“Classification and Detection of Lung Cancer Nodule using Deep Learning of CT Scan Images”: A Systematic Review

Anas Abrar (aabrar15@gmail.com)
Aligarh Muslim University
Priyanka Rajpoot
Universiti Sains Malaysia

Research Article

Keywords: Lung cancer, Deep learning, Lung nodule, Computed tomography, Image processing

Posted Date: October 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2145172/v1

Abstract

Lung cancer is considered as the common cancerous neoplasms across the globe. In 2018, the World Health Organization (WHO) statistics approximated 2.09 million lung cancer cases with 1.76 million deaths globally. Early identification is an important aspect of providing the greatest chance of healing the patients. The objective of this manuscript was to explore how Deep Learning (DL) performs when the method is evaluated on datasets that are not from LUNA 16 for detection of pulmonary nodule and categorization of computed tomography scans. This report covered only peer-reviewed, original research papers using DL technology, and only findings were included from testing on datasets other than LUNA-16 and LIDC-IDRI. Deep learning utilizes Computed-Tomography (CT) to automatically improve the precision of an initial diagnosis of lung cancer. Consequently, this manuscript presents a short yet important review of DL methods to solve the extraordinary challenges of detecting lung cancer. In addition, this paper also traces the various causes, types, and treatment procedures of lung cancer. The fundamental principles of deep learning and CT have been described. A review of the various lung cancer detection methods via deep learning has been presented. Finally, discussions have been provided for further improvisation of the deep learning method. 9 studies investigated pulmonary nodule detection performance, 10 studies investigated the classification of pulmonary nodule performance, and 16 studies documented of pulmonary nodule for both classification and detection. Some of prominent DL methods which have been successful in detection and categorization of lung cancer nodules are Computer Aided Detection (CAD), Wavelet Recurrent Neural Network (WRNN), Optimal Deep Neural Network (ODNN), Massive Artificial Neural Network (MTANN) and Convolutional Neural Network (CNN) Training. Among, these DL methods, in most cases CNN achieved higher accurate results. The reports CNN achieved results between 73%-96.73% for both classification and detection. The CNN achieved results between 76%-99.2% for lung nodules classification and also achieved the results between 74.6%-97.78% for lung nodule detection. In addition to this, it was found that other DL method i.e., MTANN achieved the accurate results between 97%-100% for detection which came out to be superior related to other DL approaches.

1. Introduction

A tumor is classified as Non- Small Cell Lung Cancer and Small Cell Lung Cancer, according to the viewpoint of pathology, treatment, and histological classification [1]. Lung cancer reports for two-thirds of all cancers and is known as one of the most severe death-related diseases [2]. Recently, according to cancer figures from the National Institute of Health (NIH), 12.9 % of the overall diagnosis of cancer has involved lung cancer in the USA [3]. Several techniques have been reported previously to create images of human body parts for treatment and diagnostic purposes. These include Positron Emission Tomography (PET), CT, Mammography, Magnetic Resonance Imaging (MRI), X-ray, and Ultrasound [4]. CT and X-ray are the two most common forms of imaging utilized to identify lung nodules [5]. However, the traditional CT examination entails radiologist evaluation, which is particularly laborious and often provides false-positive test findings [6]. Additionally, the CT procedure is not sufficient as certain parts of the chest X-ray image of an infected lung might appear normal and ultimately lead to an inaccurate diagnosis [7]. And chest radiography is also the most common form of procedure for others to initially identify and diagnose lung cancer; due to its non-invasiveness features, radiation dose and economic considerations, it is even preferred to more sensitive and reliable procedures (e.g. MRI and CT). The primary phase in Computer-Aided pulmonary nodule detection systems is a patient recognition process considered to identify a basic demonstration of lung anatomy, so that characteristics such as the lung wall and large routes are omitted, leaving only information that has a superior potential to be a nodule [8].

In recent years, researchers are paying a higher amount of attention to an alternative approach of clinical-image-processing to analyze lung cancer called Convolutional Neural Networks (CNN). By automating the initial evaluation in medical scans, DL offers us the chance to improve the precision in early detection [9], [10]. For the detection of natural camera pictures, CNN has been used quite effectively and has recently been applied in CT analysis and classification [11]. Hence, as a research direction of deep learning, CNN has become a common tool in the field of classification of images [12]. The application of CNN helps us to take into account the architectural pattern and improve the unprocessed images acquired from multiple sources [13]. Furthermore, the aid of CNN is effective in retrieving essential data. DL has newly developed as an extra intelligent and correct knowledge for image category and has been applied to classify therapeutic pictures, including CTs. To the best of understanding, when diagnosing pulmonary nodules, DL method has yet to be applied effectively in a regular medical workflow. A reason for this could be that DL processes need to be trained on information comparable to final task awareness [14].

The aim of this manuscript was to consider how DL performs when the method is evaluated on datasets that are not from LUNA 16 for pulmonary nodule detection and classification of CT scans. In addition, the analysis aimed to investigate whether when the process is evaluated on a dataset that is dissimilar from the training dataset, the efficiency of DL is reduced. A brief description type of lung cancer is, and its types are given in the Section 2. Section 3 offers details about the fundamental work of CT and the datasets and also describes the works based on CNN architecture. Section 4 gives a detailed overview of how deep learning method was used by other researchers for the detection and classification of CT datasets pertaining to lung nodule. Section 5 discusses the outcome of the review based on the utilization of deep learning techniques to detect and classify CT datasets and provides the reader with the whole panorama of this area. It gives outstanding references to this subject, and furthermore, it describes the latest state-of-the-art approaches to resolve issues of the nodules of lung cancer. Section 6 presents the conclusion of this systematic review.

2. Background Study

Lung cancer develops in the lungs and can extent to lymph nodes or different areas in the body, for example the brain. Cancer can spread to the lungs from other organs as well. There are metastases, too, as cancer cells pass from one organ to another. A brief overview of lung cancer causes, and types are explained in the following sub-sections.

2.1 Causes of Lung Cancer

The potentials contributor to lung cancer by far is smoking, predominantly cigarettes [15]. Smoke of cigarette contains at least 73 known carcinogens, amongst other items. NNK,1,3-Benzo and polonium pyrene [alpha] butadiene, polonium-210-radioactive isotopes [16]. In year 2000, around 90% of lung cancer deaths in men were due to smoking in the developing world (70% for females), smoke compensates for 85% of lung cancer cases [17]. It is possible to describe a passive smoker as someone working or living with a smoker. As per European, US and UK research [17]. Radon gas is a colourless and odourless gas formed due to radioactive radium dissolution, which is the uranium deterioration agent contained in the Earth's crust. In USA, Radon is the 2nd record important reason of lung cancer that causing around 21,000 deaths per year, and radiation products of genetic ionize material that causes transmutations that often become tumorous. [17]. The toxicity of outdoor air has a minor impact on raising the possibility of lung cancer. The extremely elevated risk is associated with minute particulates (PM2.5) and sulfate aerosols that can be emitted in traffic exhaust fumes. A marginal increase in nitrogen dioxide of 10 parts per billion surges the possibility of lung cancer by 14% [17].

2.2 Types of Lung Cancer

Generally, cancer or disease arises when cells begin to develop out of control in the body. There are two basic types of Lung Cancer called Small-Cell-Lung-Cancer (SCLC) and Non-SCLC (NSCLC) [18]. A complete classification of the various sources of Lung Cancer is portrayed in Fig. 1. Globally, around 15% of the lung cancers are SCLC while the remaining 85% are NSCLC [19].

2.2.1 Small-Cell-Lung-Cancer (SCLC)

A type of high cancer that most frequently exists inside the lung is Small-Cell Carcinoma [20]. Most cases of duodenum neuroendocrine Small-Cell-Carcinoma show rapid development of cancer and death is not stopped except by radical surgery, with or without chemotherapy. We have also described a rare sub-type of Small-Cell Lung Cancer [21]. Has been linked to heterozygosity failure at many different genetic loci, including 3p, 13q, and 17p chromosomes [22]. SCLC is of two types as follow:

- **LS-SCLC (Limited Stage Small Cell Lung Cancer)**

A form of Small Cell Lung Cancer (SCLC) is restricted to a region that is lesser enough to be included within a radiation portal is Limited-Stage Small Cell Lung Carcinoma (LS-SCLC) [23].

- **ES-SCLC (Extensive Stage Small Cell Lung Cancer)**

Platinum-based combination chemotherapy is the standardized care in ES-SCLC, with radiotherapy applied mainly to palliate symptoms such as dyspnea, liver or bone pain, or for the cure of brain metastases that usually provide a rapid, if the transient response to whole-brain radiotherapy in SCLC[24]

2.2.2 Non-Small Cell-Lung-Cancer

When irregular cells easily replicate and do not avoid reproducing, cancer happens. In the body, the disease can evolve everywhere. Treatment is based on the position thereof [25]. Typically, NSCLC spreads slower than Small-Cell Lung Cancers [26,27]. The key NSCLC forms are Squamous Cell Carcinoma, Adenocarcinoma and Large Cell Carcinoma [28]. These subtypes are classified as NSCLC since their care and prognosis (outlook) are identical [29]. Surgical resection is useful in almost all peoples with stage I (SCLC) to stage II (NSCLC), while patients with more advanced diseases are candidates for treatment without surgery. Orthodox clinical staging is most generally conducted for thorax and upper abdominal CT scans [30]. If a patient has signs like excessive coughing or chest inflammation, breathlessness, hoarseness, chest pain, or blood coughing, lung cancer may be suspected. Fever, lack of appetite, and excessive weight loss may be other signs [31]. Figure 2 shows the example of NSLC. The right bottom lobe in the Figure shows the cell carcinoma.

- **Adenocarcinoma**

This is the form of NSCLC that is most general. This is the most prevalent cause of lung cancer without without-smoke. Much more commonly, it's found in Current-Smokers or Ex-Smokers The outer edges of the lungs begin to expand. Typically, it progresses more slowly than other lung cancer type [32]. It amounts to 40 to 55 of the lung with cancer, and in many nations, it has become the most prevalent form. The peripheral form of adenocarcinoma is more clinically normal. It is primarily split into 4 forms [1].

