diversity of single primary learners is not considered, and all are from deep CNNs.

Lakshmanprabhu et al. [77] presented an optimal DNN based technique. Texture and Wavelet-based features are directly extracted. Features dimension were reduced using LDA (Linear discriminate analysis). The model was then optimized with Modified gravitational search algorithm (MGSA). Though accuracy of this technique is dependent on traditional image preprocessing pipeline.

The paper [64] suggested a multi-view knowledge-based semi-supervised adversarial classification model. It is the extension of the work proposed in [78]. The network uses semi-supervised reconstruction network

| Ref ID | Datasets                                                   | Sample size                                                | Methods                                                                                              | Performance (%)                | Nodule category                                | Limitations & Research Gaps                                                                                                                                                                                                                                   |
|--------|------------------------------------------------------------|------------------------------------------------------------|------------------------------------------------------------------------------------------------------|--------------------------------|------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [54]   | LIDC-IDRI                                                  | 1704 M, 1840 non-cancerous                                 | CNN, GAN, 2D                                                                                         | Acc 93.9, Sen 93.4, SP 93      | non-cancerous, benign, malignant               | low performance                                                                                                                                                                                                                                               |
| [75]   | Emphysema Database, ELCAP                                  | 124 and 180 slices respectively                            | external feature extraction, CNN, 2D                                                                 | Sen 69.56                      | various groups in the database                 | used data is too small containing a low number of nodules.                                                                                                                                                                                                    |
| [58]   | Cancer imaging archive                                     | 5043 CT images                                             | IDNN, intelligent spiral model                                                                       | Acc 92.1, Sen 92.3             | cancer, non-cancer                             | high processing time                                                                                                                                                                                                                                          |
| [48]   | pathological dataset (D2:lung nodules)                     | D2:67 nodules                                              | multi-channel, multi-slice 2D-CNN, 2D                                                                | Acc 77, Sen 76, SP 46          | B, M                                           | the proposed CNN models still didn't fully study all the information from the dataset. Missing slices may lead to loss of topology information.                                                                                                               |
| [76]   | LIDC-IDRI                                                  | 743 nodules                                                | Eight deep CNNs, Ten fusion methods, ensemble learners, 2D                                           | Acc 84, Sen 88.6               | B, M                                           | training time might be longer because of ensemble learners. Multi-scale feature may have better performance than multiple learners. 2D CNNs are employed in the study. The diversity of single primary learners is not considered and all are from deep CNNs. |
| [77]   | ELCAP                                                      | 50 CT scans                                                | MGSA, LDA, DNN, 2D                                                                                   | Acc 94.56, Sen 96.2, SP 94.2   | normal, B, M                                   | insufficient data.                                                                                                                                                                                                                                            |
| [64]   | LIDC-IDRI, Tianchi, LUNGx                                  | 1301 B, 644 M, 1839 unlabeled                              | GAN, Autoencoder, 2D                                                                                 | Acc 92.53, Sen 84.94, SP 96.28 | B, M, unlabelled data                          | uncertain nodules (median malignancy level=3) were neglected during training process. With the 9 mentioned views they can extract only few information from 3D to 2D. Instead of OA, HVV and HS the multi-view slices could be used                           |
| [78]   | LIDC-IDRI, TCIA                                            | 1945 nodules                                               | MV-KBC, 2D                                                                                           | Acc 91.6, Sen 86.52, SP 94     | B, M                                           | supervised model.                                                                                                                                                                                                                                             |
| [79]   | LIDC-IDRI                                                  | 2669 nodules                                               | deep CNN, AdaBoost, GLCM, 2D                                                                         | Acc 89.53, Sen 84.19, SP 92.02 | B, M                                           | more processing time.                                                                                                                                                                                                                                         |
| [80]   | LIDC                                                       | 743 nodules or 375 M, 368 B                                | hybrid CNN of LeNet and AlexNet, 2D                                                                  | Acc 82.2                       | B, M                                           | three channels of input are homogeneous                                                                                                                                                                                                                       |
| [81]   | Integrated lung nodule database, Kyoto University Hospital | 412 B, primary lung cancer 571, metastatic lung cancer 253 | CNN, transfer learning, 2D                                                                           | Acc 68                         | B, primary lung cancer, metastatic lung cancer | ignored all nodule-specific features, such as nodule size and type, only investigated the effect of smaller image sizes                                                                                                                                       |
| [82]   | LIDC-IDRI, clinical datasets from TCIA                     | NI                                                         | shallow CNN, 3D                                                                                      | Sen 80, SP 64                  | NI                                             | did not assess the effect of variability in CT acquisition parameters across patients and institutions on the performance of our models, did not assess the impact of inter-reader variability of segmentations                                               |
| [41]   | LUNA16, TCIA                                               | 1186 nodules                                               | 3D CNN                                                                                               | Sen 88                         | patient has cancer or not                      | infers the labels for each nodule of the patient                                                                                                                                                                                                              |
| [83]   | LIDC                                                       | 1226 nodules; 431 M, 795 B                                 | encoder-decoder structure, manifold learning, 3D                                                     | Acc 90, Sen 81, SP 95          | B, M                                           | model under performs in terms of sensitivity.                                                                                                                                                                                                                 |
| [84]   | LIDC-IDRI                                                  | 554 B, 450 M                                               | 3D ResNet-22, multi-stream CNN, multi-task learning                                                  | Acc 93.92, Sen 92.60, SP 96.25 | malignancy, 8 attributes                       | manually adjust the weight combination of the multi-task loss function                                                                                                                                                                                        |
| [57]   | LUNA16, ANODE09, SPH6                                      | 2361 nodules                                               | multi-region proposal network, Faster-RCNN, mRPN, 3D-DCNN                                            | Acc 98.51, Sen 99.1, SP 94     | B, M                                           | the performance was relatively less accurate in detecting micro nodules                                                                                                                                                                                       |
| [85]   | LIDC-IDRI                                                  | 635 B, 510 M                                               | CNN, Soft activation mapping, U-Net, Global maxpooling, 3D                                           | Acc 99.13, Sen 97.05, SP 99.21 | NI                                             | performance dependent on autoencoder part.                                                                                                                                                                                                                    |
| [86]   | LIDC-IDRI                                                  | 431 M, 795 B                                               | 3D CNN, handcrafted features, SVM, 3D                                                                | Acc 88.66, Sen 82.60, SP 91.95 | B, M                                           | sensitivity is low.                                                                                                                                                                                                                                           |
| [45]   | LUNA16                                                     | 1004 nodules                                               | 3D DPN, HOG, LBP                                                                                     | Acc 93.78                      | B, M                                           | evaluation of model on a single metric does not show model efficiency.                                                                                                                                                                                        |
| [53]   | LIDC                                                       | 4,252 annotations                                          | 3D CNN, HSCNN                                                                                        | Acc 84.2, Sen 70.5, SP 88.9    | B, M                                           | only 5 malignancy attribute is considered. Model achieves low accuracy (0.842) compared to multi-crop CNN (0.8714).                                                                                                                                           |
| [87]   | LUNA16                                                     | 1004 nodules (positive:450)                                | 3D contextual and spatial attention, ensemble of 3D dual path network                                | Acc 90.24, Sen 92.04           | B, M                                           | low performance.                                                                                                                                                                                                                                              |
| [55]   | LUNA16, DSB 2017                                           | 1186 nodules from LUNA, 832 from DSB                       | modified U-Net, 3DCNN, 3D RPN, Leaky Noisy-OR method                                                 | Acc 81.42                      | probability of lung cancer for that subject    | growing speed of nodules is not considered. small nodules detection accuracy is not considered.                                                                                                                                                               |
| [88]   | LIDC-IDRI, private hospital                                | 509 M, 635 B                                               | supervised: 3D CNN, graph regularized sparse representation; unsupervised: k-means, proportional SVM | Acc 78.06, Sen 77.85, SP 78.28 | B, M                                           | the average accuracy of the proposed unsupervised model is low 78.06%. The accuracy of the supervised model is also low 91.26%. The unsupervised learning methods like GANs and autoencoders are promising                                                    |
| [89]   | LIDC-IDRI                                                  | 1018 cases                                                 | CNN, QIF, Random forest, 3D                                                                          | Acc 94.6, Sen 94.8, SP 94.3    | nodule, non-nodule or likely B, likely M       | radiomic attributes may improve performance.                                                                                                                                                                                                                  |
| [95]   | HarvardRT, Radboud, Moffitt, MUMC, MSPORE, Maastro, RIDER  | 317+147+307+200+90+101 +32 patients                        | CNN, transfer learning, 3D                                                                           | NI                             | NI                                             | the retrospective nature of this study, the input data space, opaqueness of deep learning networks                                                                                                                                                            |
| [63]   | LIDC                                                       | 1011 nodules                                               | 3D-DenseNet-40, 3D-ResNet-50                                                                         | Acc 91.47, Sen 91.26, SP 91.67 | B, M                                           | non-visual attribute labels like age, gender, genomics data, family history and histological information can also be incorporated into network training                                                                                                       |
| [21]   | LIDC-IDRI, NLST                                            | 103 nodules                                                | U-net, feature pyramid network, leaky noisy-OR gate, 3D                                              | Acc 85                         | B, M                                           | features were described and observed by a single reader. The hand-crafting of the classification patterns of the activation maps                                                                                                                              |

B: Benign nodule, M: Malignant nodule

Table 6: Performance of malignancy classification models

trained on labelled and unlabelled data. The features extracted from each convolution and identity blocks of the encoder are passed through the transition layers to each respective layers of classification network. The uncertain nodules (median malignancy level=3) are not used during the training process. The multi-view slices may be a better substitute to OA, HVV, and HS patches.

Taking a similar scheme [78], a  $64 \times 64 \times 64$  sized voxels, and nine 2D views (coronal, sagittal and axial views) along the six diagonals of a cube are extracted. The U-Net is used to segment the nodules on each slice. Further, HVV (heterogeneity in voxel values), OA (Overall appearance), and HS (heterogeneity in shapes) of each nodule for each view is extracted. OA is extracted from the square-shaped region of interest. HVV is considered by making non-nodule regions of the OA patch set to 0. HS is extracted by putting nodule regions of the OA patch as 0. In a pre-trained ResNet-50 model, the last fully-connected layer is replaced by three fully-connected layers to classify each view as benign or malignant. Further, each of these views' classification is combined to give the final output. However, the nine mentioned views, can extract only some information from a 3D to 2D. Multi-view knowledge-based collaborative (MV-KBC) is a supervised model.

The authors presented a framework based on the information extracted from texture, shape, and deep model (improved 2D eight-layer DCNN LeNet-5 model) [79]. The weighted sum of likelihood is used to converge the decisions made from the three classifiers. 697 uncertain nodules were either labelled as benign or labelled as malignant or discarded. If these are taken as benign then Acc 87.74%, Sen 81.11%, SP 89.67% and AUC 94.45%. If grouped as malignant then Acc 71.93%, Sen 59.22%, SP 84.85% and AUC 81.24%. The performance of the model of the discarded group is given in Table 6. In [80] authors designed a robust AgileCNN, which combines the benefits of LeNet layer settings and AlexNet parameter settings. Since LeNet and AlexNet are originally designed for coloured images, this affects the performance because the medical images are gray coloured images. Another similar model [81] used VGG-16, that was pre-trained on IMAGENET dataset. Authors compared their model to the conventional SVM based model. The validation accuracy of the model with transfer learning is reported as 68.0%, and without transfer learning is 58.9%. The performance of CADx is somewhat affected by the image size. The larger image size can give better accuracy.

### 6.2.2 3D Classification Techniques

The difficulty associated with classification of malignant lung nodules is the presence of In [82], authors developed LungNet, a shallow convolutional neural network. It has two CNNs, the survival prediction and the malignancy prediction network. Each has three convolutional, one max-pooling, and three fully connected layers. The weights of the survival prediction network are shared by transfer learning with the malignancy prediction network. The patients' clinical variables like age, histology, sex, and cancer stage, along with CT images, are considered. They demonstrated the use of transfer learning. The model achieves AUC of 0.85, while without transfer learning AUC of 0.82 is achieved.

Furthermore, a systematic end-to-end pipeline is proposed [41]. First predicts malignancy at nodule level and then predicts lung cancer at the patient level. They used 3D CNN ResNet50, with and without batch normalization, respectively, as classifiers.

Another method uses a regularization technique that helps in preserving the information as much as possible from the original data space [83]. The encoder block encodes the input image, while the decoder network reconstructs the encoded image. Further, classification task is carried out by the four fully connected layers using the encoded image. The network optimization loss function contains both the manifold loss and the classification loss. It is challenging to train the models because much inter-class similarity and intra-class variations exist in the medical images. Moreover a study explored the internal relationship behind a benign, malignant and nodule attributes classification in an end-to-end manner [84]. A fine-tune ResNet-22 avoids the vanishing gradient problem as features from the two-stream CNNs are stacked. A new hyperparameter in their loss function adjusts the relationship between multi-task classification. Although they manually adjust the weight combination of multi-task loss function.

Another method used a cloud-based 3D DCNN [57]. Here, Faster-RCNN (Faster-Regional convolutional

neural network), mRPN (multi-view and multi-scale region proposal network), a VGG-16 model is used. Authors integrate cloud-computing into the CAD system using 12 virtual machines and 24 processing units. Because of the cloud backend, radiologists can simultaneously get feedback from multiple radiologists in real-time. However, virtual environments may sometimes lead to inaccuracy. It incurs high performance and low computational cost. There is the use of MIP images for multi-view extraction. The performance was less accurate in detecting micro-nodules, whose diameter is less than 3 mm.

A high-level feature enhancement scheme is developed in which U-Net type architecture with three down sampling and three up sampling layers are applied[85]. By combining one dimensional output from global MaxPooling (GMP) layer with feature-analysis technique (soft activation mapping, SAM) they extracted high level features named HESAM. However, the model is dependent on the high-level accuracy from the autoencoder part.