- **Squamous-Cell-Carcinoma**

This sort of NSCLC is more commonly seen in smokers or former smokers. These cancers appear to originate near the main airways (the bronchi) in the middle portion of the lungs [32]. In recent years, the occurrence of this sort has been decreasing. Lung cancer accounts for 30–40%, with 2/3 being primary and 1/3 being peripheral [1].

- **th**

This is the kind of NSCLC that is the least general. It continues to expand and spread to other organs rapidly. This will make treatment more complicated [32].

2.3 Treatment Procedures of Lung Cancer

SCLC and NSCLC follow different treatment procedures [29]. The various methods of treating lung cancer are described in Fig. 3.

Generally, lung cancer treatment seeks to eliminate the whole tumor at the boundary, leaving a small volume of natural tissue (almost 2 cm, 0.8 inches). The generic term for surgery that reaches the chest is thoracotomy, and as part of the thoracotomy, various named forms of surgical operations can be performed, such as wedge resection [32]. Combination radiotherapy and chemotherapy treatment significantly improves survival in patients with un-removable stage 3 lung cancer. Tumor Treatment Fields can enhance chemotherapy treatment. Different molecular-targeted drugs have been developed in recent years to treat advanced lung cancer [33]. One such drug targeting the Epidermal Growth Factor Receptor (EGFR) domain of tyrosine kinase is Gefitinib (Iressa; withdrawn from the U.S. market) [34], expressed in several instances of NSCLC carcinoma. Immunotherapy is a part of therapy for cancer that stimulates the cancer-fighting immune system [35]. Radiotherapy alone leads to 13–39 percent of patients lasting five years with respect to stage one or two NSLC [32]. A brief overview of lung cancer detection methods and tools are described in the following Sections 3 and also describes the databases, work based on CNN architecture.

3. Lung Cancer Detection Tool And Method

Different tools are used that is CT, MRI, and PET for Lung Tumor Recognition using DL [36]. CT-scans are more widely used than MRI and are typically less expensive. A brief overview of CT and DL is labelled in the following sub-sections.

3.1 Computed Tomography

Godfrey Newbold Hounsfield [37] Nobel Prize winner in 1979 for Physiology and Medicine) developed the 1st X-ray CT method in 1972. Future applications in the detection of medical imaging became obvious from the beginning. In 1971, a scan of a patient with a cystic frontal lobe disease at Atkinson Morley Hospital in Wimbledon (United Kingdom) was submitted into clinical practice by CT. After this, the medical profession immediately accepted CT and was frequently referred to as the most influential invention of radiological diagnosis since the discovery of X-rays [37]. A Free-breathing scan is also used to prepare radiotherapy for lung tumors [38]. This procedure is replicated many times in the body until the entire field of concern is filled. Then, through image processing techniques, multiple cross-sectional scans are collected along the body. After their inception, CT scans have been used because of their efficacy in helping the physician to diagnose numerous diseases [39]. A quantitative scale for defining radio-density on CT scans is the Hounsfield scale, named after Sir Godfrey Hounsfield [40]. A Hounsfield-Scale (HE) is used for the pixel values in CT scans [39]. **Table 1** displays standard radio densities for different areas of a CT scan. Typically, the air is about one thousand Hounsfield, lung-tissue is generally about the five hundred, water, blood, and other tissues are about zero Hounsfield, and one is generally about seven hundred (700) Hounsfield (HU). Figure 4 shows sickness and normal patients' computed tomography images [41]. In Fig. 4 (a), (b) and (c) displays the normal patient lung images. But in Fig. 4 (d) and (e) displays some carcinogenic part in circle and don’t spread outside of the lung, this type of tumor is called typical tumors. Cancer is any disease in which normal cells are damaged and as quickly as they are spread by mitosis do not experience programmed cell death. Carcinogens can increase cancer risk and cause uncontrolled, cancerous division, ultimately leading to tumor formation by directly altering cellular breakdown or spoiling DNA in cells that interferes with genetic procedures. Significant DNA spoil normally results in programmed cell expiry, but the cell would not prevent it from becoming a cancer cell by itself if the programmed cell death mechanism is disrupted.

3.1.1 Dataset Description

The Lung Imaging Consortium Imaging Array (LIDC-IDRI) dataset and the Computer Science Bowl 2017 training set are the primary CT lung scan datasets [39]. The archive comprises the whole of 1018 helical thoracic CT scans obtained retrospectively from the image archiving and communications networks of the seven participating Academic Institutions with sufficient local IRB clearance [42]. A combined dataset from eight public medical databases is 3DSeg. For either CT or MR scans [43], it protects multiple organs/tissues of concern [44]. The 2016 Lung Nodule Research database (abbreviated as LUNA) and the Computer-Science-Bowl 2017 training package (abbreviated as CSB). The LUNA [45], the dataset contains 1186 nodule marks in 888 radiologist-annotated cases, whilst the Defense-Science-Board-Dataset only provides binary labels showing whether the patient was treated with lung cancer in the year following the scanning. In its preparation, validation, and test collection, the Defense Science Board (DSB) dataset contains 1397, 198, 506 individuals. The 754 training set nodules and 78 validation set nodules are manually counted [46]. There are broad nodules with a mean diameter of about 13.68 mm in the DSB dataset, and there are smaller nodules with a diameter of about 8.31 mm in the LIDC-IDRI dataset. If this data is included in the DSB dataset to identify nodules, it is recommended that nodules lesser than 6 mm be removed from the LIDC-IDRI annotations [39]. Used a subset of data from the LDCT-arm of the National Lung Screening Trail (NLST) database. NLST study covered 3 years: a baseline scan (T0) followed by two successive scans (T1 and T2) spaced nearly 1 year apart [11]. The subset preferred for the study was split down into two thematic cohorts: cohort 1 (85 cancer pulmonary nodules and 176 positive control nodules) and cohort 2 separately (85 lung cancer nodules and 152 positive control nodules) [11].

3.1.2 Pre-processing on DICOM Database

To standardize all input data, some researchers used threshold segmentation with Hounsfield Units and CT image resampling.

Raw in out images involve noise, so pre-processing is the first stage in the identification process, which is enhancing the accuracy of an image to be used further by eliminating unnecessary image detail, known as image noise. If this problem is not adequately handled, many inaccuracies will occur in the classification [47]. The need to conduct this pre-processing is due to the poor contrast between skin lesion and stable skin surrounding, abnormal boundary and skin artifacts, which are skin lines, hairs and black frames, in addition to inaccuracies [47]. In Fig. 5, it is possible (Don’t use jargons) to illustrate four processes and is explained as follows [39].

- Installing the files is the 1ST task. One computed tomography scan/person is used in any folder in the dataset. They are in the style of DI-COM [39]. As a common pattern of clinical images, they Describe-Digital –Imaging-Communication in Medicine (DICOM). It includes the photographs and details of the patient, the features of the computed tomography scan that produced the image, and the image itself characteristics [48].
- The raw dataset scans are not on the Hounsfield scale. As a result, it is essential to do a transformation process. The value of the pixels around the thorax is 2000. The first step is to define these values to zero, which at present is equal to air. Then, it is translated to the hours-field-scale, multiplying and adding the intercept with the re-scale slope [39].
- 3rd part is re-sampling. The pixel spacing of a computed tomography scanner may be different from another one. It relies on an imaging scan. A typical way to deal with this issue is the resampling of the complete information set to a certain isotropic resolution [39]. There are several attributes in a DICOM data object, including objects such as name, ID, etc., and a special attribute that contains a pixel image data object (i.e. technically, there is no" header "as such in the main object, being merely a attributes list, including data of pixel) [48].
- After the lung nodule detection and nodule feature computation the next step is lung nodule classification. This process use for nodule classification in the form of cancerous or non- cancerous nodule. This is final stages. In this stage the Fully Connected layers (FC) are used for classification task. There is also a completely linked layer as part of the convolutional network that takes the end result of the convolution/pooling phase and reaches a decision on classification.

3.2 Deep Learning (DL)

The DL model [5] class is the Convolution-Neural-Network (CNN) [49]. An interesting approach to adaptive image processing is the relationship between general neural-networks and adaptive filters classified by Convolutional Neural Networks (CNNs). Two-Dimensional (2-D) CNN’s produce one or additional layers of 2D filter, with potential non-linear activation and/or down-sampling functions [49]. As seen from Fig. 6, Convolution, Pooling, and Fully-connected layer are the key components of the CNN. In the convolution layer, the feature map is created by applying the dot product's numerical weight grid operation over the entire content of each input data example, such as images, videos, and many others. [5]. First, the pooling-layer includes a standard down-sampling operation that decreases the function maps' in-plane dimensionality to add a transformation invariance of minor moves and deformations and minimize the quantity of learnable parameters that follow [50]. This involves further sub-processing steps in the traditional approach of nodule identification, which explicitly represents greater mistake and analytical time-period [5]. Since it is an End-to-End technique that does not rely much on segmentation, extraction of features, and preprocessing the computational range and time-period is much lower [5].

The workflow of the two-step groups (Traditional approach and CNN-based approach), lung nodule identification is shown in Fig. 3 **(a)** and **(b)** respectively. The standard image-process system is used only for lung nodule identification, while CNN is used to train and evaluate the division of nodule in the case of the CNN-based system [5].

- **Two-Dimension Convolutional Neural Network (2D-CNN)**

A simple 2-D convolution operation was used in this method to define basic characteristics from the whole-images. This CNN architecture is mainly used for its flexibility and reliability in computing. Many images are two-dimensional in the form of a matrix and contain intensity values [51]. The FP rate of the lung nodule is effectively and reliably decreased by decreasing the bias of a single network, 2D-CNN designs the correct single-view model [52].

- **Three-Dimension Convolutional Neural Network (3D-CNN)**

The simple convolution operation was implemented in 3-D (x, y, z) in the same period in the network. The 3-D convolution filters used are three-dimension and have been implemented for automated feature extraction over the 3-D input data. In terms of computational efficiency, these networks are costly than 2-D-based networks. 3-D-based conversion operations require more calculations and more memory since the three-dimension evaluation is needed to store the features extraction in the computer cache. 51, 52. A CNN architecture variant is defined in which a 3D image is input. On the kernels, they add an axis and describe a 3D max pool. Using the previous method, multiple approaches to the classification task were introduced [39].