There are a few hybrid models proposed for lung cancer diagnosis. Authors [86] discussed the fusion of both the handcrafted and CNN based features. Texture, intensity, and geometric based 29 features is selected. Radial basis kernel function (RBF) with support vector machine is used to select the significant features. The paper [45], used local binary patterns to extract texture features, a histogram of oriented gradients to extract shape features, and a 3D deep dual-path network (DPN) to extract in-depth features. The proposed DPN has benefits of both ResNet and DenseNet. Then a Gradient boosting machine algorithm is used to classify the nodules. It is not fully-automated but a hybrid model.

A systematic three-step hierarchical semantic convolutional neural network (HSCNN) is introduced [53]. The first step, model learns the general features using two 3D convolutional modules. In the next step predicts the score values for five attributes: texture, margin, sphericity, subtlety, and calcification using two fully connected, two batch normalization and one dropout layers. The third step, the outputs from the first step and the first fully connected layers of the second step are input to the module. However, only five malignancy tasks are there. Model achieves low accuracy (0.842) compared to multi-crop CNN (0.8714).

Using further the contextual and spatial attention mechanisms to denote the significance of local properties for nodule classification [87]. The ensemble of four variants of the CNN framework with weighted voted averaging, the final classification results are predicted. The result shows an F1 score of 90.45%.

Above mentioned research work detects the nodules present in the CT images. Nevertheless, the mere presence of nodules does not confirm that patient has cancer. Authors [55] provide a modified U-Net architecture which selects the five most malicious nodules, and then, using a leaky noisy-OR gate, and combines the cancer probabilities to obtain the probability of lung cancer for the subject.

A different strategy developed in [88] uses a trace norm regularization and graph structure matrix for regularization in supervised learning approach. A 3D CNN is used that is pre-trained on sports dataset. It is multi-task learning on visual attributes. Moreover a better scoring function based on the mean and standard deviation of the scores is proposed to deal with radiologists' uncertainty. However the average accuracy of the proposed model is low 78.06%. Further in the unsupervised learning approach, initial clustering is done on the appearance features to obtain an initial label. Then, using these labels, compute label proportions corresponding to each cluster. K-means clustering and Proportion-SVM are used for the above. The unsupervised learning methods like GANs and autoencoders are promising.

A systematic approach [89] for lung cancer diagnosis in which first the detection of candidate nodules, and then classification according to candidates' malignancy. Here, two CNN networks take  $47 \times 47 \times 5$  and  $21 \times 21 \times 5$  as input. A random forest classifier classifies the images based on fifty radiological quantitative image features (QIF). Nodule classification was performed separately on the two CNNs with and without QIF features. The best performance combined with QIF is in Table 6. If it would consider attributes malignancy also then model performance may have increased. Another qualitative approach adopted in [63] applying a multi-attribute lung nodule classification network (ALNC) using DenseNet. It is designed to output labels for each attribute and also classifies it based on malignancy.

In the paper [21], authors used a U-Net encoder-decoder architecture, leaky-noisy OR gate and a 20-layer deep residual CNN network. The occlusion and saliency maps are used to understand the internal process of

| Ref ID | Datasets            | Sample size  | Methods                            | Performance (%)             | Nodule category                         |
|--------|---------------------|--------------|------------------------------------|-----------------------------|-----------------------------------------|
| [91]   | clinical dataset    | 126 patients | 3D CapsNet                         | Acc 81.3, Sen 82.2, SP 80.7 | adenocarcinoma, squamous cell carcinoma |
| [90]   | clinical dataset    | 1419 samples | 10 ML algorithms, VGG16, 2D        | Acc 84.1, Sen 91.8, SP 80.9 | adenocarcinoma, squamous cell carcinoma |
| [92]   | clinical dataset    | 111 patients | 2D RDL                             | Acc 91.3, Sen 73.4, SP 96.4 | patterns of Adenocarcinoma              |
| [93]   | NSCLC Radiogenomics | 211 patients | 2D ensemble CNN, bidirectional RNN | Acc 96.29, Sen 96.09        | adenocarcinoma, squamous cell carcinoma |

Table 7: Correlative studies of histologic findings with CT findings

malignancy estimation. Using heat maps, they found that heat inside a nodule is generally a malignant feature. In contrast, peripheral heat, heat in the adjacent plane, and satellite heat were useful to classify nodule as benign. The model has a 22.2% false-positive rate for fatal cases and an 8% false-positive rate for mild cases. The probability of lung cancer in patients was predicted by taking the top 5 malignant nodules.

### 6.3 Bridging the Gap: Correlating CT Scans with Histological Analysis

Researchers developed various DNN models for studying correlation of histologic findings with CT findings. Authors [90] evaluated VGG16, ten machine learning (ML) algorithms and ten feature selection techniques for the differentiation of histological subtypes of NSCLC. The ML algorithms used are: Linear discriminant analysis, support vector machine, Random Forest classifier, AdaBoost, Naive Bayes, Bagging, multi-layer perceptron and decision tree. The experiments in this paper concluded that VGG16 has outperformed the conventional ML algorithms with radiomics.

A DNN model based on capsule net for the subtype discrimination of NSCLC is proposed [91]. Another method combined radiomics with deep learning (RDL) to predict patterns in adenocarcinoma nodules specifying ground glass opacifications [92]. Results showed radiomics combined with deep learning has better performance compared to radiomics alone. Furthermore a hybrid of convolutional and bidirectional recurrent neural network is introduced to detect NSCLC [93]. They ensembled and then averaged the best 10 models. The results of ensembled model were better than the single model. Table 7 shows the brief summary of the described papers.

## 7 Discussion and Future Directions

A total of 52 articles that are published in well-known journals, are focused in this review. This review work focuses on methodology for training and validation. Deep learning's ability to learn from data has revolutionized diagnostics, achieving unmatched accuracy over traditional methods. While integrating DNN technology into medical practice offers exciting possibilities, it also presents significant challenges. We will explore both aspects below.

### 7.1 Eliminating Bias: Ensuring Fair and Accurate Data

From the analysis, we see that most of the studies (a total of 38 selected articles) have used LIDC-IDRI and LUNA16 datasets, while some have used clinical based private datasets, which are not available publicly and thus restricts the research process [48, 58, 68, 41, 55, 57, 82, 21, 71, 88, 95].

The performance of models that are trained and tested on only LIDC-IDRI dataset: Lei2020 has highest accuracy of 99.13%, sensitivity of 97.05% and specificity of 99.21%. Here the formed four categories of dataset. These datasets differed in, what kind of information each contained. Like one category ignored the edge variations while the other contains shape variations and multi-scale information. But here it should be noted that uncertain nodes were removed from the training and testing set. One of the possible reasons for the low sensitivity of work published [85] is that all the nodules with average malignancy score of 3 are considered to be benign. [74] shows good performance in the detection of nodules. [84, 47, 54, 49] have low performance on all the metrics.