3.2.1 Convolutional Layer

As an operation on two functions, a convolution is defined. In image processing, one function comprises of input values at the image location, such as pixel values, and a filter function is the second function (or kernel) [53]. The convolution process is described by the * symbol where:

- *K = Kernel*
- *I (t) = Input*
- *O (t) = Output*
- *t = Integer values*
- *a, b, n, m = number of matrix*

An output (or feature map) *O (t)* is well-defined below when input *I(t)* is convolved with a kernel K(a) .

- *O(t) = (I * K) (t). (1)*

If only integer values can be taken up by t, the discretized convolution is given by:

- *O(t) =* $\sum_a I\left(a\right).K(t-a)$*. (2)*

A kernel *K (a, b)* and Input *I (m, n)* is defined as:

- $O(t) = \Sigma_a \Sigma_b I\,(a, b).\ K\,(m - a, n - b).\ (3)$

The kernel is flipped and the above is equivalent to by the commutative law:

- $O(t) = \Sigma_a \Sigma_b I\,(m - a, n - b).\ K\,(a, b).\ (4)$

The cross-correlation function, which is the same as convolution, is implemented by Neural Networks without flipping the kernel:

- $O(t) = \Sigma_m \Sigma_n I\,(m + a, n + b).\ K\,(a, b).\ (5)$

3.2.2 Rectified Linear Unit Function

A RELU sheet is an instigation function which sets minus input values to zero. This simplifies and speed up measurements and preparation and helps to prevent the gradient dilemma that is fading. Tanh, Sigmoid, Randomized RELUs, Leaky RELUs and Parametric RELUs are other activation functions. [53].

Where:

- *x = Input to neurons*

This simplifies and speeds up computations and planning and supports to avoid the gradient issue from declining. It is described statistically as:

- $f\,(x) = max\,(0, x)\ (6)$

3.2.3 Pooling Layer

The pool layer among ReLU layers and the convolution is added to minimize the parameters number to be measured, in addition to the image size (height and width but not depth). Max-pooling is generally used; in additional pooling layers, average pooling and L2-normalization pooling are used. Max-pooling proceeds the greatest input value in a filter and removes the other values; it effectively sums up the highest activations in a field [53].

3.2.4 Fully Connected Layer

The output layer of a CNN is the Connected Layer, which means that in the Completely Connected Layer, every neuron in the previous layer is connected to every neuron. Dependent on the degree of characteristic abstraction required, there may be 1 or more FC layers, with convolution, RELU, and pooling layers. This layer takes as its input the result of the prior layer (RELU, Convolution or Pooling) and calculates a probability score for classification into the various classes obtainable [53]. Figure 7 shows the process of lung detection using CNN.

3.3 Works Based on CNN Architecture

Setio et al. [54] a simple 2-D convolution operation was used in this approach to define local features from the entire image. For its simplicity and computational effectiveness, this CNN architecture is also used. Most of the pictures are 2D [54]. This network can automatically learn multiple 2-D filters/functions from the train dataset. Setio et al. [54] presented the several Conv-Net 2-D streams (Table 2). Initially, for nodule candidate detection, three candidate detector algorithms built specifically for solid, sub-solid and significant nodules were combined. Then a set of 2-D patches of 64 * 64 pixels were collected from differently oriented planes for each candidate. For feature extraction, many 2-D Conv-Nets were then used. The authors have suggested that the late fusion approach offers greater recognition operation in contrast to committee mixture and mixed fusion techniques. A committee mixture is a combined neural network resulting from other neural networks. Various algorithms may be used to assimilate expert into a single output such as decision. W. Shen et al. [55] by developing a "multi-crop" pooling approach for automatic lung nodule characterization, Shen et al. have developed a Multi-Crop Convolutional Neural Network (MC-CNN) [55]. They suggested a multi-crop-based approach for pooling operations. Three pooling operations consist of the multi-crop pooling approach. Convolution features (R0), acquired either from the pooled features or from the original input image, were the input to the first pooling process. The concatenated nodule-centered feature $f = [f_0, f_1, f_2]$ consists of three nodule-centered feature patches R_0, R_1, R_2. Specifically, R_0 size be $l * l * n$, where $l * l$ is the feature map dimension and n is the number of feature maps:

$$f_{\,i = max-Pool}{}^{(2-i)} \{R_i\}.\ i = (0, 1, 2)\ (7)$$

Where $R1, R2$ is a middle region with a scale of $(l/2) * (l/2) * n$ and $(l/4) * (l/4) * n$. The "max-pool" superscript indicates the frequency of the max-pooling process used on Ri. $R1$ was the central region of $R0$, and $R2$ was the middle region of $R1$. After that, $R0$ has max pooled double and $R1$ has max pooled once to create $f0$ and $f1$, the pooled feature maps, respectively. The final multi-crop characteristics were figured by concatenating $f0, f1$, and $f2$. The study used Rectified Linear Units (ReLU) as a Stochastic Gradient Descent (SGD) learning algorithm activation function that minimizes cross-entropy learning. Q. Dou et al. [56] suggested the 3-DCNN-based architecture (Table 2) for detection. Investigators develop the 3-D CNNs to use hierarchical architecture to encrypt spatial details and representative features. An area of size 20* 20* 6 was used for the first CNN model, size 30 * 30 * 10 for the second architecture, and 40* 40 * 26 for the third architecture. Finally, for nodule detection, the 3 CNNs have been combined and serve as a feature extractor. The suggested architecture was validated and achieved result in the form of sensitivity 94.4% for nodule detection by contributing in the LUNA16 (LUNG Nodule Analysis 2016) competition. In Table 2 shows the CNN-based DL approach in lung CT images.

Five lung nodule forms, viz., juxtapleural, well-circumscribed, pleural tail, GGO and juxtavascular using CNN, have been classified by Yuan et al. [57]. Researchers used a multi-scale, multi-view CNN, with the Visual Geometry Group (VGG) network. Next, the descriptor of the SIFT function was computed and encoded in the Fisher Vector type (FV). To measure the Fisher geometrical characteristics, the Gaussian mixture model was used. Finally, using Multiple Kernel Learning, all handcrafted features were combined with CNN extracted features (MKL). An overall precision of 93.1% was achieved by Yuan et al.[57]. While Paul et al.[58] suggested the latest CNN architecture for characterization of lung nodule. Scholars used a pre-trained VGG network trained with the Image-net dataset. For this reason, three networks, viz., VGG-s, VGG-m and VGG-f were used. The s, m, and f stand for slow, medium, and fast here and refer to time for training. Researchers have registered an accuracy using pre-trained CNN of 76.79%[58]. Ciompi et al.[59] suggested a 2-D CNN architecture for nodule detection called "Over-Feat". 6 convolutional layers were used in the architecture, with filter sizes ranging from 7 *7 to 3 * 3. Finally, a two-stage classifier has been used by researchers for perifissural nodule detection [59].

Table 2
– CNN Based DL method in lung CT images

Author	Purpose	Method	Result
R. Paul, et al. [58]	Nodule classification	VGG-network with CNN Network Merge	ROC = 0.87, Accuracy = 76.79%
F. Ciompi et al. [59]	Nodule’s detection	Named as “OverFeat” 2-D convolution-based architecture	Area under curve (AUC) = 0.868
Yang F, et al. [55]	Nodule classification	MC-CNN (Multi-crop convolutional neural network)	Sensitivity = 77%, Specificity = 93%, Accuracy = 87.14%, AUC = 0.93
J. Yuan et al. [57]	Nodule classification	CNN-based,VGG network	Accuracy = 93.1% (LIDC dataset) and 93.9% (ELCAP dataset)
Q. Dou et al. [56]	Nodule detection nodule	With fusion technology, three 3-D CNN added	Sensitivity = 92.2%, Sensitivity = 94.4%
Setio et al. [54]	Nodule detection nodule	Multiple streams of 2-D Conv-Nets	Sensitivity = 85.4% with 1 FP/scan and 90.1% with 4 FPs/scan

Deep Learning's benefits or advantages are as follows: features are deduced dynamically and optimally tuned for the desired result. Extracting functionality in advance is not appropriate. This eliminates methods of machine learning that are time-intensive. Robustness in the outcomes is taught automatically against normal variants [36]. For several different applications and kinds of data, the same neural network-based approach can be applied. Graphics Processing Units (GPUs) can be used to perform large parallel computations and are ideal for vast amounts of data. Also, since the data volume is high, it provides enhanced performance results. To adapt to emerging problems in the future, the architecture for deep learning is scalable [36].

The drawbacks of Deep Learning in terms of better performance than other method is that it needs very large volumes of data. Due to complicated data structures, it is incredibly costly to practice. In comparison, deep learning calls for pricey GPUs and hundreds of processors. It boosts prices for consumers [36]. In choosing the right deep learning methods, there is no standard theory to help you as it involves knowledge of topology, method of training, and other parameters. Based on mere learning, it is not simple to understand performance and requires classifiers to do so. Convolution algorithms based on the neural network execute those functions [36].

3.3.1 Image-Processing

Medical image processing has seen intense development and has become an interdisciplinary area of study that has drawn expertise from many parts of engineering and science. Diagnostic therapy with computed tomography has proven to play an important role in the clinical routine [60]. Many problems emerge, complemented by a surge of emerging high-technology growth and the use of several imaging modalities; for example, how to store and interpret a big-amount of images so that high-quality knowledge can be collected for disease detection and treatment [60]. The gain of greater visibility reduced noise, and distortion of the machine tomography image is the main advantage. In this regard, one can conveniently determine the mean and variance [61]. The calculated value is very similar to the original value. Image denoising algorithms may be the most commonly used in image processing [61]. The purpose of image improvement is to enhance the generalizability or representation of images for examination purpose or to provide 'measured' data for other optical image processing at the level of image preprocessing. However, both images have undergone multiple pre-processing steps [62].

3.4 Measurement Parameter

The exact categorization of the positive class is True Positive (TP), e.g., if an image shows cancer cells and the model effectively segments the cancer component and the cancer presence is classified by the result. The correct categorization of the negative class is True Negative (TN), as there is no cancer in the picture, and the model declares that cancer is not present after categorization. The incorrect positive prediction is False Positive (FP), e.g., the image of cancer cells, but the model classifies that the image does not have cancer. The wrong estimate of the negatives is False Negative (FN). Three measurement parameters describe in following subsection.

3.4.1 Accuracy

The accuracy of first test is calculated by equating the findings of a medical test (positive or negative) with the patient's current disorder or disease (presence or absence) [63]. Sensitivity (SENS) and Specificity

(SPES) are the two fundamental metrics for quantifying the diagnostic accuracy of a procedure [63].