The datasets have to be appropriately annotated by the well-experienced radiologists. If it is clinical data, we see that very few radiologists annotate it, and in some data collected, the researchers themselves annotate these data and prepare the ground truth themselves. The researchers compared to radiologists, have less experience regarding the CT scans. So, their segmentation of nodules is prone to errors.

Though a substantial amount of data is available in these datasets, there is still a need for the variability in nodules shape, structure, and dimensions. Some of the studies use GAN, augmentation-based techniques to bring balance in the datasets. Studies validated with less image data are [94, 48, 77, 21]. For better generalization, the models must be trained and tested on different datasets available from various geographical locations. Better algorithms are still needed to deal with missing, unlabelled, ambiguous IDs, and noisy data [96].

## 7.2 Role of 2D and 3D Imaging in Early Detection

Owing to 3D nature of CT scans, 3D models are more appropriate than their 2D counterparts. In 3D model the inter channel correlation is exploited to refine the extraction of spectral data while in a 2D model only spatial correlation of the channels in the image are considered to obtain spectral data. In addition to volume, 3D medical imaging provides a clearer picture of blood vessels and crisper images of bones. 3D models can efficiently extract spatial information due to 3D-based convolution and pooling functions. More remarkably, training a 3D model from scratch may not certainly give better performance than transfer learning, fine-tuned from natural image or video to medical image in 2D.

Figure 6 illustrates the comparative view in terms of accuracy and sensitivity of the methodologies used in 2D based models and 3D based models. Only those methodologies are represented in the graph whose both accuracy and sensitivity are provided in the respective papers. From Figure 6 2D models, it is visible that the best accuracy of 98% is achieved by two methodologies, that is RNN GSO (Recurrent neural network glow-worm swarm optimization) [8] and hybrid ensemble CNN [44]. But hybrid ensemble CNN has less sensitivity of 93.5% while RNN GSO has sensitivity of 96%. So, we can conclude that RNN GSO achieves the better accuracy and sensitivity. However, RNN GSO is trained on small lung cancer dataset having 32 records. While ensemble CNN is trained on LIDC-IDRI dataset.

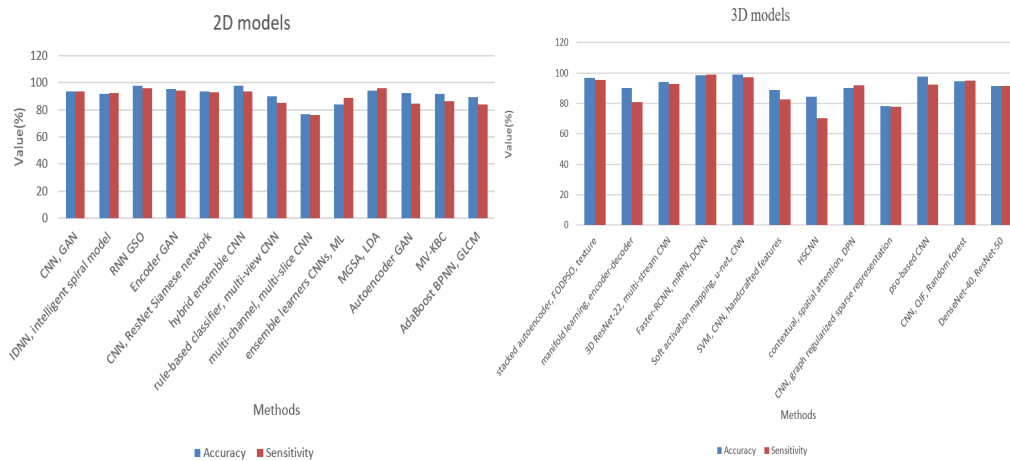


Figure 6: Accuracy and sensitivity of the reviewed methods

From Figure 6 3D models, it is visible that the methodologies Faster-RCNN, mRPN, DCNN [57] and Soft activation mapping (SAM), U-Net, CNN [85] achieves better accuracy and sensitivity compared to other methodologies. However, FP/scan is 2.1 which is high. Further, MIP images are used but the good thing is that it is tested on three different datasets. The methodologies that have comparatively performed poor are CNN, graph regularized sparse representation [88] and HSCNN [53]. The HSCNN has low sensitivity of 70.5% and CNN,

graph regularized sparse representation has low accuracy of 78.06%. The objective of this paper [88] is to show the efficiency of supervised and unsupervised techniques.

Although it is difficult to say which methodology is best among all, but still from Figure 6 it can be said that optimization techniques, hybrid ensemble techniques, evaluation based on multiple regions and SAM are the ones which have shown good results.

### 7.3 Integration of Supervised and Unsupervised Learning

During the study, we found that few work uses unsupervised techniques. Some have utilized the benefits of both unsupervised and supervised techniques [64]. Though unsupervised models are computationally complex, but for medical images where annotation of data is less available and background of the images are highly similar then, both the supervised and unsupervised techniques in the model might help improve sensitivity. Instead of discarding nodules with average malignancy label '3' and thus losing some amount of information, the features in this can be learnt from unsupervised techniques. There is many work [18, 58] that are supervised techniques based. Work [47, 64, 65, 88] have both supervised and unsupervised methods.

### 7.4 Advantages and Limitations of DNN

Table 5 and Table 6 provides a concise comparative analysis of the techniques reviewed in this work. The highest sensitivity and accuracy achieved by a 2D segmentation task model is 99.43% and 99.66% [4]. While for cancer classification task [59] has highest accuracy of 96.67%. Most of the reviewed 3D models can achieve more than 90% accuracy. Among 3D models, the work [85] achieves the highest accuracy of 99.13%, the sensitivity of 97.05%, and specificity of 99.21%. The work in [46, 85, 57, 4, 8, 77] has sensitivities higher than 95%, which means that these models can better detect malignant nodules. Studies [63, 74, 49, 58, 4] having Acc more than 90% [54, 47, 59, 8, 64, 62, 77, 43, 78, 46, 85, 57, 83, 84, 87, 45, 89].

Over the years the accuracy of the models has increased. The low sensitivity of 69.56% is because of different dataset used. The method is based on both deep and external features. The highest AUC achieved is 98.4%. the work shows better detection of nodules. The same methodology achieves AUC of 94.3% when deals with malignancy classification tasks. Minimum AUC is 66%, the aim of this paper has been to deal with small pathological datasets. Adopting ensemble techniques for classification tasks such as a four-channel CNN has been proposed, which shows better performance than a single classifier. Machine learning-based classifiers are there in proposed hybrid schemes, but instead of using handcrafted features, the features extracted from multiple CNN models are the input.

GANs are used to generate augmented 2D images [47, 54, 64, 59]; further, it can thus, produce 3D volumes of images using these 2D continuous images. Following studies [64, 46, 83] have used autoencoders. The highest accuracy achieved in the malignancy classification tasks is 94.6% while in detection tasks is 98.51%. Further researchers may consider these techniques in order to achieve the next level of progress in lung cancer diagnosis based on CT images.