3.4.2 Sensitivity and Specificity

Based on how many individuals (not the whole population) have the disorder, susceptibility is determined. Using the equation: sensitivity = number of TP / (number of TP + number of FN), it can be measured. Based on how many persons do not have the disorder, precision is determined. The equation can be used to quantify it: precision = number of TN (number of TN + number of FP) [64].

- *Sensitivity = TP/(TP + FN) (8)*
- *Specificity = TN/(FP + TN) (9)*
- *Accuracy = TP + TN/(TP + TN + FP + FN) (10)*

4. Review Of Deep Learning Measurement Methods

This Section explains the findings of the various datasets for the lung nodules detection (Section 4.2), the classification of lung nodules (Section 4.3) and the classification and lung nodules detection in Section 4.4. In reports using dissimilar test and training datasets, the performance of the DL technology was analogous to analyses using some types of test and training dataset. Figure 9 illustrates the study as a flowchart for PRISMA.

4.1 Literature Search Results

In this study, there were a total of 30 studies included. 16 studies have described outcomes for both detection and classification operation. The efficacy of the various nodule classification processes is shown in Table 3 when arranged according to unique types of output measurements.

10 studies analyzed nodule classification efficiency (Table 4) and 9 studies investigated the use of DL in Table 5, i.e. nodule or non-nodule, for nodule detection. The studies listed four distinct DL algorithms: Massive Training Artificial Neural Network (MTANN), Convolution Neural Network (CNN), Wavelet Recurrent Neural Network (WRNN), Computer Assisted Detection/Diagnosis (CAD) and Optimal Deep Neural Network (ODNN). Both MTANN and CNN are machine-learning end-to-end processes, meaning that inputs are full pixelated images and are processed and trained using backpropagation without established components of precise feature detection. Compared to, for example, CNN, the benefits of stacked auto-encoders involve less training instances, since stacked auto-encoders are capable to produce latest pictures from image characteristic vector features [50]. The MTANN filter was trained with real nodules in CT images to enhance real nodule configurations. The sensitivity and precision of the CAD scheme has been greatly enhanced by the use of the MTANN method. DeepLNAnno has a different 3-tier

working mechanism and also develops the precision of the labels compared to some other annotation systems.

In order to systematically search articles and literature for review-based investigation, the Preferred Reporting an Items for Systematic Reviews and Meta-Analyses (PRISMA) instrument or framework uses a range of methods. In addition, in any sort of research that systematically assesses the quality of selected papers and either includes them for the study or excludes them for the study, PRISMA is often focused on formulated inclusion and exclusion criteria. In Fig. 8 shows the PRISMA flowchart. 2778 studies imported for screening. In the 2778 studies 996 studies duplicates removed then 1782 studies were screened. In the 1782 studies 1567 studies irrelevant, then 1567 studies were removed from 1782 studies. 215 studies full text assessed for eligibility and 180 studies are excluded with reasons then finally 35 studies were included.

4.2 Lung Nodule Detection

Xu et al. [65], F. Liao et al.[46], S.Kar et al.[66] and M. Liu et al.[67] proposed the CNN based algorithms for nodule detection. X. Xu et al.[65] tested their CNN based method 3D-CNN on case the DeepLNdataset, while F. Liao et al.[46] tested their method on Data Science Bowl 2017 competition [46]. The 3rd study by A. Masood et al[68] tested their Fully CNN (F-CNN) based on method from the LIDC-IDRI, Lung CT-Diagnosis, Lung Nodule Analysis (LUNA) 2016 datasets and achieved the result 77.6% and 84.58%. All the results reached between 76.6-97.59%. Only N. Tajbakhsh et al.[69] published a sensitivity result which was 100% using MTANN method.

S. Chen et al.[70] and R. Jones et al[71] proposed the D-CNN and MTANN respectively. S. Chen et al.[70] S. Cheng et al.[70] suggests a web-based annotation approach for lung nodules named DeepLNAnno. DeepLNAnno has a special three-tier working mechanism and tones of features such as semi-automatic annotation, which not only makes it much easier for doctors to annotate, but also develops the precision of the labels compared to some other annotation systems and achieved the result in the form of accuracy 87.5%. R. Jones et al[71] presented the MTANN method. The sensitivity and precision of the CAD scheme has been greatly enhanced by the use of the MTANN method. With a database of 69 lung cancers, nodule candidate detection by the MTANN filter achieved 97% sensitivity with 6.7 false positives per section [71]. Classification MTANNs were applied for further reduction of FPs. The MTANN classification removed 60% of FPs with a loss of 1 TP; thus, it achieved a 96% sensitivity with 2.7 FPs per segment. Overall, with the CAD system based on the MTANN filter and classification MTANNs, an 84% sensitivity with 0.5 FPs per segment was achieved [71]. S.Kar et al.[66] and Setio et al.[54] both examined their algorithm on the same types of databases (LIDC) and achieved the sensitivity 85.4%-90.1% and 97.59% respectively.

4.3 Lung Nodule Classification

For the classification of lung nodule total 10 studies were included. All studies reported of sensitivity and specificity, which ranged between 76.5%-99.9% and 80.1%-98.7% respectively. Five of these studies, C. J.

Lin et al. [72], M. Usman et al. [73], Y.Wang et al. [74], R.Pual et al.[75], S. A. El-Regaily et al.[76] and G. Jakimovski et al.[77] has CNN architecture based on DL, while only Lakshmanp et al.[78] used Optimal Deep Neural Network (ODNN).

S. Wang et al.[79] indicated the 3D-CNN on CT data obtained from the Fudan University Shanghai Cancer Center of 1,545 patients with pre-invasive and invasive lung cancer, the sensitivity, specificity and AUC automatic classification system were 88.5 percent, 80.1 percent and 89.2 percent, respectively. Y.Wang et al.[74] tested and trained the CT scan from the LIDC/LUNA databases. A total of 888 CT scans of LUNA were tested and final score was 0.69% and then further tested Intelligent Imaging Layout System (IILS) on LIDC/IDRI database with 1018 CT scans, while Zhang et al.[80] used an algorithm that was both tested and trained on the Lung Nodule Research database (LUNA 16) and achieved the sensitivity of 96.0% and specificity 88.0% for their accurate performing. S. A. El-Regaily et al.[76] used the CNN based algorithm using CAD system and achived the sensitivity of 85.25% and specificity of 90.2%.

4.4 Lung Cancer Nodule for Both (Classification and Detection)

For both detection and classification total 16 studies were included. Two of the studies [81],[4] reported accuracy value of 80% and 77% respectively on same dataset (LUNA16) for both detection and classification of lung nodule, while Harsono et al.[82] used the Ratine U-Net 3D. the result obtained by training I3DR-Net model on 1009 LIDC dataset for both detection and classification.

The Wavelet Recurrent Neural Network (WRNN) was introduced by Nurtiyasari et al.[83] where the Wavelet model was used for the lung diagnosis method and the Recurrent Neural Network for the classification phase. The outcome sensitivity, precision and accuracy of the lung nodule classification using WRNN were 93.75%, 66.6% and 84% for training and 88.2%, 75% and 84% for testing results. N. Tajbakhsh et al.[84] collected information of the false positive (FP) when 100% result in form of sensitivity was achieved with Massive-Training-Artificial Neural Network (MTANN), which result in 22.7 FP in %.

Gu et al.[85] used 3D deep CNN with multi-scale prediction to detect lung nodule after lung segmentation from chest CT scans for detection efficiency assessment, with a systematic procedure used. At one and four false positives/scan, the sensitivity of the proposed method with the primary alternative reached 87.94% and 92.93%, respectively. Paul et al.[11] tested on dataset from The National Lung Screening Trail (NLST) and achieved the result AUC of 96.0% and accuracy 89.45%.

Eun et al.[86] introduced a new FP reduction framework for pulmonary nodule detection, the whole of single-view 2D CNNs with completely automatic without nodule categorization, and achieved the result of 92.2% metric score. Cao et al.[87] tested on the LUNA16 dataset for nodule detection using Multi-Branch Ensemble Learning architecture based on the 3D CNN (MBEL-3DCNN). On the detection they reached a sensitivity 72.9%. Winkels et al.[88] tested and trained on NLST or LIDC datasets for both detection and classification of lung nodule and also defined the Free-Response Operating Characteristic (FRCO) curve.

Table 3
Recent literature for classification and detection

Author	Year	Method	Result
(D. Nurtiyasari et al.) [83]	2017	Wavelet-RNN	93.75% = Sensitivity, 66.67% = Specificity
(N. Tajbakhsh et al.)[84]	2017	MTANN	100% = Sensitivity
(Gu et al.) [85]	2018	3D-DCNN	87.94%=Sensitivity
(Paul et al.) [11]	2018	Transfer, Ensemble CNN	Accuracy = 84%,76.49%
(Coundray et al.) [89]	2018	Inception v3	Accuracy = 73%, 85%
(Eun et al.) [86]	2018	Ensemble 2D-CNN	92.2%=Metrics score
(Xi et al.) [90]	2019	2D-CNN	86.42%=Sensitivity
(Winkels et al.) [88]	2019	3D-roto- CNN's	(FROC) curve
(Kim et al.) [91]	2019	MGI-CNN's	Average CPM = 98.8%
(Li et al.) [92]	2019	R-CNN	85.2%=Sensitivity
(Zuo et al.) [93]	2019	Multi Resolution-CNN	Accuracy = 96.73%
(Cao et al.) [87]	2019	MBEL-3DCNN	Sensitivity = 72.9%
(Wang et al.) [94]	2019	Classifier based CNN	Sensitivity = 96.8%
(Sharma et al.) [4]	2019	3D-CNN	Accuracy = 77%
(T. Ahmed et al.) [81]	2020	3D-CNN	Accuracy = 80%
(Harsono et al.)[82]	2020	3D-Conv Net	mAP = 49%, 22.86% with AUC 81%

Table 4
Classification performance result in Sensitivity and Specificity

Researcher	Method	Year	Sensitivity	Specificity
S. Wang *et al.*[79]	3D-CNN	2018	88.5%	80.1%
G. Jakimovski *et al.* [77]	CDNN	2018	99.9%	98.7%
Y.Wang *et al.* [74]	Res-Net	2019	76.5%	89%
Zhang *et al.*[80]	3D-CNN	2019	96.00%	88.00%
Polat *et al.* [95]	3D-CNN	2019	88.5%	94.2%
Lakshmanp *et al.* [78]	ODNN	2019	96.2%	94.2%
M. Usman *et al.* [73]	U-Net	2020	89.02%	N/A
S. A. El-Regaily *et al.*[76]	CNN based CAD	2020	85.25%	90.2%
C. J. Lin *et al.* [72]	CNN	2020	99.2%	N/A
R.Pual *et al.*[75]	CNN	2020	90.2%	N/A

Table 5
Detection performance result in Sensitivity and Specificity

Author	Method	Year	Sensitivity	Specificity
R. Jones *et al.* [71]	MTANN	2014	97.00%	N/A
Setio *et al.*[54]	2D-CNN	2016	85.4%-90.19%	N/A
N. Tajbakhsh *et al.* [69]	MTANN	2017	100%	N/A
A. Masood *et al* [68]	F-CNN	2018	74.6%	86.5%
S. Chen *et al.* [70]	D-CNN	2019	94%-97%	N/A
F. Liao *et al.*[46]	3D-CNN	2019	85.6%	N/A
M. Liu *et al.* [67]	3D-CNN	2020	85.6%	N/A
X. Xu *et al.* [65]	3D-CNN	2020	96.65%	N/A
S.Kar *et al.*[66]	CNN	2020	97.59%	97.78%

5. Discussion

A total of 30 analyses that tested DL methods on various datasets were discussed in this proposed work. Several other studies on broad, publicly accessible databases such as LUNA 16, DSB 16, and LIDC have trained and tested DL models, and a systematic review of the different studies tested on these databases has been written. The algorithm with CNN reached accuracy between 73–96% (Table 3) on detection and classification. There was no observed difference in detection accuracy compared to a prior analysis using deep learning processes on CT scans from a different database. Sensitivity and Specificity for detection discovered in this study were between 77.6%-100% (Table 5) and for classification 76.5%-99.2% (Table 4), respectively. This manuscript also explains the various architecture of CNN. The aim of this study is the applications of DL algorithms for lung Nodule Cancer, performance measurements are given for various tests. In this study, numerous DL algorithms are discussed for the classification and detection of lung nodule. Several major corporations have invested in general image recognition of day by day artifacts in studying deep learning and several manufacturers have pushed towards automated recognition in clinical radiology. It would be prudent to contain DL into medical practice with the growing prevalence of AI emerging in health care and the rising capacity for radiologists. All previous studies reported classification of sensitivity, only the result of sensitivity of 2017–2019 was included in their work. The sensitivity ranges from 76.5% – 96.2% for work trained and tested on different datasets, but in our proposed work we have included the sensitivity result of 2020 also. Classification sensitivity ranges from 85% – 99.2% for 2020. The same case-in sensitivity result for the detection of lung nodule 2017–2019 was included. Sensitivity of detection ranges from 74.6–100% for research tested and educated on various datasets. In this proposed work also included the result of 2020 for sensitivity detection, which ranged from 85.6% – 97.59%. In the proposed work also included the accuracy results ranged from 73% – 96.73% from 2017 to 2020 (Table 3). Cao et al.[87] tested on LUNA 16 datasets, had a noticeably low sensitivity result (72.9%). Li et al. 92, tested on the same dataset (LUNA 16), had a noticeable good sensitivity result (85.2%). An equivalent mission was completed by Coudray et al. 89, however, to classify sub-types of lung cancer and to predict common genetic changes, histopathological images were used. Knowing genetic transformations helps to forecast the period of survival and to direct chemotherapy collection. The task was completed by U. Pastorino et al.96 however, the identification of nodules whose size is less than 6 mm is still a difficult task. Using voxel-wise detection is the alternative, but more computing power is needed. In CT, the lung nodule is classified as an oval or round shape of a tissue with a diameter of less than 30 mm. Pulmonary nodules typically have a diameter of more than 4 mm and pulmonary nodules are known as micro-nodules. Blood vessels are frequently misled as nodules that are non-nodules due to the identical presence, resulting in increased false-positive Lung nodules during diagnosis.

The record latest papers were chosen for the research published in Science Direct, IEEE Xplore, and Scopus Indexed journals. It is noted that maximum research performed for the identification of lung tumor have been used in LUNA 16, DSB 16, LIDC datasets and only a few have been used in those datasets. Although some techniques achieve a major development in the outcome of accurate sensitivity or minimum FP, it is found that several problems remain to be solved by measuring the new procedures used to find cancer of the lung nodule. For the detection of lung cancer as primarily as possible, therefore, the effective method is still relevant, for which this article in the review report would be useful for researcher.

6. Conclusion

The analysis of depth learning approaches used in computed tomography scans to distribute with cancer lungs is provided in the proposed review. In the medical community, it is notable that this mission is very relevant because this disorder contributes to many deaths every year. The leading approach to this task is deep learning. Another important consideration for the outcome of the early warning program is competitive mortality threats, particularly where peril is determined by smoking-position, as smoking triggers copious other lung cancer fatal-illness. Regardless of whether the DL methods were trained and evaluated on the similar kind or on dissimilar types of datasets, a pattern of equivalent implementation levels was observed. In the future, convolution neural network will become a powerful implement to support radiologists in their diagnostic work, offering additional correct and time- effective identification and diagnosis of pulmonary nodules; yet extra analyses and improvement are demanded. In this proposed work CNN achieves the accurate results for both classification and detection task and also noted that CNN has more popular recently with the majority of the newer publications using DL method. The proposed research presents useful elements for determining the influence of co-morbidity on the average life expectancy of traditional populations at higher - peril. Most of the studies performed to detect lung cancer have used LIDC, LUNA, DSB, COPD-GENE datasets and only a few have used their datasets. For nodule classification based on the growth rate, various algorithms could be proposed, which will eventually assist in early cancer identification and diagnosis.

References

1. Liu Z, Yao C, Yu H, Wu T. Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things. *Future Generation Computer Systems*. 2019;97:1-9. doi:10.1016/j.future.2019.02.068
2. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*. 2015;136(5):E359-E386. doi:10.1002/ijc.29210
3. Ponnada VT, Naga Srinivasu S V. Efficient CNN for lung cancer detection. *International Journal of Recent Technology and Engineering*. 2019;8(2):3499-3503. doi:10.35940/ijrte.B2921.078219
4. Sharma S, Kaur M, Saini D. Lung cancer detection using convolutional neural network. *International Journal of Engineering and Advanced Technology*. 2019;8(6):3256-3262. doi:10.35940/ijeat.F8836.088619
5. Kumar S B, Vinoth Kumar M. Detection of Lung Nodules using Convolution Neural Network: A Review. Published online 2020:590-594. doi:10.1109/icirca48905.2020.9183183
6. Zhang C, Sun X, Dang K, et al. Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network. *The Oncologist*. 2019;24(9):1159-1165.

doi:10.1634/theoncologist.2018-0908

7. Techniques L. healthcare Analyzing Lung Disease Using Highly E ff ective Deep. Published online 2020:1-21.
8. Lee SLA, Kouzani AZ, Hu EJ. Automated detection of lung nodules in computed tomography images: A review. *Machine Vision and Applications*. 2012;23(1):151-163. doi:10.1007/s00138-010-0271-2
9. Rossetto AM, Zhou W. Deep Learning for Categorization of Lung Cancer CT Images. *Proceedings - 2017 IEEE 2nd International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2017*. Published online 2017:272-273. doi:10.1109/CHASE.2017.98
10. Bhattacharyya S, Chaki N, Konar D, et al. *Advanced Computational and Communication Paradigms*. Vol 475. Springer Singapore; 2018. doi:10.1007/978-981-10-8240-5
11. Paul R, Hall L, Goldgof D, Schabath M, Gillies R. Predicting Nodule Malignancy using a CNN Ensemble Approach. *Proceedings of the International Joint Conference on Neural Networks*. 2018;2018-July. doi:10.1109/IJCNN.2018.8489345
12. Essaf F, Li Y, Sakho S, Gadosey PK. Improved Convolutional Neural Network for Lung Cancer Detection. *ACM International Conference Proceeding Series*. Published online 2020:48-54. doi:10.1145/3398329.3398337
13. Kriegsmann M, Haag C, Weis CA, et al. Deep learning for the classification of small-cell and non-small-cell lung cancer. *Cancers*. 2020;12(6):1-15. doi:10.3390/cancers12061604
14. Li D, Vilmun BM, Carlsen JF, et al. The performance of deep learning algorithms on automatic pulmonary nodule detection and classification tested on different datasets that are not derived from LIDC-IDRI: A systematic review. *Diagnostics*. 2019;9(4). doi:10.3390/diagnostics9040207
15. Biesalski HK, Bueno De Mesquita B, Chesson A, et al. Consensus statement on lung cancer. *European Journal of Cancer Prevention*. 1997;6(4):316-322. doi:10.1097/00008469-199708000-00002
16. Hecht SS. Lung carcinogenesis by tobacco smoke. *International Journal of Cancer*. 2012;131(12):2724-2732. doi:10.1002/ijc.27816
17. Mustafa M, Azizi ARJ, IIIzam EL, Nazirah A, Sharifa S, Abbas SA. Lung Cancer: Risk Factors, Management, And Prognosis. *IOSR Journal of Dental and Medical Sciences*. 2016;15(10):94-101. doi:10.9790/0853-15100494101
18. Oncology T. Lung cancer Lung cancer Lung cancer. *Conn's Current Therapy 2020*. 2013;2030(November):133-141.
19. Roche. Lung cancer A guide for journalists on Non-Small Cell Lung Cancer (NSCLC) and its treatment. Published online 2013.
20. Oncology S. Small-cell carcinoma.
21. Sata N, Tsukahara M, Koizumi M, et al. Primary small-cell neuroendocrine carcinoma of the duodenum - A case report and review of literature. *World Journal of Surgical Oncology*. 2004;2:1-7. doi:10.1186/1477-7819-2-28