### 7.5 Challenges Associated with Evaluation of Nodule Attributes

Few studies [49, 41, 42, 53, 88, 74, 63] have considered the radiologists annotated attribute scores as an essential criterion for classification. Each attribute focus on specific details of a nodule. The qualitative semantic information of a nodule like "highly spiculated", "lobulation index", "calcification index", "moderate heterogeneity" is captured. Some architectures predict the malignancy scores for each attribute and thus helps in improving the malignancy classification tasks. The benefits of such design are that gradients from semantic label prediction can also improve the detection and classification task. The purpose behind this is to improve the accuracy and consistency of medical image diagnosis through computational support. The work [85] rank malignancy attributes and choose only best ones to train their model. However, DNNs hold immense promise for the future of lung cancer detection and diagnosis.

## 8 Conclusion

This review analyzed 52 recent research papers on Deep Neural Network (DNN) methods for lung cancer diagnosis on chest CT scans. The studies explored nodule segmentation, detection, and malignancy classification. Key details like CT imaging characteristics, datasets, data pre-processing, and augmentation techniques were examined. The review categorized the methods based on tasks (nodule detection vs. malignancy classification) and input dimensions (2D or 3D). Reported accuracies were high, with nodule detection reaching 98% [8] and malignancy classification reaching 99.13% [85]. Advantages and limitations of different approaches were also discussed.

While significant progress has been made in automated lung cancer diagnosis using DNNs, particularly CNNs and their variants, there's room for improvement. The review suggests exploring emerging machine learning algorithms like reinforcement learning. These algorithms can incorporate cognitive decision-making alongside learned knowledge, potentially leading to the next level of computer-aided lung cancer diagnosis. The reviewed methods provide valuable insights for researchers aiming to improve early lung cancer detection.

## References

- [1] Cui, L., H. Li, W. Hui, S. Chen, L. Yang, Y. Kang, Q. Bo, J. Feng, "A deep learning-based framework for lung cancer survival analysis with biomarker interpretation", *BMC bioinformatics* 21(1), 1, 2020. <https://doi.org/10.1186/s12859-020-3431-z>
- [2] "Lung Cancer", <http://cancerindia.org.in/lung-cancer/>, 2020 (Accessed: 2020-09-10)
- [3] Jitendra, A. <http://www.univarta.com.html>, (Accessed: 2020-09-10).
- [4] Chen Y, Wang Y, Hu F, Wang D. "A lung dense deep convolution neural network for robust lung parenchyma segmentation", *IEEE Access*, 8:93527-47, 2020. <https://doi.org/10.1109/access.2020.2993953>
- [5] Khan, M.A., S. Rubab, A. Kashif, M.I. Sharif, N. Muhammad, J.H. Shah et al., "Lungs cancer classification from ct images: An integrated design of contrast based classical features fusion and selection", *Pattern Recognition Letters* 129 (77), 2020. <https://doi.org/10.1016/j.patrec.2019.11.014>
- [6] Dey, S.. <https://timesofindia.indiatimes.com/india/Cancer-cases-in-India-likely-to-soar-25-by-2020-ICMR/articleshow/52334632.cms>.
- [7] Chen, S. and S. Wu, "Identifying lung cancer risk factors in the elderly using deep neural networks: Quantitative analysis of web-based survey data", *Journal of Medical Internet Research* 22(e17695), 2020. <https://doi.org/10.2196/17695>
- [8] Selvanambi, R., J. Natarajan, M. Karuppiah, S.H. Islam, M.M. Hassan and G. Fortino, "Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization", *Neural Computing and Applications* 32(4373), 2020. <https://doi.org/10.1007/s00521-018-3824-3>
- [9] Moitra, D. and R.K. Mandal, "Classification of non-small cell lung cancer using one-dimensional convolutional neural network", *Expert Systems with Applications*, 113564, 2020. <https://doi.org/10.1016/j.eswa.2020.113564>
- [10] Lai, Y. H., W.-N. Chen, T.-C. Hsu, C. Lin, Y. Tsao and S. Wu, "Overall survival prediction of non-small cell lung cancer by integrating microarray and clinical data with deep learning", *Scientific reports* 10 (1), 2020. <https://doi.org/10.1038/s41598-020-61588-w>

- [11] Lee, B., S.H. Chun, J.H. Hong, I.S. Woo, S. Kim, J.W. Jeong et al., "DeepBts: prediction of recurrence-free survival of non-small cell lung cancer using a time-binned deep neural network", *Scientific reports* 10(1), 2020. <https://doi.org/10.1038/s41598-020-58722-z>
- [12] Armato, S., A. Starkey, H. MacMahon, G. McLennan, D. Yankelevitz, M. McNitt-Gray et al., "Toward a definition of lung nodule: a visual nodule library resource", *Radiology* 233(708), 2004. <https://dctd.cancer.gov/programs/cip>
- [13] Sim, Y., M.J. Chung, E. Kotter, S. Yune, M. Kim, S. Do et al., "Deep convolutional neural network-based software improves radiologist detection of malignant lung nodules on chest radiographs", *Radiology* 294(199), 2020. <https://doi.org/10.1148/radiol.2019182465>
- [14] Shaffer, K., "Deep learning and lung cancer: Ai to extract information hidden in routine ct scans", 2020. <https://doi.org/10.1148/radiol.2020201366>
- [15] Bhandary, A., G.A. Prabhu, V. Rajinikanth, K.P. Thanaraj, S.C. Satapathy, D.E. Robbins et al., "Deep-learning framework to detect lung abnormality—a study with chest x-ray and lung ct scan images", *Pattern Recognition Letters* 129 (271), 2020. <https://doi.org/10.1016/j.patrec.2019.11.013>
- [16] Zhang, G., S. Jiang, Z. Yang, L. Gong, X. Ma, Z. Zhou et al., "Automatic nodule detection for lung cancer in ct images: A review", *Computers in biology and medicine* 103 (287), 2018. <https://doi.org/10.1016/j.compbiomed.2018.10.033>
- [17] Ather, S., T. Kadir and F. Gleeson, "Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications", *Clinical radiology* 75 (13), 2020. <https://doi.org/10.1016/j.crad.2019.04.017>
- [18] Cao, H., H. Liu, E. Song, G. Ma, R. Jin, X. Xu et al., "A two-stage convolutional neural networks for lung nodule detection", *IEEE Journal of Biomedical and Health Informatics* P1, 2020. <https://doi.org/10.1109/jbhi.2019.2963720>
- [19] Lessmann, N., B. van Ginneken, M. Zreik, P.A. de Jong, B.D. de Vos, M.A. Viergever et al., "Automatic calcium scoring in low-dose chest ct using deep neural networks with dilated convolutions", *IEEE transactions on medical imaging* 37(615), 2017. <https://doi.org/10.1109/tmi.2017.2769839>
- [20] Kim, B. C., Y.S. Sung and H.-I. Suk, "Deep feature learning for pulmonary nodule classification in a lung ct", in *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, pp. 1–3, IEEE, 2016. <https://doi.org/10.1109/iww-bci.2016.7457462>
- [21] Venugopal, V.K., K. Vaidhya, M. Murugavel, A. Chunduru, V. Mahajan, S. Vaidya et al., "Unboxing ai-radiological insights into a deep neural network for lung nodule characterization", *Academic Radiology* 27 (88), 2020. <https://doi.org/10.1016/j.acra.2019.09.015>
- [22] Zhang, C., X. Sun, K. Dang, K. Li, X. Guo, J. Chang et al., "Toward an expert level of lung cancer detection and classification using a deep convolutional neural network", *The Oncologist* 24(1159), 2019. <https://doi.org/10.1634/theoncologist.2018-0908>
- [23] Zhang J, Xia Y, Cui H, Zhang Y. "Pulmonary nodule detection in medical images: A survey", *Biomedical Signal Processing and Control*, 43:138-47, 2018. <https://doi.org/10.1016/j.bspc.2018.01.011>
- [24] Chaturvedi P, Jhamb A, Vanani M, Nemade V. "Prediction and classification of lung cancer using machine learning techniques", In *IOP Conference Series: Materials Science and Engineering*, 1099(1),012059, 2021. <https://doi.org/10.1088/1757-899x/1099/1/012059>