22. Uebelacker LA. HHS Public Access.*Physiology & behavior*. 2017;176(1):139-148.
23. Lung L stage SC. Limited-stage small cell lung carcinoma.
24. Byers LA, Rudin CM. Small cell lung cancer: Where do we go from here? *Cancer*. 2015;121(5):664-672. doi:10.1002/cncr.29098
25. Lung T, Dna SC, Five-year NS, et al. Non-small-cell lung carcinoma.
26. Liang Z, Xu J, Ma Z, Li G, Zhu W. MiR-187 suppresses non-small-cell lung cancer cell proliferation by targeting FGF9. *Bioengineered*. 2020;11(1):70-80. doi:10.1080/21655979.2019.1706287
27. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine*. 2018;24(10):1559-1567. doi:10.1038/s41591-018-0177-5
28. Lukeman JM. What Is Lung Cancer? *Perspectives in Lung Cancer*. Published online 2015:30-40. doi:10.1159/000400400
29. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clinic Proceedings*. 2008;83(5):584-594. doi:10.4065/83.5.584
30. European S for MO. What is Non-Small-Cell Lung Cancer? Let us answer some of your questions. *ESMO Patient Guide Series*. Published online 2018:7.
31. Moorthy C, Bharanikumar S, Bharath K, Gobinath S, Gokulraj S. Prognosis of Lung Cancer System using 3D Deep Convolutional Neural Network. 2020;26(5):142-146.
32. Gadgeel SM, Ramalingam SS, Kalemkerian GP. Treatment of Lung Cancer. *Radiologic Clinics of North America*. 2012;50(5):961-974. doi:10.1016/j.rcl.2012.06.003
33. Sulu E, Tasolar O, Takir HB, Tuncer LY, Karakurt Z, Yilmaz A. Delays in the diagnosis and treatment of non-small-cell lung cancer. *Tumori*. 2011;97(6):693-697. doi:10.1700/1018.11083
34. Function R, Erythrocytes R. Tyrosine Kinase. *Encyclopedia of Cancer*. Published online 2011:3822-3822. doi:10.1007/978-3-642-16483-5_6079
35. Cooley ME, Short TH, Moriarty HJ. Symptom prevalence, distress, and change over time in adults receiving treatment for lung cancer. *Psycho-Oncology*. 2003;12(7):694-708. doi:10.1002/pon.694
36. Kota VM, Manoj Kumar V, Bharatiraja C. Deep Learning - A Review. *IOP Conference Series: Materials Science and Engineering*. 2020;912(3). doi:10.1088/1757-899X/912/3/032068
37. Hernández Muñiz S, Mitjavdla Casanovas M. Introduction to computed tomography. *Revista Espanola de Medicina Nuclear*. 2006;25(3):206-214. doi:10.1157/13088421
38. Wolthaus JWH, Schneider C, Sonke JJ, et al. Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients. *International Journal of Radiation Oncology Biology Physics*. 2006;65(5):1560-1571. doi:10.1016/j.ijrobp.2006.04.031
39. Borja MGB, Huauya R, Lazo C. A brief survey on deep learning based methods for lung cancer classification using computerized tomography scans. *IEEE CHILEAN Conference on Electrical,*

Electronics Engineering, Information and Communication Technologies, CHILECON 2019. Published online 2019:1-5. doi:10.1109/CHILECON47746.2019.8987722

40. Johnson M, Surgeon CU, Trust F. diagnosis , pathology & prognosis Plan for today :
41. Alakwaa W, Nassef M, Badr A. Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *International Journal of Biology and Biomedical Engineering*. 2017;11(8):66-73. doi:10.14569/ijacsa.2017.080853
42. Armato S, Mclennan G, Mcnitt-gray M, et al. WE-B-201B-02: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Public Database of CT Scans for Lung Nodule Analysis. *Medical Physics*. 2010;37(6):3416-3417. doi:10.1118/1.3469350
43. Zbontar J, Knoll F, Sriram A, et al. fastMRI: An Open Dataset and Benchmarks for Accelerated MRI. Published online 2018:1-35.
44. Chen S, Ma K, Zheng Y. Med3D: Transfer Learning for 3D Medical Image Analysis. 2019;(November).
45. Fb QHLHL, B SBJQ. . 2i2 + iBM ; GmM ; * M + 2 ` G2bBQMb BM * h AK ; 2b mbBM ; Published online 2019:114-118.
46. Liao F, Liang M, Li Z, Hu X, Song S. Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network. *IEEE Transactions on Neural Networks and Learning Systems*. 2019;30(11):3484-3495. doi:10.1109/TNNLS.2019.2892409
47. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. Cancer diagnosis using deep learning: A bibliographic review. *Cancers*. 2019;11(9):1-36. doi:10.3390/cancers11091235
48. Imaging D. [2] [3] [4]. Published online 2017.
49. Browne M, Ghidary SS. Convolutional neural networks for image processing: An application in robot vision. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2003;2903(May):641-652. doi:10.1007/978-3-540-24581-0_55
50. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*. 2018;9(4):611-629. doi:10.1007/s13244-018-0639-9
51. Halder A, Dey D, Sadhu AK. Lung Nodule Detection from Feature Engineering to Deep Learning in Thoracic CT Images: a Comprehensive Review. *Journal of Digital Imaging*. 2020;33(3):655-677. doi:10.1007/s10278-020-00320-6
52. Lin CJ, Jeng SY, Chen MK. Using 2D CNN with taguchi parametric optimization for lung cancer recognition from CT images. *Applied Sciences (Switzerland)*. 2020;10(7). doi:10.3390/app10072591
53. Ker J, Wang L, Rao J, Lim T. Deep Learning Applications in Medical Image Analysis. *IEEE Access*. 2017;6:9375-9379. doi:10.1109/ACCESS.2017.2788044
54. Setio AAA, Ciompi F, Litjens G, et al. Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks. *IEEE Transactions on Medical Imaging*. 2016;35(5):1160-1169. doi:10.1109/TMI.2016.2536809
55. Shen W, Zhou M, Yang F, et al. Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition*. 2017;61:663-673.

doi:10.1016/j.patcog.2016.05.029

56. Dou Q, Chen H, Yu L, Qin J, Heng PA. Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection. *IEEE Transactions on Biomedical Engineering*. 2017;64(7):1558-1567. doi:10.1109/TBME.2016.2613502
57. Yuan J, Liu X, Hou F, Qin H, Hao A. Hybrid-feature-guided lung nodule type classification on CT images. *Computers and Graphics (Pergamon)*. 2018;70:288-299. doi:10.1016/j.cag.2017.07.020
58. Paul R, Hawkins SH, Schabath MB, Gillies RJ, Hall LO, Goldgof DB. Predicting malignant nodules by fusing deep features with classical radiomics features. *Journal of Medical Imaging*. 2018;5(01):1. doi:10.1117/1.jmi.5.1.011021
59. Ciompi F, de Hoop B, van Riel SJ, et al. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Medical Image Analysis*. 2015;26(1):195-202. doi:10.1016/j.media.2015.08.001
60. Inage T, Nakajima T, Yoshino I, Yasufuku K. Early Lung Cancer Detection. *Clinics in Chest Medicine*. 2018;39(1):45-55. doi:10.1016/j.ccm.2017.10.003
61. Sharma D, Jindal G. Identifying Lung Cancer Using Image Processing Techniques. *International Conference on Computational Technique and Artificial Intelligence*. Published online 2011:115-120.
62. Chaudhary A, Singh SS. Lung cancer detection on CT images by using image processing. *Proceedings: Turing 100 - International Conference on Computing Sciences, ICCS 2012*. Published online 2012:142-146. doi:10.1109/ICCS.2012.43
63. Mandrekar JN, Ph D, Mandrekar SJ. Statistical Methods in Diagnostic Medicine using SAS ® Software. *Analysis*. Published online 1990:1-8.
64. Swift A, Heale R, Twycross A. What are sensitivity and specificity? *Evidence-Based Nursing*. 2020;23(1):2-4. doi:10.1136/ebnurs-2019-103225
65. Xu X, Wang C, Guo J, et al. DeepLN: A framework for automatic lung nodule detection using multi-resolution CT screening images. *Knowledge-Based Systems*. 2020;189:105128. doi:10.1016/j.knosys.2019.105128
66. Kar S, Das Sharma K, Maitra M. Adaptive weighted aggregation in Group Improvised Harmony Search for lung nodule classification. *Journal of Experimental & Theoretical Artificial Intelligence*. 2020;32(2):219-242. doi:10.1080/0952813x.2019.1647561
67. Liu M, Jiang X, Liu Y, Zhao F, Zhou H. A semi-supervised convolutional transfer neural network for 3D pulmonary nodules detection. *Neurocomputing*. 2020;391:199-209. doi:10.1016/j.neucom.2018.12.081
68. Masood A, Sheng B, Li P, et al. Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images. *Journal of Biomedical Informatics*. 2018;79(December 2017):117-128. doi:10.1016/j.jbi.2018.01.005
69. Tajbakhsh N, Suzuki K. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition*. 2017;63(October 2016):476-486. doi:10.1016/j.patcog.2016.09.029