- [25] Abdullah DM, Ahmed NS. "A review of most recent lung cancer detection techniques using machine learning", *International Journal of Science and Business*, 5(3):159-73, 2021. <https://ideas.repec.org/a/aif/journal/v5y2021i3p159-173.html>
- [26] Binczyk F, Prazuch W, Bozek P, Polanska J. "Radiomics and artificial intelligence in lung cancer screening", *Translational lung cancer research*, 10(2):1186, 2021. <https://doi.org/10.21037/tlcr-20-708>
- [27] Monkam, P., S. Qi, H. Ma, W. Gao, Y. Yao and W. Qian, "Detection and classification of pulmonary nodules using convolutional neural networks: A survey", *IEEE Access* 7 (78075), 2019. <https://doi.org/10.1109/access.2019.2920980>
- [28] Sathyakumar K, Munoz M, Singh J, Hussain N, Babu BA. "Automated lung cancer detection using artificial intelligence (AI) deep convolutional neural networks: A narrative literature review", *Cureus*, 12(8), 2020. <https://doi.org/10.7759/cureus.10017>
- [29] Zhang, G., S. Jiang, Z. Yang, L. Gong, X. Ma, Z. Zhou et al., "Automatic nodule detection for lung cancer in ct images: A review", *Computers in biology and medicine* 103 (2018): 287-300. 12, 2018. <https://doi.org/10.1016/j.compbiomed.2018.10.033>
- [30] Ma, J., Y. Song, X. Tian, Y. Hua, R. Zhang and J. Wu, "Survey on deep learning for pulmonary medical imaging", *Frontiers of Medicine* 1, 2019. <https://doi.org/10.1007/s11684-019-0726-4>
- [31] Liu X, Li KW, Yang R, Geng LS. "Review of deep learning based automatic segmentation for lung cancer radiotherapy", *Frontiers in Oncology*, 11:2599, 2021. <https://doi.org/10.3389/fonc.2021.717039>
- [32] Riquelme D, Akhloufi MA. "Deep learning for lung cancer nodules detection and classification in CT scans", *AI* 1(1):28-67, 2020. <https://doi.org/10.3390/ai1010003>
- [33] Ahmed T, Parvin MS, Haque MR, Uddin MS. "Lung cancer detection using CT image based on 3D convolutional neural network", *Journal of Computer and Communications* 2020 Mar 5;8(03):35. <https://doi.org/10.4236/jcc.2020.83004>
- [34] "Types of Lung Cancer", <https://webmd.com/>, 2020 (Accessed: 2020-09-10).
- [35] Gerard, S.E., J. Herrmann, D.W. Kaczka, G. Musch, A. Fernandez-bustamante and J.M. Reinhardt, "Multi-resolution convolutional neural networks for fully automated segmentation of acutely injured lungs in multiple species", *Medical Image Analysis* 60(101592), 2020. <https://doi.org/10.1016/j.media.2019.101592>
- [36] Nafti, N. and M. Othmani, "A novel approach of lung tumor detection using deep neural network", in *2019 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)* pp. 198–202, IEEE, 2019. <https://doi.org/10.1109/iintec48298.2019.9112100>
- [37] Orobinskyi, P., D. Petrenko and V. Lyashenko, "Novel approach to computer-aided detection of lung nodules of difficult location with use of multifactorial models and deep neural networks", in *2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM)*, pp. 1–5, IEEE, 2019. <https://doi.org/10.1109/cadsm.2019.8779340>
- [38] Huang, X., J. Shan and V. Vaidya, "Lung nodule detection in ct using 3d convolutional neural networks", in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 379–383, IEEE, 2017. <https://doi.org/10.1109/isbi.2017.7950542>
- [39] Essaf, F., Y. Li, S. Sakho and P.K. Gadosey, "Improved convolutional neural network for lung cancer detection", in *Proceedings of the 2020 International Conference on Computing, Networks and Internet of Things*, pp. 48–54, 2020. <https://doi.org/10.1145/3398329.3398337>

- [40] Hussein, S., R. Gillies, K. Cao, Q. Song and U. Bagci, "Tumornet: Lung nodule characterization using multi-view convolutional neural network with gaussian process", in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 1007–1010, IEEE, 2017. <https://doi.org/10.1109/isbi.2017.7950686>
- [41] Bonavita, I., X. Rafael-Palou, M. Ceresa, G. Piella, V. Ribas and M.A.G. Ballester, "Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline", *Computer Methods and Programs in Biomedicine* 185(105172), 2020. <https://doi.org/10.1016/j.cmpb.2019.105172>
- [42] Wang, J., X. Chen, H. Lu, L. Zhang, J. Pan, Y. Bao et al., "Feature-shared adaptive-boost deep learning for invasiveness classification of pulmonary subsolid nodules in CT images", *Medical Physics* 47(1738), 2020. <https://doi.org/10.1002/mp.14068>
- [43] Huang, X., W. Sun, T.-L.B. Tseng, C. Li and W. Qian, "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(25), 2019. <https://doi.org/10.1016/j.compmedimag.2019.02.003>
- [44] Omigbodun, A.O., F. Noo, M. McNitt-Gray, W. Hsu and S.S. Hsieh, "The effects of physics-based data augmentation on the generalizability of deep neural networks: Demonstration on nodule false-positive reduction", *Medical Physics* 46(4563), 2019. <https://doi.org/10.1002/mp.13755>
- [45] Zhang, G., Z. Yang, L. Gong, S. Jiang and L. Wang, "Classification of benign and malignant lung nodules from ct images based on hybrid features", *Physics in Medicine & Biology* 64(125011), 2019. <https://doi.org/10.1088/1361-6560/ab2544>
- [46] Naqi, S.M., M. Sharif and A. Jaffar, "Lung nodule detection and classification based on geometric fit in parametric form and deep learning", *Neural Computing and Applications* 32(4629), 2020. <https://doi.org/10.1007/s00521-018-3773-x>
- [47] Kuang, Y., T. Lan, X. Peng, G.E. Selasi, Q. Liu and J. Zhang, "Unsupervised multi-discriminator generative adversarial network for lung nodule malignancy classification", *IEEE Access* 8(77725), 2020. <https://doi.org/10.1109/access.2020.2987961>
- [48] Zhang, S., F. Han, Z. Liang, J. Tan, W. Cao, Y. Gao et al., "An investigation of cnn models for differentiating malignant from benign lesions using small pathologically proven datasets", *Computerized Medical Imaging and Graphics* 77, 2019. <https://doi.org/10.1016/j.compmedimag.2019.101645>
- [49] Liu, L., Q. Dou, H. Chen, J. Qin and P.A. Heng, "Multi-task deep model with margin ranking loss for lung nodule analysis", *IEEE Transactions on Medical Imaging* 39 (718), 2020. <https://doi.org/10.1109/tmi.2019.2934577>
- [50] Choe, J., S.M. Lee, K.-H. Do, G. Lee, J.-G. Lee, S.M. Lee et al., "Deep learning-based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses", *Radiology* 292(365), 2019. <https://doi.org/10.1148/radiol.2019181960>
- [51] Souza, J.C., J.O.B. Diniz, J.L. Ferreira, G.L.F. da Silva, A.C. Silva and A.C. de Paiva, "An automatic method for lung segmentation and reconstruction in chest x-ray using deep neural networks", *Computer methods and programs in biomedicine* 177(285), 2019. <https://doi.org/10.1016/j.cmpb.2019.06.005>
- [52] Kumar, D., A. Wong and D.A. Clausi, "Lung nodule classification using deep features in CT images", in *2015 12th Conference on Computer and Robot Vision*, pp. 133–138, IEEE, 2015. <https://doi.org/10.1109/crv.2015.25>