70. Chen S, Guo J, Wang C, Xu X, Yi Z, Li W. DeepLNAnno: a Web-Based Lung Nodules Annotating System for CT Images. *Journal of Medical Systems*. 2019;43(7):1-9. doi:10.1007/s10916-019-1258-9
71. Jones R. NIH Public Access. *Bone*. 2014;23(1):1-7. doi:10.1088/0031-9155/54/18/S03.Supervised
72. Lin CJ, Li YC. Lung nodule classification using taguchi-based convolutional neural networks for computer tomography images. *Electronics (Switzerland)*. 2020;9(7):1-9. doi:10.3390/electronics9071066
73. Usman M, Lee BD, Byon SS, Kim SH, Lee B il, Shin YG. Volumetric lung nodule segmentation using adaptive ROI with multi-view residual learning. *Scientific Reports*. 2020;10(1):1-15. doi:10.1038/s41598-020-69817-y
74. Wang Y, Yan F, Lu X, et al. IILS: Intelligent imaging layout system for automatic imaging report standardization and intra-interdisciplinary clinical workflow optimization. *EBioMedicine*. 2019;44:162-181. doi:10.1016/j.ebiom.2019.05.040
75. Paul R, Schabath M, Gillies R, Hall L, Goldgof D. Convolutional Neural Network ensembles for accurate lung nodule malignancy prediction 2 years in the future. *Computers in Biology and Medicine*. 2020;122(June):103882. doi:10.1016/j.compbiomed.2020.103882
76. El-Regaily SA, Salem MAM, Abdel Aziz MH, Roushdy MI. Multi-view Convolutional Neural Network for lung nodule false positive reduction. *Expert Systems with Applications*. 2020;162:113017. doi:10.1016/j.eswa.2019.113017
77. Jakimovski G, Davcev D. Using double convolution Neural Network for lung cancer stage detection. *Applied Sciences (Switzerland)*. 2019;9(3). doi:10.3390/app9030427
78. Lakshmanaprabu SK, Mohanty SN, Shankar K, Arunkumar N, Ramirez G. Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*. 2019;92:374-382. doi:10.1016/j.future.2018.10.009
79. Wang S, Wang R, Zhang S, et al. 3D convolutional neural network for differentiating pre-invasive lesions from invasive adenocarcinomas appearing as ground-glass nodules with diameters ≤3 cm using HRCT. *Quantitative Imaging in Medicine and Surgery*. 2018;8(5):491-499. doi:10.21037/qims.2018.06.03
80. Zhang C, Sun X, Dang K, et al. Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network. *The Oncologist*. 2019;24(9):1159-1165. doi:10.1634/theoncologist.2018-0908
81. Ahmed T, Parvin MstS, Haque MR, Uddin MS. Lung Cancer Detection Using CT Image Based on 3D Convolutional Neural Network. *Journal of Computer and Communications*. 2020;08(03):35-42. doi:10.4236/jcc.2020.83004
82. Harsono IW, Liawatimena S, Cenggoro TW. Lung nodule detection and classification from Thorax CT-scan using RetinaNet with transfer learning. *Journal of King Saud University - Computer and Information Sciences*. 2020;(xxxx). doi:10.1016/j.jksuci.2020.03.013

83. Nurtiyasari D, Rosadi D, Abdurakhman. The application of Wavelet Recurrent Neural Network for lung cancer classification. *Proceeding - 2017 3rd International Conference on Science and Technology-Computer, ICST 2017*. 2017;(July):127-130. doi:10.1109/ICSTC.2017.8011865
84. Tajbakhsh N, Suzuki K. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. *Pattern Recognition*. 2017;63(January 2016):476-486. doi:10.1016/j.patcog.2016.09.029
85. Gu Y, Lu X, Yang L, 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*. 2018;103(May):220-231. doi:10.1016/j.compbiomed.2018.10.011
86. Eun H, Kim D, Jung C, Kim C. Single-view 2D CNNs with fully automatic non-nodule categorization for false positive reduction in pulmonary nodule detection. *Computer Methods and Programs in Biomedicine*. 2018;165:215-224. doi:10.1016/j.cmpb.2018.08.012
87. Cao H, Liu H, Song E, et al. Multi-Branch Ensemble Learning Architecture Based on 3D CNN for False Positive Reduction in Lung Nodule Detection. *IEEE Access*. 2019;7:67380-67391. doi:10.1109/ACCESS.2019.2906116
88. Winkels M, Cohen TS. Pulmonary nodule detection in CT scans with equivariant CNNs. *Medical Image Analysis*. 2019;55:15-26. doi:10.1016/j.media.2019.03.010
89. Xie H, Yang D, Sun N, Chen Z, Zhang Y. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. *Pattern Recognition*. 2019;85:109-119. doi:10.1016/j.patcog.2018.07.031
90. Kim BC, Yoon JS, Choi JS, Suk H Il. Multi-scale gradual integration CNN for false positive reduction in pulmonary nodule detection. *Neural Networks*. 2019;115:1-10. doi:10.1016/j.neunet.2019.03.003
91. Li Y, Zhang L, Chen H, Yang N. Lung nodule detection with deep learning in 3D thoracic MR images. *IEEE Access*. 2019;7:37822-37832. doi:10.1109/ACCESS.2019.2905574
92. Zuo W, Zhou F, Li Z, Wang L. Multi-resolution cnn and knowledge transfer for candidate classification in lung nodule detection. *IEEE Access*. 2019;7(c):32510-32521. doi:10.1109/ACCESS.2019.2903587
93. Wang J, Wang J, Wen Y, et al. Pulmonary Nodule Detection in Volumetric Chest CT Scans Using CNNs-Based Nodule-Size-Adaptive Detection and Classification. *IEEE Access*. 2019;7:46033-46044. doi:10.1109/ACCESS.2019.2908195
94. Polat H, Mehr HD. Classification of pulmonary CT images by using hybrid 3D-deep convolutional neural network architecture. *Applied Sciences (Switzerland)*. 2019;9(5). doi:10.3390/app9050940
95. Pastorino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-Year results. *Lancet*. 2003;362(9384):593-597. doi:10.1016/S0140-6736(03)14188-8

Table 1

Table 1 is available in Supplementary Files section.