- [53] Shen, S., S.X. Han, D.R. Aberle, A.A. Bui and W. Hsu, "An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification", *Expert systems with applications* 128 (84), 2019. <https://doi.org/10.1016/j.eswa.2019.01.048>
- [54] Suresh, S. and S. Mohan, "ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis", *Neural Computing and Applications*, 2020. <https://doi.org/10.1007/s00521-020-04787-w>
- [55] Liao, F., M. Liang, Z. Li, X. Hu and S. Song, "Evaluate the malignancy of pulmonary nodules using the 3-D deep leaky noisy-or network", *IEEE Transactions on Neural Networks and Learning Systems* 30 (3484), 2019. <https://doi.org/10.1109/tnnls.2019.2892409>
- [56] Jiang, H., H. Ma, W. Qian, M. Gao and Y. Li, "An automatic detection system of lung nodule based on multigroup patch-based deep learning network", *IEEE Journal of Biomedical and Health Informatics* 22 (1227), 2018. <https://doi.org/10.1109/jbhi.2017.2725903>
- [57] Masood, A., P. Yang, B. Sheng, H. Li, P. Li, J. Qin et al., "Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT", *IEEE Journal of Translational Engineering in Health and Medicine* 8, 2020. <https://doi.org/10.1109/jtehm.2019.2955458>
- [58] Shakeel, P.M., M. Burhanuddin and M.I. Desa, "Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier", *Neural Computing and Applications* 1, 2020. <https://doi.org/10.1007/s00521-020-04842-6>
- [59] Yu, H., Z. Zhou and Q. Wang, "Deep learning assisted predict of lung cancer on computed tomography images using the adaptive hierarchical heuristic mathematical model", *IEEE Access* 8 (86400), 2020. <https://doi.org/10.1109/access.2020.2992645>
- [60] Hu, Q., L.F. Lu 'is, G.B. Holanda, S.S. Alves, F.H. Francisco, T. Han et al., "An effective approach for CT lung segmentation using mask region-based convolutional neural networks", *Artificial Intelligence in Medicine* 103, 2020. <https://doi.org/10.1016/j.artmed.2020.101792>
- [61] Cao, H., H. Liu, E. Song, C.C. Hung, G. Ma, X. Xu et al., "Dual-branch residual network for lung nodule segmentation", *Applied Soft Computing Journal* 86, 2020. <https://doi.org/10.1016/j.asoc.2019.105934>
- [62] Kasinathan, G., S. Jayakumar, A.H. Gandomi, M. Ramachandran, S.J. Fong and R. Patan, "Automated 3-d lung tumor detection and classification by an active contour model and cnn classifier", *Expert Systems with Applications* 134 (112), 2019. <https://doi.org/10.1016/j.eswa.2019.05.041>
- [63] Dai, Y., S. Yan, B. Zheng and C. Song, "Incorporating automatically learned pulmonary nodule attributes into a convolutional neural network to improve accuracy of benign-malignant nodule classification", *Physics in Medicine and Biology* 63, 2018. <https://doi.org/10.1088/1361-6560/aaf09f>
- [64] Xie, Y., J. Zhang and Y. Xia, "Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT", *Medical Image Analysis* 57 (237), 2019. <https://doi.org/10.1016/j.media.2019.07.004>
- [65] Gu, Y., X. Lu, L. Yang, B. Zhang, D. Yu, Y. Zhao et al., "Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs", *Computers in biology and medicine* 103 (220), 2018. <https://doi.org/10.1016/j.combiomed.2018.10.011>
- [66] El-Regaily, S.A., M.A.M. Salem, M.H.A. Aziz and M.I. Roushdy, "Multi-view convolutional neural network for lung nodule false positive reduction", *Expert systems with applications* 113017, 2019. <https://doi.org/10.1016/j.eswa.2019.113017>

- [67] Xie, H., D. Yang, N. Sun, Z. Chen and Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks", *Pattern Recognition* 85 (109), 2019. <https://doi.org/10.1016/j.patcog.2018.07.031>
- [68] Li, G., W. Zhou, W. Chen, F. Sun, Y. Fu, F. Gong et al., "Study on the detection of pulmonary nodules in CT images based on deep learning", *IEEE Access* 8 (67300), 2020. <https://doi.org/10.1109/access.2020.2984381>
- [69] Zheng, S., J. Guo, X. Cui, R.N.J. Veldhuis, M. Oudkerk and P.M.A. van Ooijen, "Automatic pulmonary nodule detection in ct scans using convolutional neural networks based on maximum intensity projection", *IEEE Transactions on Medical Imaging* 39 (797), 2020. <https://doi.org/10.1109/tmi.2019.2935553>
- [70] Ozdemir, O., R.L. Russell and A.A. Berlin, "A 3d probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans", *IEEE Transactions on Medical Imaging* 39 (1419), 2020. <https://doi.org/10.1109/tmi.2019.2947595>
- [71] Huang, W., Y. Xue and Y. Wu, "A CAD system for pulmonary nodule prediction based on deep three-dimensional convolutional neural networks and ensemble learning", *Plos one* 14(e0219369), 2019. <https://doi.org/10.1371/journal.pone.0219369>
- [72] Pezeshk, A., S. Hamidian, N. Petrick and B. Sahiner, "3-d convolutional neural networks for automatic detection of pulmonary nodules in chest CT", *IEEE Journal of Biomedical and Health Informatics* 23(2080), 2019. <https://doi.org/10.1109/jbhi.2018.2879449>
- [73] Silva, G.L. da, O.P. da Silva Neto, A.C. Silva, A.C. de Paiva and M. Gattass, "Lung nodules diagnosis based on evolutionary convolutional neural network", *Multimedia Tools and Applications* 76 (19039), 2017. <https://doi.org/10.1007/s11042-017-4480-9>
- [74] Silva, G.L.F. da, T.L.A. Valente, A.C. Silva, A.C. de Paiva and M. Gattass, "Convolutional neural network-based pso for lung nodule false positive reduction on CT images", *Computer Methods and Programs in Biomedicine* 162 (109), 2018. <https://doi.org/10.1016/j.cmpb.2018.05.006>
- [75] Srivastava, V. and R.K. Purwar, "Classification of CT scan images of lungs using deep convolutional neural network with external shape-based features", *Journal of digital imaging* 33 (252), 2020. <https://doi.org/10.1007/s10278-019-00245-9>
- [76] Zhang, B., S. Qi, P. Monkam, C. Li, F. Yang, Y.-D. Yao et al., "Ensemble learners of multiple deep cnns for pulmonary nodules classification using ct images", *IEEE Access* 7(110358), 2019. <https://doi.org/10.1109/access.2019.2933670>
- [77] Lakshmanaprabu, S., S.N. Mohanty, K. Shankar, N. Arunkumar and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images", *Future Generation Computer Systems* 92(374), 2019. <https://doi.org/10.1016/j.future.2018.10.009>
- [78] Xie, Y., Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham et al., "Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT", *IEEE transactions on medical imaging* 38(991), 2018. <https://doi.org/10.1109/tmi.2018.2876510>
- [79] Xie, Y., J. Zhang, Y. Xia, M. Fulham and Y. Zhang, "Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT", *Information Fusion* 42(102), 2018. <https://doi.org/10.1016/j.inffus.2017.10.005>
- [80] Zhao, X., L. Liu, S. Qi, Y. Teng, J. Li and W. Qian, "Agile convolutional neural network for pulmonary nodule classification using ct images", *International journal of computer assisted radiology and surgery* 13(585), 2018. <https://doi.org/10.1007/s11548-017-1696-0>

- [81] Nishio, M., O. Sugiyama, M. Yakami, S. Ueno, T. Kubo, T. Kuroda et al., "Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning", *PloS one* 13 (e0200721), 2018. <https://doi.org/10.1371/journal.pone.0200721>
- [82] Mukherjee, P., M. Zhou, E. Lee, A. Schicht, Y. Balagurunathan, S. Napel et al., "A shallow convolutional neural network predicts prognosis of lung cancer patients in multi-institutional computed tomography image datasets", *Nature Machine Intelligence* 2 (274), 2020. <https://doi.org/10.1038/s42256-020-0173-6>
- [83] Ren, Y., M.-Y. Tsai, L. Chen, J. Wang, S. Li, Y. Liu et al., "A manifold learning regularization approach to enhance 3d ct image-based lung nodule classification", *International Journal of Computer Assisted Radiology and Surgery* 15(287), 2020. <https://doi.org/10.1007/s11548-019-02097-8>
- [84] Zhao, J., C. Zhang, D. Li and J. Niu, "Combining multi-scale feature fusion with multi-attribute grading, a cnn model for benign and malignant classification of pulmonary nodules", *Journal of Digital Imaging* 2020. <https://doi.org/10.1007/s10278-020-00333-1>
- [85] Lei, Y., Y. Tian, H. Shan, J. Zhang, G. Wang and M.K. Kalra, "Shape and margin-aware lung nodule classification in low-dose CT images via soft activation mapping", *Medical Image Analysis* 60, 2020. <https://doi.org/10.1016/j.media.2019.101628>
- [86] Li, X., B. Hu, H. Li and B. You, "Application of artificial intelligence in the diagnosis of multiple primary lung cancer", *Thoracic Cancer* 10 (2168), 2019. <https://doi.org/10.1111/1759-7714.13185>
- [87] Jiang, H., F. Gao, X. Xu, F. Huang and S. Zhu, "Attentive and ensemble 3d dual path networks for pulmonary nodules classification", *Neurocomputing* 2019. <https://doi.org/10.1016/j.neucom.2019.03.103>
- [88] Hussein, S., P. Kandel, C.W. Bolan, M.B. Wallace and U. Bagci, "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches", *IEEE transactions on medical imaging* 38,1777, 2019. <https://doi.org/10.1109/tmi.2019.2894349>
- [89] Causey, J.L., J. Zhang, S. Ma, B. Jiang, J.A. Qualls, D.G. Politte et al., "Highly accurate model for prediction of lung nodule malignancy with ct scans", *Scientific reports* 8(1), 2018. <https://doi.org/10.1038/s41598-018-27569-w>
- [90] Han Y, Ma Y, Wu Z, Zhang F, Zheng D, Liu X, Tao L, Liang Z, Yang Z, Li X, Huang J. "Histologic subtype classification of non-small cell lung cancer using PET/CT images", *European journal of nuclear medicine and molecular imaging* 48(2):350-60, 2021. <https://doi.org/10.1007/s00259-020-04771-5>
- [91] Liu H, Jiao Z, Han W, Jing B. "Identifying the histologic subtypes of non-small cell lung cancer with computed tomography imaging: A comparative study of capsule net, convolutional neural network, and radiomics", *Quantitative Imaging in Medicine and Surgery* ;11(6):2756, 2021. <https://doi.org/10.21037/qims-20-734>
- [92] Wang X, Zhang L, Yang X, Tang L, Zhao J, Chen G, Li X, Yan S, Li S, Yang Y, Kang Y. "Deep learning combined with radiomics may optimize the prediction in differentiating high-grade lung adenocarcinomas in ground glass opacity lesions on CT scans", *European journal of radiology*, 129:109150, 2020. <https://doi.org/10.1016/j.ejrad.2020.109150>
- [93] Moitra D, Mandal RK. "Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network", *Journal of Digital Imaging* 2020 Aug;33(4):895-902. <https://doi.org/10.1007/s10278-020-00337-x>

- [94] Chae, K.J., G.Y. Jin, S.B. Ko, Y. Wang, H. Zhang, E.J. Choi et al., "Deep learning for the classification of small (2 cm) pulmonary nodules on ct imaging: A preliminary study", *Academic Radiology* 27, 2020. <https://doi.org/10.1016/j.acra.2019.05.018>
- [95] Hosny, A., C. Parmar, T.P. Coroller, P. Grossmann, R. Zeleznik, A. Kumar et al., "Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study", *PLoS medicine* 15, 2018 e1002711. <https://doi.org/10.1371/journal.pmed.1002711>
- [96] Li, D., B.M. Vilmun, J.F. Carlsen, E. Albrecht-Beste, C.A. Lauridsen, M.B. Nielsen et al., "The performance of deep learning algorithms on automatic pulmonary nodule detection and classification tested on different datasets that are not derived from linc-idri: A systematic review", *Diagnostics* 12, 2019. <https://doi.org/10.3390/diagnostics9040207>