Figures

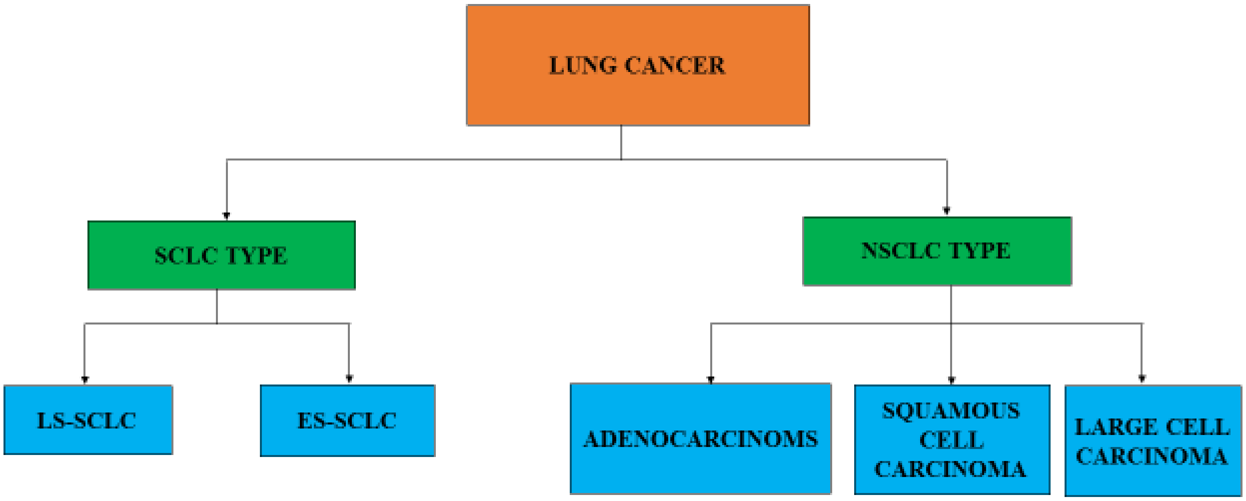

Figure 1

Types of lung cancer

Figure 2

NSCLC in the right bottom lobe squamous cell carcinoma

4

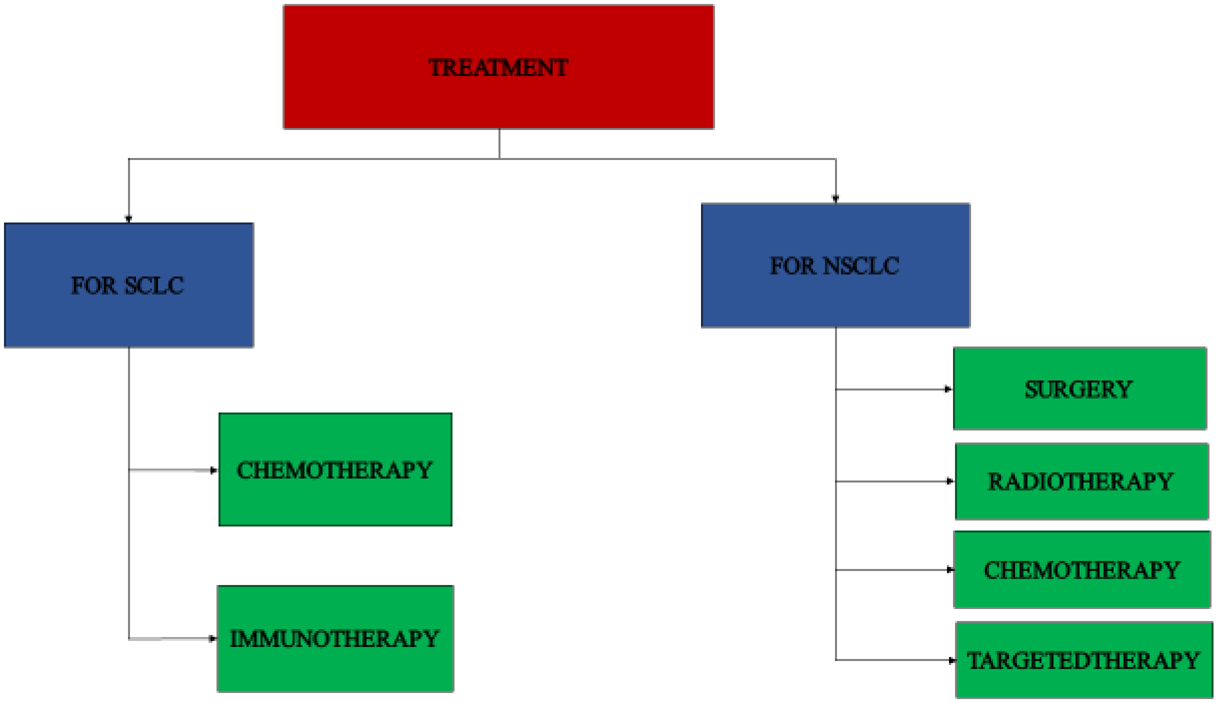

Figure 3

Type of treatment

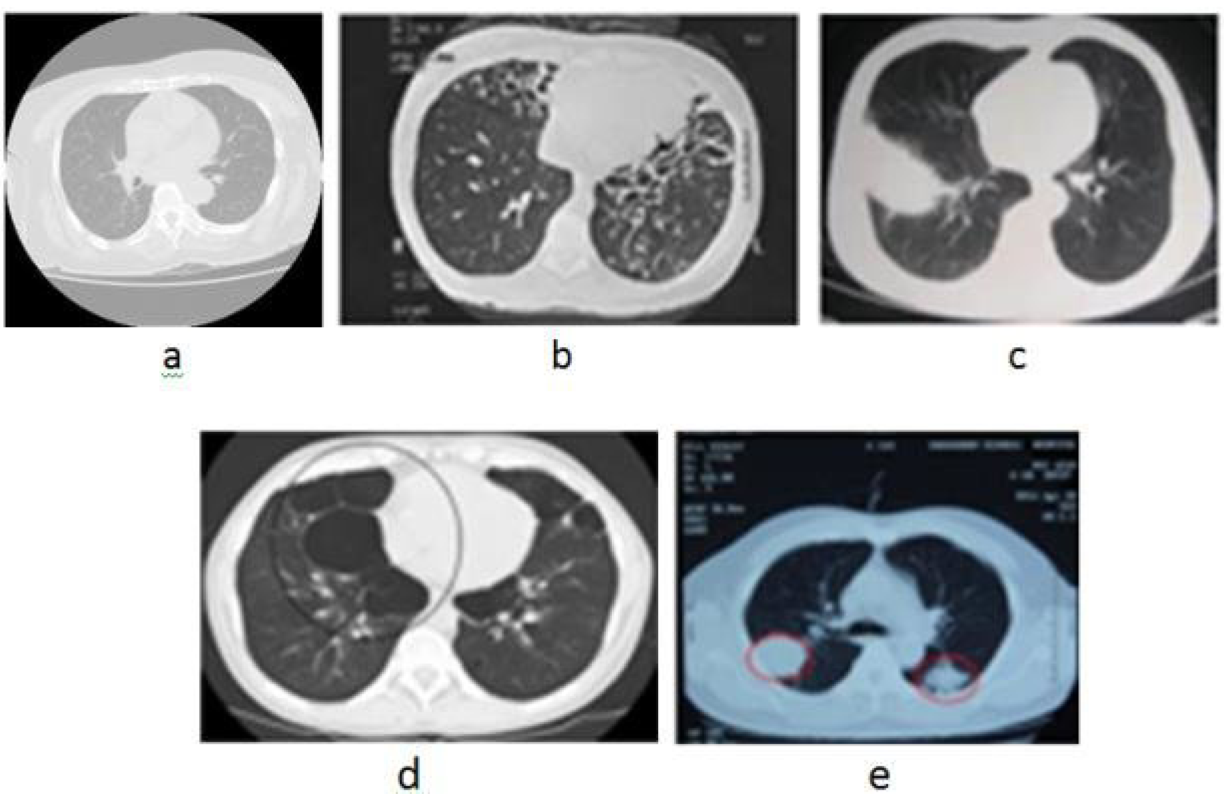

Figure 4

CT image of lung diagnosis (d), (e) and Lung normal patients (a), (b), (c)

41

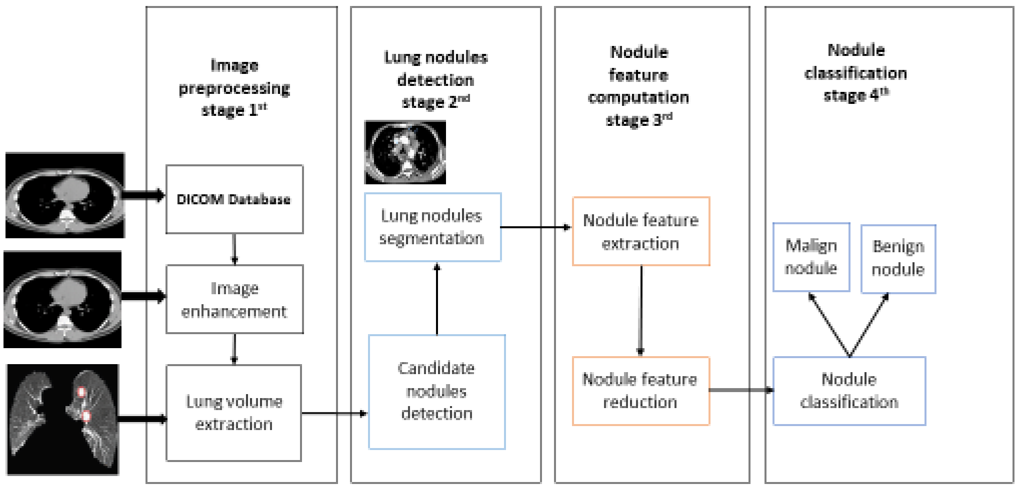

Figure 5

Pre-processing on DICOM database

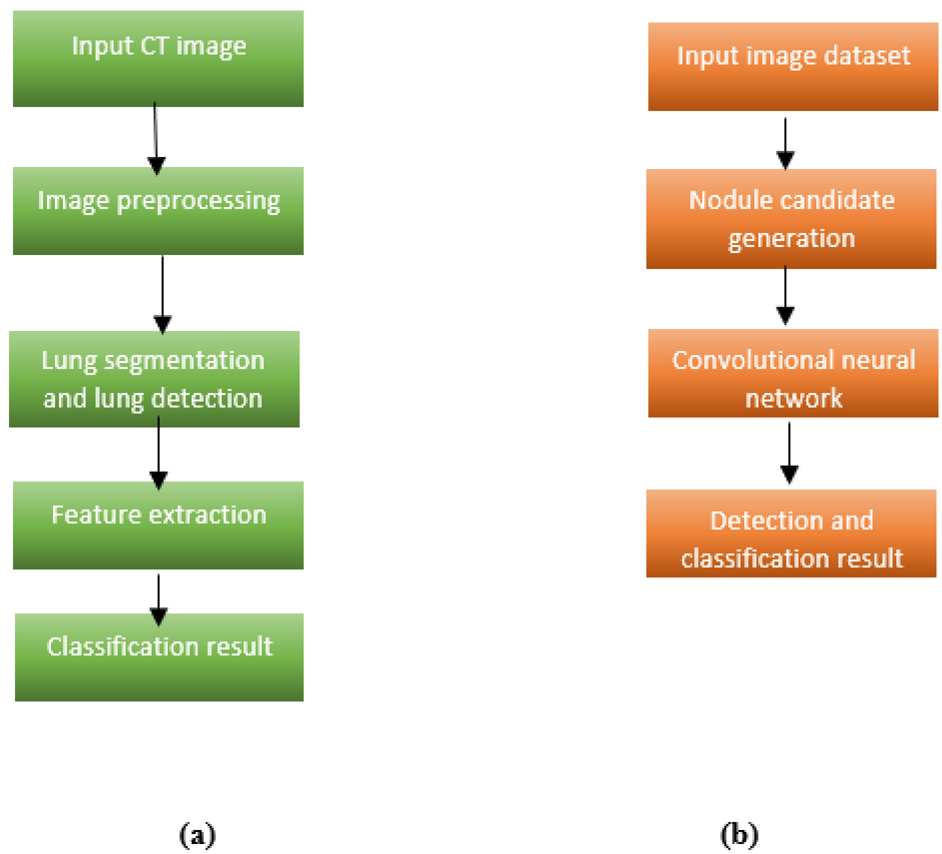

Figure 6

Lung nodule identification (a) traditional; and (b) CNN respectively

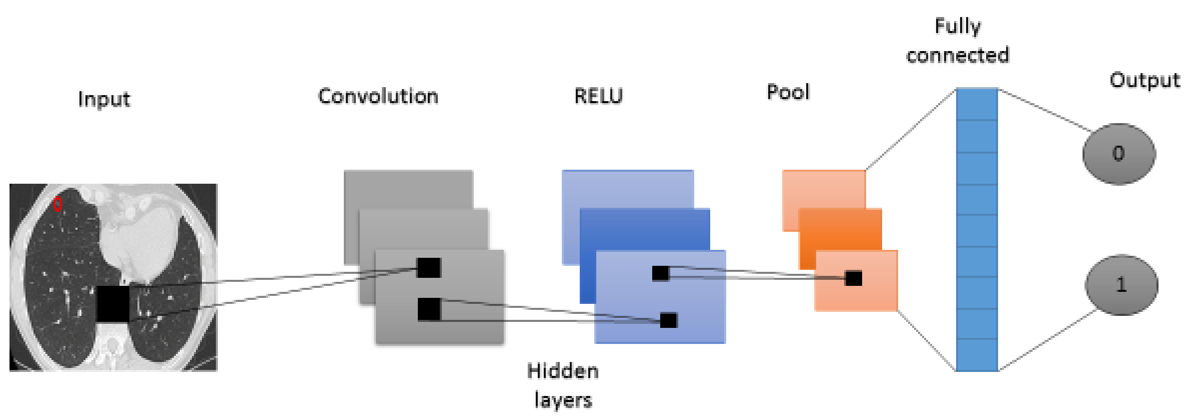

Figure 7

Structure of CNN using CT scan

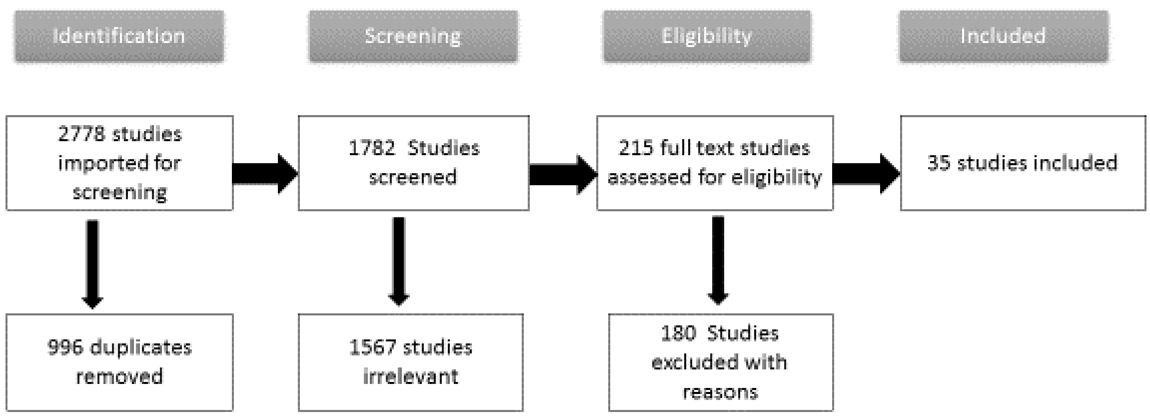

Figure 8

Flowchart of the literature and research collection of Preferred Reporting and Objects for Systematic Reviews and Meta-Analyses (PRISMA).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx