RESTORATION, ENHANCEMENT AND ANALYSIS OF LUNG NODULAR IMAGES FOR PROMPT DETECTION OF ABNORMALITIES

Thesis

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

SAVITHA G

DEPARTMENT OF MATHEMATICAL AND COMPUTATIONAL SCIENCES
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA,
SURATHKAL, MANGALORE - 575 025

MARCH 2020

Dedication

To The Memory of My Dear Mother

DECLARATION

By the Ph.D. Research Scholar

I hereby declare that the Research Thesis entitled **RESTORATION, ENHANCEMENT AND ANALYSIS OF LUNG NODULAR IMAGES FOR PROMPT DETECTION OF ABNORMALITIES** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy** in **Mathematical and Computational Sciences** is a *bonafide report of the research work carried out by me*. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

Place: NITK, Surathkal. (SAVITHA G.)

Date: 13-03-2020 148053-MA14F06

Department of Mathematical and Computational Sciences

CERTIFICATE

This is to *certify* that the Research Thesis entitled **RESTORATION, ENHANCEMENT AND ANALYSIS OF LUNG NODULAR IMAGES FOR PROMPT DETECTION OF ABNORMALITIES** submitted by **Ms. SAVITHA G**, (Register Number: 148053-MA14F06) as the record of the research work carried out by her is *accepted as the Research Thesis submission* in partial fulfillment of the requirements for the award of degree of **Doctor of Philosophy**.

Dr. JIDESH P
Research Guide

Dr. B R SHANKAR
Research Co-Guide

Chairman - DRPC
(Signature with Date and Seal)

ACKNOWLEDGEMENTS

I am profoundly grateful to my research advisors, Dr. Jidesh P. and Dr. B R Shankar for their invaluable and enlightening guidance throughout the course.

I thank the Research Progress Assessment Committee members, Dr. Sreejith A, School of Management, and Dr. Chandhini G, Dept. of MACS, National Institute of Technology Karnataka, for their insightful suggestions and encouragement.

I thank HOD, Prof. Shyam S. Kamath, Department of MACS, for all the support extended to me. I am grateful to NITK for providing me an opportunity for carrying out this research study.

I express my sincere gratitude to all the faculty and staff of MACS department for their help and cooperation.

I owe my deep debt of gratitude to my father Prof. S. G. Mayya and my brothers Shreesha and Girisha for their moral support and continuous encouragement.

It is my pleasure to thank all my friends and other research scholars who shared their valuable time with me.

Place: NITK, Surathkal SAVITHA G.

Date: 13-03-2020

ABSTRACT OF THE THESIS

Detection of lung cancer in the Computed Tomography (CT) images when the lung nodules are in the sub-solid state (early stage) results in higher survival rate of the patients. Two Computer Aided Detection (CAD) systems for identifying the sub-solid nodules in lung CT images are developed as a part of this thesis.

The first system adopts a pipeline approach which is carried-out in two phases. The first phase employs a series of algorithms for denoising, segmentation of region of interest and feature selection followed by a classification to separate nodules and non-nodules. In the second phase, Histogram of Gradients method is used to categorize the nodules identified in first phase as solid or sub-solid. Sensitivity of the system is observed to be more than 90% with just 3 false positive observations per scan. Both supervised and unsupervised classification models adopted for identifying sub-solid nodules give consistent and reliable results with an average accuracy above 93% when tested with Lung Image Database Consortium (LIDC) and International - Early Lung Cancer Action Program (I-ELCAP) databases. The accuracy of the system is categorically higher compared to the present state-of-the-art models employed for sub-solid nodule classification.

The second system adopts a deep learning approach for identifying sub-solid nodules, making use of a Deep Convolution Neural Network (DCNN) incorporated within the Conditional Random Field (CRF) framework. Adopting CRF framework reduces the occurrence of false positives. It is further observed that the overall accuracy of the system is increased from 83 to 89.5 percentage when tested with LIDC/IDRI and I-ELCAP databases. Though, the accuracy of the system is lower than the pipeline based model (mentioned above), the model does not demand any pre or post processing technique including the region of interest segmentation. The accuracy obtained for this system is comparatively higher than the state of the art deep learning models employed for sub-solid nodule classification. Moreover, a detailed cross comparative analysis of the systems proposed in this thesis is done to analyze their performance.

Keywords: Computer Aided Detection System, pulmonary nodule detection, sub-solid/part-solid nodule identification, Computed Tomography Images, Gray Level Co-variance Matrix, Deep Learning Convolution Neural Network, Conditional Random Field,

Table of Contents

List of Figures

List of Tables

List of Notations

U	Original image
U_0	Observed image
$U(x,y)$	Intensity of U at the location x and y
i,j	Intensity values(Gray value)
∇	Gradient operator
Δ	Laplacian matrix
$\|.\|$	Absolute value
$\prod$	Product
$\sum$	Summation
Ω	Image Domain
$\mathbb{R}$	Set of real numbers
$\forall x$	For all the values of x
∂	Partial derivative
Ip	Transmitted beam intensity
I_0	Beam intensity
T	Material Thickness
Q_0	Attenuation coefficient of the material
$P_{KL}(P,Q)$	KL divergence of distribution P and Q
$\nabla_{NL}u(x,y)$	Non-Local gradient for pair of pixels x and y
$\phi(.)$	Level set function
g_k	Gaussian kernal function
$d_1(x_i)$	Average intensity value inside the curve

$d_2(x_i)$	Average intensity value outside the curve
$Jaccard(R,S)$	Jaccard coefficient of two sets R and S
$d_J(R,S)$	Dissimilarity between two sets R and S
$prob(x,y)$	Probability of x and y
μ	Mean
σ	Standard Deviation
$*$	Convolution operator
$K(r,r')$	Radial basis function for data r and support vector r'
$\parallel x_i - y_j \parallel$	Euclidean distance between x_i and y_j
$\varphi(x_i)$	Unary potential energy
$\lambda(x_i,x_j)$	Pairwise potential energy
$E(U)$	Gibbs energy

List of Abbreviations

WHO	World Health Organisation
CT	Computed Tomography
AAH	Adenomatous Hypoplasia
MIA	Minimally Invasive Adenocarcinoma
$I-ELCAP$	International - Early Lung Cancer Action Program
VIA	Vision Image Analysis
$LIDC/IDRI$	Lung Image Database Consortium - Image Database Resource Initiative
CAD	Computer Aided Detection
GGN	Ground Glass Nodule
HoG	Histogram of Gradients
FLD	Fisher Linear Discriminant
$SPVA$	Segmentation-based Partial Volume Analysis
$3D$	Three-Dimensional
$MAGIC5$	Medical Applications in a Grid Infrastructure Connection 5
GP	Genetic Programming
LDA	Linear Discriminant Analysis
ROC	Receiver Operating Characteristic
$FROC$	Free-response Receiver Operating Characteristic
CNN	Convolutional Neural Network
$MCCNN$	Multi-crop Convolutional Neural Network
ROI	Region of Interest
PDF	Probability Density Function
KL	Kullback-Leibler
$NLTV$	Non-Local Total Variation Minimization
$PSNR$	Peak Signal to Noise Ratio

MSE	Mean Square Error
MAP	Maximum A Posteriori
AA	Aubert Ajol
BV	Bounded Variations
GLCM	Gray Level Covariance Matrix
PCA	Principal Component Analysis
SVM	Support Vector Machine
RF	Random Forest
RBF	Radial Basis Function
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
ROC	Receiver Operating Characteristics
AUC	Area Under Curve
OOB	Out-Of-Bag error
SIFT	Scale Invariant Feature Transform
SURF	Speeded Up Robust Features
DCNN	Deep Convolution Neural Network
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent
CRF	Conditional Random Field
MIoU	Mean Intersection over Union
PA	Pixel Accuracy

CHAPTER 1

Introduction

Diseases remain as a major cause of human fatality. Among them, cancer related deaths are more prominent and frequent. Cancer is a condition involving abnormal cell growth with the potential to invade or spread to other parts of the body. In general, all the body cells have a defined life span and are replaced by new cells from time-to-time. This process is called as Appoptosis (Su et al., 2015). Cancerous cells lack this capability where the cells fail to die but continue to divide uncontrollably. As a result, the cells pile up in the body which may form tumors or impair the immune system. These cancerous tumors may be in a particular area or spread to other parts of the body. There are various types of cancer such as bladder, colon and rectal, endometrial, kidney, leukemia etc. (Mohammad et al., 2015). Table 1.1 shows the statistics of most common cancer cases found in 2018 according to World Health Organization(WHO).

Table 1.1 Statistics of most common cancer cases

Cancer Type	No. of Cases	No. of Deaths
Lung Cancer	2.09 million	1.76 million
Breast Cancer	2.09 million	627000
Colo-rectal	1.80 million	86200

Lung cancer is the most common type of cancer that originates in lung cells. It is observed that number of deaths due to lung cancer is increasing and will remain so till the year 2030 (courtesy: World Health Organization 2019 report).

Lung cancer is referred to as primary lung cancer which manifests itself as tumors (nodules) in lungs. A pulmonary mass refers to the small round or oval-shaped growth in lungs. They may also appear as small spots in the lungs with diameter less than three centimeters. If the pulmonary mass has growth more than three centimeters then they are known as pulmonary nodules which are more likely to represent a cancer. The nodules are visible because of size however needs to be further analysed to conclude whether it is cancerous or not. Hence, they are easy to find but can be hard to diagnose. There are two main types of pulmonary nodules namely malignant (cancerous) and benign (noncancerous). Benign nodules are non invasive. They do not spread to the other parts of body unlike malignant tumor. They are not harmful and do not need any treatment. Over 90% of pulmonary nodules that are smaller than two centimeters (around 3/4 inch) in diameter are benign. These are non-cancerous nodules which are caused by previous infections or old surgery scars. If a nodule is found to be growing over the period of time it is termed as malignant which needs immediate treatment (Larici et al., 2017).

Malignant Pulmonary nodules are categorized into solid and sub-solid nodules depending on the degree of calcification (attenuation). The process of accumulation of calcium content on the soft tissue is termed as calcification. Solid nodules are characterized by concentrated soft tissue attenuation whereas sub-solid nodules are further classified into liquid or pure Ground Glass and part-solid nodules. Liquid nodules are also known as pure Ground Glass nodules. They occur as hazy areas with little attenuation and without clear boundary. Part-solid nodules are found to have attenuation more than liquid nodules but less than solid nodules. Figure 1.1 shows the categorization of pulmonary lung nodules. Visual representations of different kinds of nodules are shown in Figure 1.2.

According to Fleischner Society recommendation (MacMahon et al., 2005), lesions with diameter less than 4 mm do not require any further invasive or non-invasive measures or evaluations. However, lesions having a diameter of more than 8 mm should be evaluated with follow-up CT scans and biopsy since they are generally malignant.

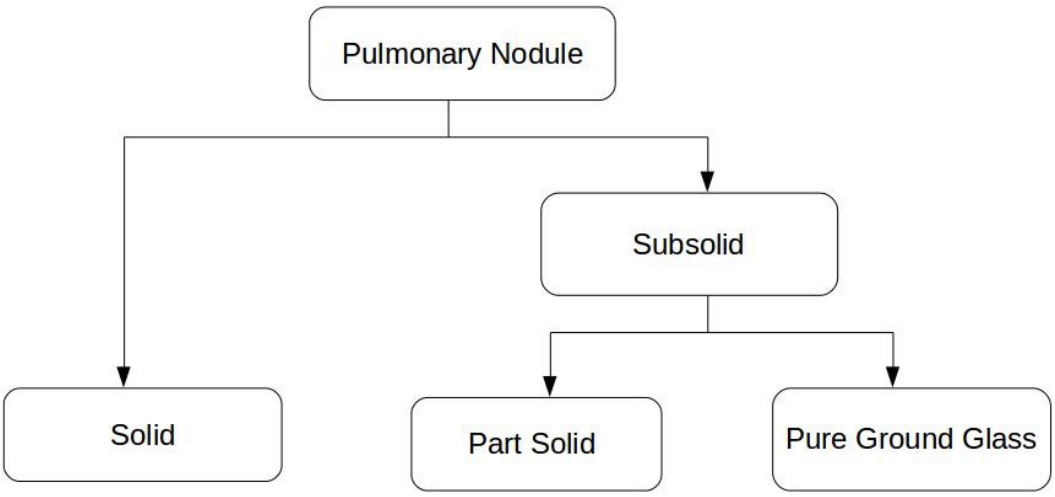
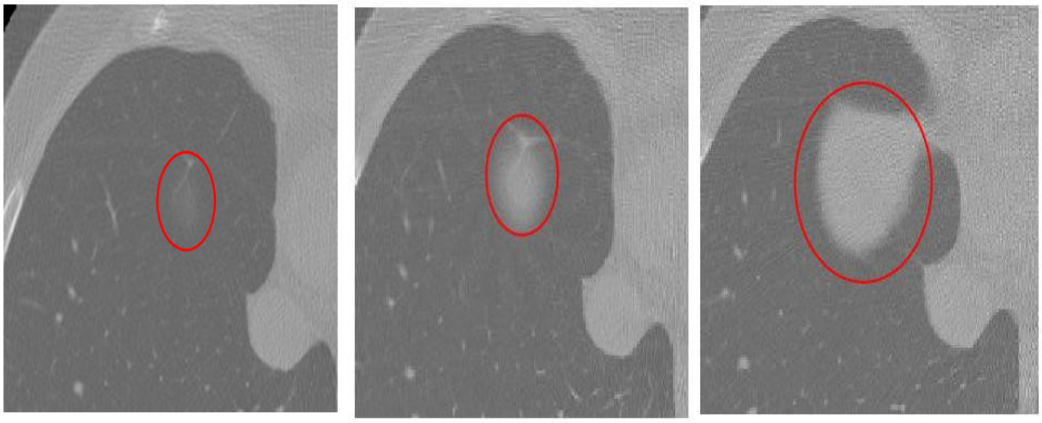


Figure 1.1 Nodule categorization



Liquid Nodule Sub-solid Nodule Solid Nodule

Figure 1.2 Nodules appearance in Lung CT Images (Marked in red circle)

For lesions having diameter between 4 mm and 8 mm, CT follow-up with a limited number of scans is recommended. Sub-solid nodules are often visible during screening studies or incidentally on CT scans. They are characterized by their appearance as rounded areas of lesions or calcifications. Sub-solid nodules can regress, persist, or grow. Small persistent non-solid nodules are often a typical adenomatous hypoplasia (AAH) and focal fibrosis, whereas non-solid and part-solid nodules that increase in size typically indicate malignancy, which can include adenocarcinoma, Minimally Invasive Adenocarcinoma (MIA), and invasive adenocarcinoma disease. A small percentage of ground-glass opacities (AAH) develop into larger lesions that can eventually comprise

on solid components and invades to alveolar or walls of lungs indicating the advanced stage of lung cancer (MacMahon et al., 2005).

American cancer society in the year 2016 conducted a survey and found that lung cancer is more severe and common. The details of this study is summarized in Table 1.2. It is the leading cause of cancer deaths in the world.

Table 1.2 Study of survival rate for lung cancer (American Society, 2016)

Different Stages	Nodules present	5 year Survival Rate
Local	Liquid, Sub-solid	52
Regional	Solid	17
Distant	Other parts	4

It is reported that 5 year survival rate for all the stages combined is only 16 % while for cases detected in the initial stage it is about 52 %. But only 15 % of lung cancers are diagnosed at this early stage (Society, 2016). Treatment in advanced (solid nodules) stage may require the removal of entire lungs which patients cannot withstand. It is always advantageous to identify the tumors or nodules in their initial stage when they manifests themselves as pulmonary (sub-solid) nodules. Therefore, early detection of lung cancer is important since patients survival can be extended by appropriate treatment. Liquid biopsy technique for lung cancer detection is used for identifying the lung cancer which is a painful procedure. Hence, there is a need for identification of lung cancer from obtained Computed Tomography (CT) images without causing any troubles to the patients. Also, there is a requirement for early detection of lung cancer to increase the chance of survival of the patients. Two automated Computer Aided Detection (CAD) systems for identifying the lung nodules in early stage (part solid or sub-solid nodules) from lung CT images are developed as a part this thesis. The common imaging modality used to capture lung nodules is CT scans. The procedure of formation of CT images is described in detail as follows.

1.1 Formation of Computed Tomography Images

Lung Nodules are identified from the captured lung Computed Tomography images. High-resolution images that depicts the composition of body tissues are produced by CT imaging. Individual views or x-ray images are collected from around the body. A single slice is formed by using the collected views. Images are formed by the interaction of high energy photons and body tissues. The contrast in a CT image is created by identifying the attenuation coefficient of the material. Attenuation coefficient refers to the absorption properties of the elements of body tissue. The Beer-Lambert law describes the attenuation of an individual x-ray passing through the body and is given as,

$$I = I_0 e^{-Q_0 T}, \tag{1.1.1}$$

where, I_0 is the beam intensity, I is the transmitted beam intensity, T is the material thickness, and Q_0 is the attenuation coefficient of the material.

CT image formation has three phases namely Scanning phase, Reconstruction phase and Digital to Analog conversion phase (Bracco et al., 2019). The formation of CT images in three phases is shown in Figure 1.3.

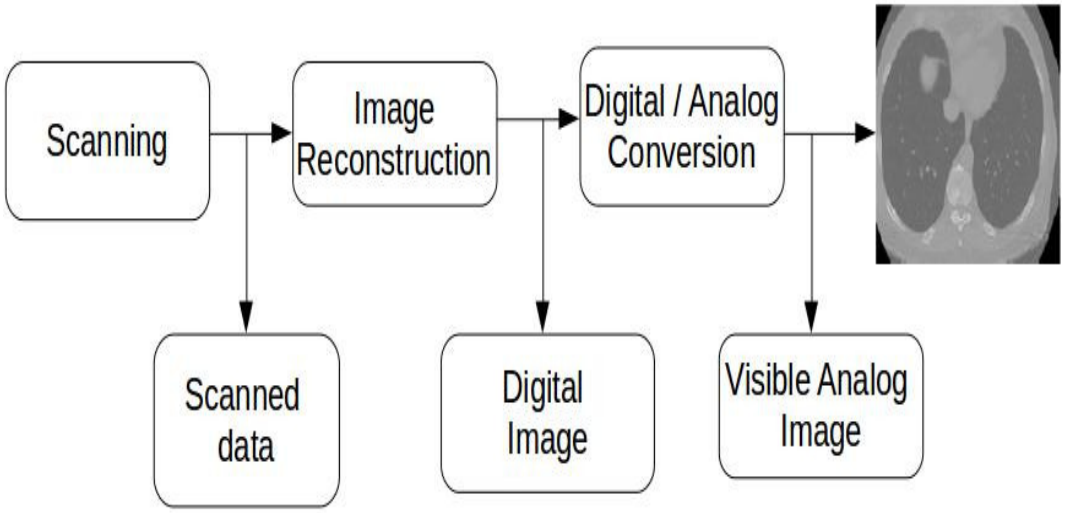


Figure 1.3 Process of CT Image Formation

The scanning phase identifies the data which is not in the image format. A fan shaped x-ray beam is scanned around the body. Detectors are used for identifying the amount of x-ray passing through the body. Projection of x-ray from one specific focal point position produces one view. Many views captured by projecting the x-ray beam from different positions are necessary to reconstruct an image. Many projections spanning 180 degrees of the object make an entire set of slice. This captured data comprises of total attenuation of the x-rays passing all through the body (Willemink and Noël, 2019). One scan produces data for one slice of image. Once this entire slice is collected in different views, it is displayed in the sinogram which depicts the relation between each projection and its sinogram. Sinogram is an array of captured projections. Projection angle and radial distance from the origin are the two parameters used for construction of sinogram as shown in Figure 1.4 where three projections in the angles 0, 45 and 90 degrees are captured. Spatial resolution and circumferential resolution are determined by number of rays and number of projections, respectively. A horizontal line is plotted through the sinogram for each object view. Each point depicts the brightness corresponding to the attenuation level experienced by the ray when passing through the body.

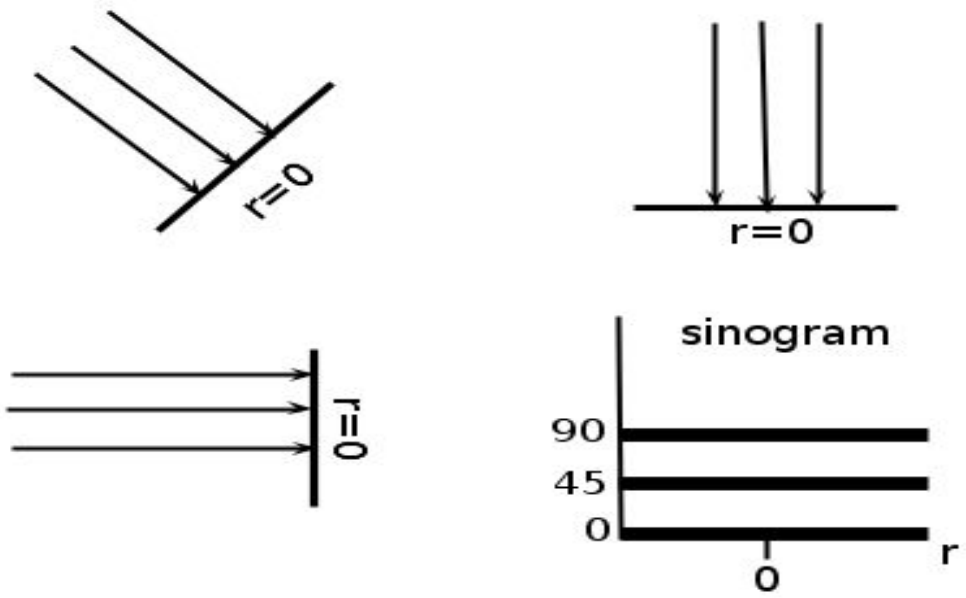


Figure 1.4 Formulation of Sinogram from individual projection views

The acquired data is processed in the reconstruction phase to form a digital image. The reconstruction phase may be repeated for producing images with different slice thickness and position of image slices. Single or multiple detectors determines the slice thickness. The width of the collimator detector decides the size of the slice thickness in single detectors whereas width of slices can be adjusted by binning adjacent detectors together in the case of multiple detectors.

Digital CT images are produced by the three popular methods such as Algebraic method, Fourier Transform method, Iterative and Back-projection method. Generally, reconstruction of medical CT images are achieved by using filtered back projection technique (Hoffman et al., 2016). If a particular position has a high intensity value in a CT view, then intensity values are assigned to all the values along that position. When the process is repeated with different angles, the signal intensities intersection point represents the true location of the point of interest as shown in Figure 1.5. To nullify the effect of blurring caused in this process, filters such as convolution, ramp etc. are used in the spatial domain since frequency domain operation require the Fourier transform equations. The detailed procedure is given below,

- Transforming each projection into frequency space.

$$\mathscr{F}_{1D}\{p(r,\theta)\} = P(\omega,\theta) \tag{1.1.2}$$

- Multiplying the filter

$$P(\omega,\theta) * |\omega| = F(\omega,\theta) \tag{1.1.3}$$

- Transforming each filtered projection into image space.

$$\mathscr{F}_{1D}^{-1}\{F(\omega,\theta)\} = f(r,\theta) \tag{1.1.4}$$

- Back-projecting the filtered projections to obtain the final image.

$$f(r,\theta) \xrightarrow{backprojection} U(i,j) \tag{1.1.5}$$

where, $|\omega|$ is the ramp filter, $p(r,\theta)$ is the projection with the radiation intensity r made to pass at an angle θ. $\mathscr{F}_{1D}$ represent the 1 Dimensional Fourier transform.

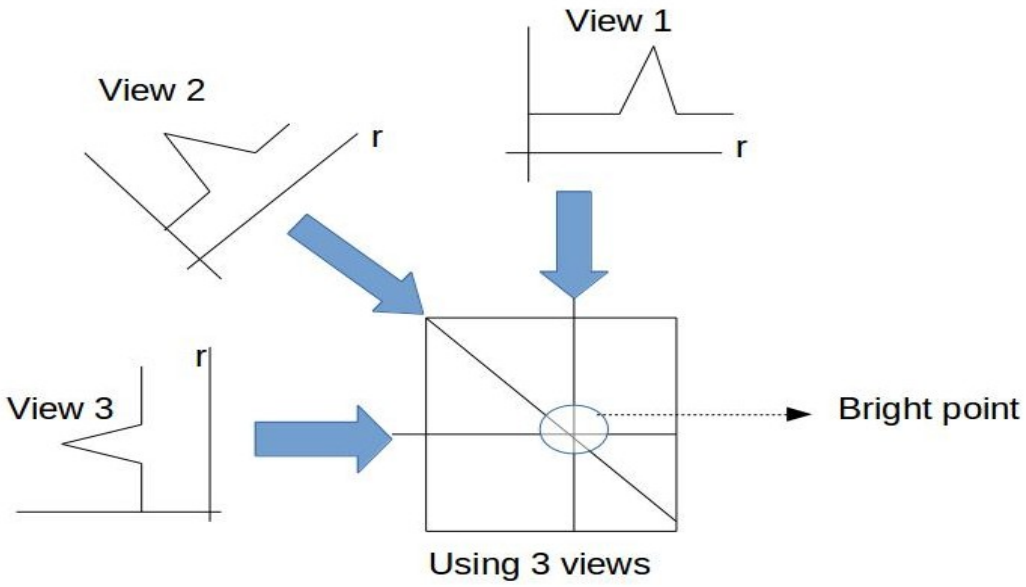


Figure 1.5 CT Image Reconstruction using Filtered Back-projection method

Image quality is improved or the image quality characteristics such as detail, texture etc. are improvised using the image processing methods in the filtered back projection approach. The data produced in this step is not a complete image but a profile of the x-ray attenuation by the objects. Back-projection of the profile on to an image surface

produces the digital image. Many views are used to produce CT images. The CT number is calculated which helps in reconstructing the digital image.

X-ray attenuation coefficient is computed in the reconstruction process which in-turn is used to calculate the CT number values. These CT numbers are computed for every individual tissue voxel. It is found and stated that CT number of air is -1000 and water is 0 (Willemink and Noël, 2019).

$$\text{CT number} = \frac{V_{voxel} - V_{water}}{V_{water}} \times 1000, \tag{1.1.6}$$

where, V_{voxel} is the attenuation coefficient of tissue, V_{water} is the attenuation coefficient of water.

CT numbers are measured in Hounsfield Units. Tissues with attenuation (density) greater than water will have positive CT numbers whereas those which are less dense have negative CT numbers. Density, atomic number of tissues and energy of the x-ray photons decide the x-ray attenuation coefficient. A high Voltage (like 120-140kV) and heavy beam filtration is used for CT images. This high voltage minimizes the photoelectric interactions in the tissue. Hence CT numbers are computed based on density of tissues.

The last digital-to analog conversion phase produces the visible image in shades of gray. The digital images made up of pixels and having a CT number are converted to visible image with different shades of gray using the windowing technique. The windowing technique transforms CT numbers into gray scale values ([0,255]). There are other adjustable factors like level and width controls which will have effects on the quality of reconstructed images. The tissues within the window will have different shades of gray i.e. they have different CT numbers. All the tissues with CT number above the window will be represented as white within the window range and the ones having value below the window will be in black color within the window range. Level control is used to adjust the center of window and the width control adjusts the range of CT numbers to be displayed in the reconstructed image. Hence reducing the window width will increase the image contrast in tissues. Window level for imaging the lungs is observed to be -500 (Heverhagen, 2016).

1.2 Dataset

The data for the present study is collected from International - Early Lung Cancer Action Program (I-ELCAP) database and Lung Image Database Consortium - Image Database Resource Initiative (LIDC/IDRI) database(Armato III et al., 2011).

International - Early Lung Cancer Action Program (I-ELCAP) database is developed by Vision Image Analysis (VIA). This database contains 500 low-dose documented lung CT scans which are obtained in a single breath hold with slice thickness of 1.25 mm produced using x-ray.

Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands developed the LIDC / IDRI database. 888 CT scans of 124 GB are available in LIDC-IDRI database. Slice thickness greater than 2.5 mm is considered while others are discarded. It contains annotations that are collected during two phase annotation process by radiologists. In this process three types of nodules are marked namely, nodule < 3 mm, nodules > 3 mm and nodules of 3 mm. The nodule exceeding 3 mm are considered as malignant nodules by radiologists.

The CT images thus obtained are processed for identifying the part-solid/sub-solid nodules. These obtained CT images are considered as raw data for further process. They contain unwanted information (noise) which appears due to the defects in the capturing environment, artifacts in the instruments etc. The Digital Image Processing techniques are adopted to separate the sub-solid nodule region from the rest of the image portion. This includes preprocessing techniques for removal of noise followed by identification of nodules through segmentation and classification process.

1.3 Organization and Contribution of the Thesis

Considering the severity of lung cancer and the lack of an efficient CAD systems for identifying it in early stages, two comprehensive automated CAD systems are developed as a part of this thesis work. These systems help in identifying the lung cancer in its early stages, when it manifests itself as part-solid/sub-solid nodules in lung CT scans.

The remaining parts of the thesis are organized as follows. Chapter 2 summarizes the state-of-the-art models available till date along with identified research gaps in these proposals and a brief overview of the steps taken to overcome the gaps. This chapter includes the details of medical aspects of lung cancer and the way the radiologists identify it till date. This chapter also summarizes the CAD system available till date that are mainly oriented towards identifying the solid nodules which are clearly visible in lung CT scans.

The proposed pipeline model for early detection of lung cancer is described in Chapter 3. Various algorithms are incorporated in series for the purpose of identification of part-solid/ sub-solid nodules. The model developed has two phases. The first phase is used for segregating the nodule and non-nodules followed by identification of solid and sub-solid nodules in the second phase. The preprocessing methods which includes identification and removal of noise, segmentation, feature identification and classification into nodules and non nodules comprise the first phase (Phase I). The features that are specific to sub-solid nodule are identified using Histogram of Gradients (HoG) method, followed by classification into solid and part-solid/sub-solid nodules in the second phase (phase II) of the developed CAD model.

The proposed deep learning approach for early detection of lung cancer is detailed in Chapter 4 where a series of convolution neural network layers are incorporated. Back propagation algorithm is used which reduces the error in the developed model.

Finally, the summary of the entire work, its applications followed by future research avenues are highlighted in Chapter 5.

CHAPTER 2

Literature Review

The present study envisages to develop an automated Computer Aided Detection (CAD) system for early detection of lung cancer or carcinoma. Sub-solid nodules signifies the initial stage of lung carcinoma. Though there are some efforts done to identify solid nodules in CT images (Murphy et al. (2009), Messay et al. (2010), Magdy et al. (2015)), not many published works are available in the literature dedicated to identify the presence of sub-solid lung nodules. Therefore a detailed review of the existing models are performed in this chapter.

The authors in Hansell et al. (2008) have established that part-solid nodules consist both ground-glass and solid components and stated that nodules identify themselves as well or poorly defined, irregular opacities which are 3 cms in diameter. They further distinguished the nodules into three types namely liquid, sub-solid/part solid and solid according to their growth and severity. They describe the structure and appearance of nodules which eases the identification of nodules and its growth rate. A homogeneous soft-tissue attenuation is termed as solid nodules, the one which consists Ground Glass Nodules and soft tissue attenuation is called as the sub-solid nodule and hazy structures with only Ground Glass attenuation is termed as liquid nodules.

The findings of Henschke et al. (2002) have revealed that sub-solid/part-solid nodules are less common in screenings which makes their identification tedious during the initial stages of lung carcinoma. Their experiment shows that 81 % of solid nodules and 19 % of sub-solid nodules are found in the screening. It is also established that 50 % of lung cancer originates from sub-solid nodules and hence identification of sub-solid / part-solid nodules play an important role in identifying the lung cancer in its early

stage.

Erasmus et al. (2000) have established that lung cancer identifies itself as pulmonary nodules in the initial stage. These nodules are usually small round shaped spots present in the lungs. They are mainly categorized into two groups based on their size. As already mentioned earlier, the nodules smaller in size having a diameter less than 3 centimeter, in general, are grouped as Benign Pulmonary nodules and usually evolve due to inflammation, injuries, etc. They are non-cancerous nodules remaining as calcified lesions forever. If the pulmonary nodules detected are more than 3 centimeters in diameter, they are termed as malignant nodules and treated as cancerous.

A CAD system is developed by Tan et al. (2011) for the detection of solid lung nodules in CT images. A mixture of features and classification algorithms are adopted therein. Segmentation of nodules is carried out with the help of nodule and vessel enhancement filters followed by computing the divergence features to locate the clusters of nodules. Classification algorithms differentiate between real nodules and some form of blood vessels by using the features identified which are defined on the gauge system.

Vlahos et al. (2018) observed that the nodules present in central locations are small and difficult to be detected when compared to nodules attached to the vessels or pleural surfaces of lungs. It is also established that central nodules are indications to early stage cancer. Small and not well developed nodules indicate the early stage of lung cancer. They are generally sub-solid or liquid nodules which represent initial stage of lung cancer. Identifying them helps in detecting lung carcinoma in the initial stage. Hence, nodules present in the center locations are to be attended rather than the ones present near the walls of lungs as they represent advanced stage of lung cancer.

Messay et al. (2010) developed a CAD system for the detection of pulmonary nodules in thoracic computed tomography imagery in the early stages. Segmentation of entire lung region is achieved and the boundaries are marked clearly. For the purpose, it combines intensity thresholding and morphological processing to detect the lung region and subsequently the lung nodules. Features are extracted and classified using Fisher Linear Discriminant (FLD) classifier and a quadratic classifier. This system is found to give 80% accuracy with an average of 517.5 nodule candidates per

case/scan (517.5 ± 72.9). A 7-fold cross-validation performance analysis using the LIDC database only shows CAD sensitivity of 82.66% with an average of 3 False Positives (FP) per CT scan/case.

Murphy et al. (2009) developed a scheme for the automatic detection of nodules in thoracic CT scans. The algorithm uses the local image features of shape index and curvedness in order to detect candidate structures in the lung volume and applies two successive k-nearest-neighbor classifiers for the reduction of false-positives. The nodule detection system is trained and tested on LIDC/IDRI database and achieves a sensitivity of 80% with an average of 4.2 false-positives per scan.

Jacobs et al. (2011) established that Ground Glass Nodules (GGNs) occur less frequent in CT scans than solid nodules but have a much higher chance of being malignant. They have developed a complete system for computer-aided detection of GGNs which incorporates the segmentation steps, candidate detection, feature extraction and a two-stage classification process. Features extracted are a set of intensity, shape and context features. Two-stage classification approach using a linear discriminant classifier and a Gentle-Boost classifier are adopted to classify the nodule regions. The system is trained and independently tested on 140 scans that contains one or more GGNs from around 10,000 scans obtained in a lung cancer screening trial. The system shows a high sensitivity of 73% with one false positive per scan.

An early detection of Ground Glass Nodule in lung CT images is developed by Tao et al. (2009). A multi-level learning-based framework for automatic detection and segmentation of GGN in lung CT images is developed. This model integrates segmentation and detection to improve the overall accuracy for GGN detection followed by two levels of classification namely voxel-level and object-level. The method is demonstrated on a dataset of 1100 subvolumes (100 containing GGNs) extracted from about 200 subjects and is found to give 80% accuracy.

Magdy et al. (2015) proposed a computer-aided diagnostic (CAD) systems for identifying and classifying the normal and cancer lungs which in-turn concentrates on identifying solid nodules since they are clearly visible. For this, Weiner filtering followed by histogram analysis with thresholding and morphological operations to segment and ex-

tract the lung region is employed. Amplitude-Modulation and Frequency-Modulation (AM-FM) is adopted to extract the features of the region of interest. Classifiers like SVM, K-Means, Naive Bayes and some other linear classifiers are used for differentiating the normal and cancerous lungs. It is established that 95% accuracy is achieved when linear classifiers are adopted.

Kuhnigk et al. (2006) have done volumetric growth assessment of pulmonary lesions crucial to both lung cancer screening and oncological therapy monitoring. This is an automated segmentation method that is based on morphological processing and is suitable for both small and large lesions. It also addresses clinical challenges to volume assessment such as variations in imaging protocol by introducing a method of segmentation-based partial volume analysis (SPVA) that follows the segmentation procedure. Both systematic and absolute errors were shown to be reduced substantially by the SPVA method. The method is especially successful in accounting for slice thickness and reconstruction kernel variations, where the median error is more than halved in comparison to the conventional approach.

De Nunzio et al. (2011) developed a fully automated and three-dimensional (3D) segmentation method for identification of the pulmonary parenchyma in thorax X-ray CT datasets. This is mainly designed as the preprocessing step in the CAD system for malignant lung nodule detection that is being developed by the Medical Applications in a Grid Infrastructure Connection (MAGIC-5) Project. Segmentation of the external airways (trachea and bronchi), is obtained by 3D region growing with wavefront simulation and suitable stop conditions, thus allowing an accurate handling of the lung region. This algorithm is tested with LIDC / IDRI database and has produced segmentation accuracy of 96% with incorrect segmentation of external airways in the remaining cases.

Choi and Choi (2012) developed an automated pulmonary nodule detection system which can assist radiologists in detecting lung abnormalities at an early stage. They have proposed a novel pulmonary nodule detection system based on a genetic programming (GP) based classifier. The system works in three steps with segmentation by thresholding method in the first step followed by optimal multiple thresholding and

rule based pruning to detect and segment nodule candidates in the second step. Finally, genetic programming based classifier is trained and used to classify nodules and non-nodules. The performance of the system is evaluated using the Lung Image Database Consortium database. It is found that this model reduces the number of false positives in the candidate nodules, achieving a 94.1% sensitivity at 5.45 false positive rates per scan.

A Computer-Aided Diagnosis system is developed by Suárez-Cuenca et al. (2009) to detect pulmonary nodules on thin-slice helical computed tomography images. The capability of an iris filter to discriminate between nodules and false-positive findings is verified. Suspicious regions are characterized with features based on the iris filter output, gray level and morphological features, extracted from the CT images. Functions based on linear discriminant analysis (LDA) are used to reduce the number of false-positives. The system is evaluated on CT scans containing 77 pulmonary nodules. Results for a test set, evaluated with free-response receiver operating characteristic (FROC) analysis has yielded a sensitivity of 80% at 7.7 false positive rate per scan.

Shen et al. (2017) investigated the problem of lung nodule malignancy suspiciousness (the likelihood of nodule malignancy) classification using thoracic CT images. Their system has directly modeled raw nodule patches with an end-to-end machine learning architecture for classifying lung nodule malignancy suspiciousness. A Multi-crop Convolution Neural Network (MC-CNN) is developed to automatically extract nodule salient information by employing a novel multi-crop pooling strategy which crops different regions from convolution feature maps and then applies max-pooling at different times. Accuracy of the system developed is found to be 70%.

A Computer-Aided Detection model is proposed by Setio et al. (2016) for pulmonary nodules using multi-view convolution networks (ConvNets), for which discriminative features are learned automatically from the training data. A collection of solid, sub-solid and liquid nodules marked in the datasets are given as input to the network in training phase. The proposed architecture comprises multiple streams of 2-D ConvNets, for which the outputs are combined using a dedicated fusion method to get the final classification. Cropped regions of nodules are given as input in the training phase

and testing phase. Hence this system is limited in capacity to classify the nodules into solid, sub-solid and liquid.

(Zhou et al., 2006) proposed a classifier for Ground Glass Opacity detection in lung CT images. They have proposed a model for segmentation of Ground Glass Nodules (sub-solid nodules) by adopting the K-means classifier. The classifier performance is boosted by measuring the euclidean distance between the nonparametric density estimates of two regions. The selected regions are segmented by analyzing the 3D texture likelihood map in the region.

Research has been carried out in segmenting and classification of nodules using popular contour detection algorithms by Magdy et al. (2015); De Nunzio et al. (2011); Van Rikxoort et al. (2009) and Choi and Choi (2012). There have been some works in identifying the lung nodules by adopting neural networks or deep learning methods. Some of the identified works in this field are summarized here. Nithila and Kumar (2017) have developed a model for identifying the solitary nodules in given CT lung images, using a swarm intelligence optimized neural network approach. Here lesions are detected by applying region-based contour segmentation method. The output of segmentation is given to back-propagation neural network which identifies the solitary nodules. A multilevel convolution neural network is proposed by Setio et al. (2016) to identify the features of nodules. However, the training and execution time taken in these models are more and it also requires the knowledge of the nodule's center location in advance. This in turn relies on the manual segmentation process.

A multi-scale Convolution Neural Network (CNN) model has been developed by Shen et al. (2015) only for feature extraction. Nodule patches are cropped and given as input for the CNN model. Features extracted by CNN is combined with a state of art method; SVM classifier for identifying the nodules. Though, this method gives an accuracy of 85 %, the requirement of cropped nodule region as an input is a major limitation of the model. Van Ginneken et al. (2015) have developed a model for detecting nodules in CT scans. An existing CAD system developed by MeVis Medical Solutions AG, Bremen, Germany has been considered to find the location of a nodule in lungs. CNN features obtained from identified nodules are used in detecting the benign and

malignant nodule.

A model has been developed by Kumar et al. (2015) which identifies lung nodule based on deep features extracted from auto-encoder. The features are then classified using a binary tree classifier. The overall accuracy of the model is found to be 75 %. However, it does not specify the nodule type(benign or malignant), which is the main limitation of the model.

Song et al. (2017) developed a deep learning model for the purpose of classification of nodules. Deep features are learned by the model to distinguish between the benign and malignant nodules. This model works with an accuracy of 65% but the model takes more time to learn the features in the training phase which turns out to be a major limitation. A hybrid CAD system is proposed by Teramoto et al. (2016) for identifying the nodule region by adopting an active contour filter. False-positive images obtained are reduced using an ensemble method by extracting shape features and features obtained from Convolution Neural Network. Rule-based and SVM classifiers are used in the final stage to classify the region as a nodule or a non-nodule.

Various existing deep convolution neural network architectures are studied and evaluated for thoraco-abdominal lymph node detection and interstitial lung disease classification by Hoo-Chang et al. (2016). They have tried to use the existing models like U-net, Segnet, Google-net etc.. It is observed that the average accuracy of all the models is close to 60% in identifying the various lung interstitial diseases. The authors Alam et al. (2019) developed a convolution neural network model for hyper-spectral image classification where both spatial and spectral information along with conditional random field algorithm are considered. They observed that for hyper-spectral images adopting conditional random field algorithm produces much better results with less false predictions. This algorithm tries to eliminate the false positives and false negatives by adopting the graph partition technique which in-turn is explained in detail by Boykov and Jolly (2001). Boykov and others developed an algorithm for segmenting the objects or regions in an image by using the optimal graph cut method.

It has been observed from the literature that the existing systems are confined to one task such as segmentation, feature selection etc.. However, there is a lack of an end to

end automated CAD system for identifying the part-solid nodules in lung CT images.

Table 2.1 highlights the snap-short of various systems available in the literature till date and their salient features. The pros and cons of each model have been briefly described here. Moreover, the research gaps are identified and scopes for improvement are analyzed from this comparative study. The next chapter introduces a CAD system which addresses some of the setbacks of the existing models.

Table 2.1 A bird's eye view of the existing CAD systems

Sl. No.	CAD systems	Aim	Method adopted	Dataset	Performance	Limitations
1	Tan et al. (2011)	solid nodule identification	nodule enhancement filters, divergence features	125 CT scans: LIDC	sensitivity of 87.5% at 4 FP/scan	advance stage cancer identification
2	Messay et al. (2010)	sub-solid nodule identification	thresholding and morphological operation	84 cases: LIDC	sensitivity of 82.6% at 3 FP/scan	More anatomical structures are detected increasing the error
3	Murphy et al. (2009)	sub-solid nodule identification	shape index & curvedness features	800 CT scans: LIDC	Sensitivity of 80% 4.2 FP/scan	vessel structure or fissures falsely identified as nodule
4	Jacobs et al. (2011)	sub-solid nodule	intensity, texture feature selection, Gentle-Boost classifier	140 scans of LIDC database	sensitivity of 73% at 1 FP/scan	cropped nodule region is given as input
5	Magdy et al. (2015)	normal lung & cancer lung identification	Weiner filter & histogram analysis SVM, K-Means, Linear classifier	166 scans of LIDC database	sensitivity of 94%	does not identify the individual nodules

Table 2.1 continued

6	Kuhnigk et al. (2006)	segmentation of solid pulmonary lesions	segmentation-based partial volume analysis (SPVA)	selected scans of LIDC database	Segmentation accuracy of 91.4%	identifies only solid lesions
7	Choi and Choi (2012)	pulmonary nodule detection system	genetic programming based classifier	84 CT scans of LIDC database	sensitivity of 94.1% at 5.45 FP/scan	limited to nodule & non-nodule identification
8	Suárez-Cuenca et al. (2009)	CAD system to detect pulmonary nodules	iris filter is used	77 CT scans LIDC	sensitivity of 80% at 7.7 FP/scan	identifies only the solid nodule
9	(Zhou et al., 2006)	classifier for Ground Glass Opacity detection	thresholding by the intensity histogram analysis	600 scans of LIDC database	error rate of 3.7%	concentrates on segmenting nodule regions
10	De Nunzio et al. (2011)	pulmonary parenchyma identification	Three-Dimensional (3D) segmentation method	ANODE09 images- 55 scans LIDC database- 84 scans	0.96 of mean overlap degree	concentrates only on segmentation of lungs & solid nodules
11	Shen et al. (2017)	lung nodule malignancy identification	multi-crop Convolution Neural Network	1010 scans of LIDC database	classification accuracy of 85%	Input: cropped nodule patches

Table 2.1 continued

12	Setio et al. (2016)	multi-view convolution networks (ConvNets)	discriminative features from training set	888 scans of LIDC database	sensitivity of 85.4% at 5 FP/scan	only classification of soild & sub-solid
13	Nithila and Kumar (2017)	solitary nodules identification	swarm intelligence optimized neural network	LIDC-IDRI & SPIE-AAPM database	accuracy: 98% for solid nodules, 97% for part-solid, 97% for liquid nodule	knowledge of the nodule's center location in advance
14	Song et al. (2017)	classification of nodules as benign & malignant	Deep learning architecture	LUNA16 dataset & DSB dataset	Accuracy: 65%	more time needed feature learning phase
15	Teramoto et al. (2016)	nodule and non-nodule region identification	active counter filter & features from CNN model	104 PET/CT images from random screening programs	90.1%, with 4.9 FPs/case	segregates nodule & non-nodule region
16	Hoo-Chang et al. (2016)	lung disease identification	Alex-nett Google-net	LIDC & NELSON dataset	Alex-net: 53% accuracy Google-net:79% accuracy	only distinguishes between normal & abnormal lung

CHAPTER 3

Pipeline Approach for Identifying the Lung Nodules in CT Images

3.1 Introduction

The images collected through CT scans need to be systematically processed to identify sub-solid nodules. The raw images also contain information that are not relevant to the present study. They are referred to as noise and need to be reduced as it hinders the analysis and diagnosis of the image data. After denoising, the image data is further analyzed to mark the nodule regions. The sub-solid nodules are then identified from these already marked nodules. The entire process of identifying the sub-solid nodules from raw images follows a well defined procedure. The methodology adopted for identification of sub-solid nodules is explained in this chapter.

The method adopted in the present analysis comprises of two phases. In the first phase, the process involves removal of noise in the lung CT images, followed by identification of Region Of Interest (ROI) and its segmentation in the given set of images. Appropriate features are identified and accordingly ROI is classified into nodules and non-nodules. The second phase involves analysis of already identified nodule regions leading to the separation of sub-solid nodules. For the purpose, more features related to sub-solid nodules are identified by extracting the Histogram of Gradients (HoG) features. These extracted features are again given as input to the state-of-art classifiers to separate sub-solid and solid nodules. The reliability of the proposed system is substantiated by analyzing the results. The proposed model described in the form of a block diagram is given in Figure 3.1.

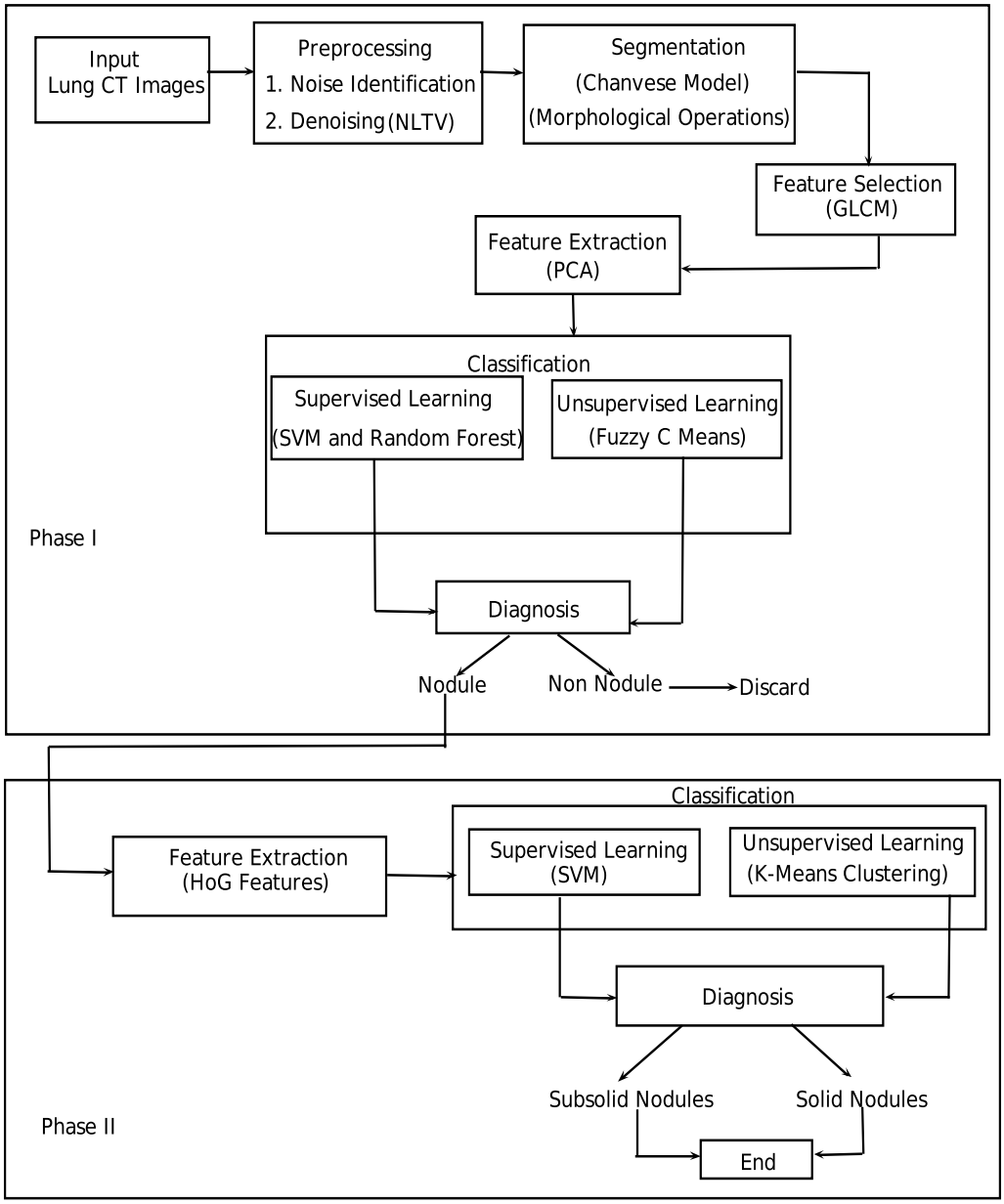


Figure 3.1 Block diagram of the proposed system

3.2 Phase I

3.2.1 Data Preprocessing

3.2.1.1 Image Noise

Noise is an integral part of data. Unwanted and unknown modifications that a signal might suffer during capture, storage, transmission, processing or conversion are termed

as noise. Presence of noise in images distorts the image data and is bound to hamper the analysis of data leading to inaccurate segmentation of regions, which ultimately reduces the accuracy and reliability of the system thus developed. Thus, the image denoising serves as an important preprocessing activity to ensure a proper analysis of the data and it's subsequent diagnosis.

Noise is classified by its statistical properties. The noise present in the data is generally additive or multiplicative with the data. As the name indicates, the additive noise gets added to the data and they are independent of the data whereas, multiplicative noise gets multiplied with the data, therefore it is data-dependent, as the high intensity regions are affected more by the noise. $U_0 = U + n$, and $U_0 = U * n$, represents the additive and multiplicative degradation models respectively, where U and U_0 denotes original and observed images, respectively and n is the noise (a random variable following a specific or mixed distribution(s)). The data-dependent nature of the noise makes the restoration process more tedious.

3.2.1.2 Analysis of Noise

Many of the medical images are observed to be distorted by noise intervention and further the noise being a random variable, its distribution function follows one of the probability distribution functions such as Gamma, Poison, Gaussian and Rayleigh (Gravel et al., 2004).

In most of the studies, the noise distribution is assumed prior to the analysis phase. However, a prior assumption of the noise distribution results in low performance of the pre-processing system. Therefore, an automated noise analysis is done in this thesis work to identify the noise distribution from given input images. This prior knowledge about the noise distribution helps in the design of appropriate restoration models to process the data more efficiently. A large number of noisy input images are considered for the analysis. Low oscillatory regions referred to as constant / homogeneous intensity region from the input data are extracted. The variation in these regions are generally due to noise, as the data is less oscillatory. The Probability Density Function (PDF) of the distribution evaluated from the histogram of such regions gives the approximate distribution of the noise. It is observed that the distribution follows a Gamma law from

the experiments conducted on the input dataset.

To analyze the distribution empirically, an assumption is made about the noise distribution and tested for the viability of the assumption. For instance it is assumed that the noise follows a Gamma distribution, then the noise parameters are the shape (k) and scale (σ). These parameters are to be estimated from the input data using Maximum Likelihood Estimator (MLE) approach. Once the parameters are estimated, then the next step is to analyze the divergence of the distributions (i.e. the one which is assumed and the one which is estimated adaptively). This is done using a curve fitting and divergence of the probability distribution function. Furthermore, a plot of the PDF (derived using the histogram) of the selected region and its estimated PDF (from the parameters) are shown in Figure 3.2.

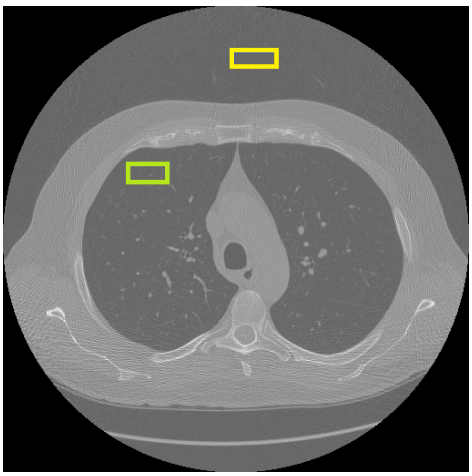
It is observed that the plot of the data from the estimated parameters fits well to the plot of the distribution from the raw data. Therefore, it can be concluded that the noise in these data sets more likely follow a Gamma distribution, the details follow.

Different homogeneous intensity regions in a lung CT image are presented in Figure 3.2(a). Yellow and Green colors indicate two different homogeneous regions selected for the analysis. The PDF of the intensity distribution of yellow colored region (derived using the histogram of the region) fitted with Gamma curve using the estimated parameters such as shape (2.23) and scale (0.0563) is presented in Figure 3.2(b). Similarly the PDF of the green colored region (derived using the histogram of the region) fitted with Gamma curve derived using the estimated parameters shape (2.276) and scale (0.076) is presented in 3.2(c).

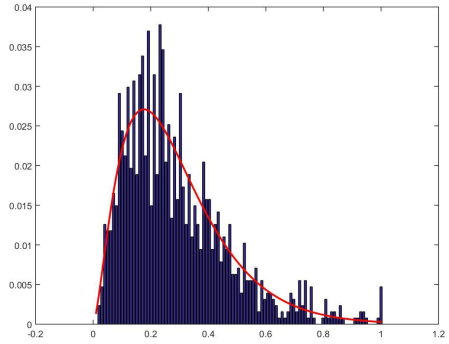
Gamma Noise: Gamma noise is generally seen in the laser images. Probability density function for the gamma noise is given as follows,

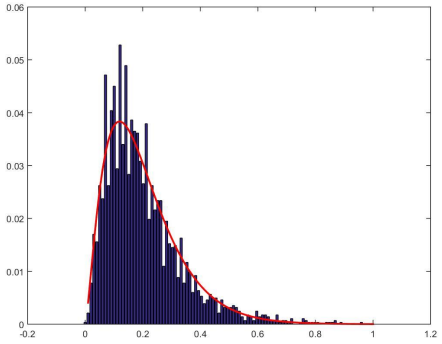
$$PD(g) = \begin{cases} \frac{a^b g^{b-1} e^{-ag}}{(b-1)!} & \text{for} \quad g \geq 0 \\ 0 & \text{for} \quad g \leq 0 \end{cases} \tag{3.2.1}$$

where, a and b are gray values, $\mu = \frac{b}{a}$ and variance $\sigma^2 = \frac{b}{a^2}$. The distribution plot of Gamma curve is given in Figure 3.3.



(a)





(b) (c)

Figure 3.2 (a). Lung CT image showing homogeneous intensity regions marked in yellow and green colors; (b) & (c) PDF plots of Gamma curve fitting.

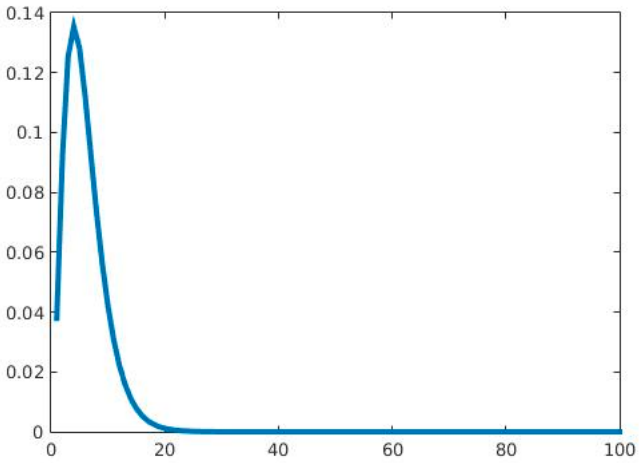


Figure 3.3 Plot of the Gamma distribution

As already pointed out a quantitative analysis to substantiate the nature of the distribution is carried out using Kullback-Leibler (KL) divergence. It explains how one probability distribution diverges from the other one. Typically a zero KL divergence indicates that the distributions are pretty close to each other. The KL divergence for two distributions P and Q of a continuous random variable is,

$$D_{KL}(P,Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx, \tag{3.2.2}$$

where p and q are the densities of P and Q respectively. The KL divergence for each of these region is evaluated and tabulated in Table 3.1.

Table 3.1 KL divergence for different regions considered in Figure 3.2(a)

Regions	KL divergence values
Red	0.0876
Blue	0.0324

It can be observed from Table 3.1 that the values of KL divergence are closer to zero indicating the divergence from one distribution to the other is very less. Hence, as assumed in the beginning of the analysis phase, it can be concluded that the Gamma noise is present in the input image.

It can be seen from literature that variety of filters or combination of filters are extensively used for denoising of nodular images. Magdy et al. (2015) adopted wiener filter to remove noise while preserving the edges and fine details of the images. Combination of wiener filter and median filter is also used for denoising images.

Since the noise follows gamma law, it is source-dependent and multiplicative in nature. Hence, general methods designed for reducing data-independent noise (like white noise) do not perform well in case of data-correlated distribution. Considering the data-correlated nature of the noise as well as the need for analyzing the distribution, a Non-Local Total Variation Minimization (NLTV) model is derived using the Maximum A Posteriori (MAP) estimator for gamma distributed noise, adopting the concepts from Aubert Aujol (AA) method (Aubert and Aujol, 2008). The same is considered in the present study.

3.2.1.3 Aubort Aujol model

(Aubert and Aujol, 2008) proposed a model (AA model) for multiplicative gamma noise where data fitting term is derived based on the MAP estimator and a diffusion term from TV regularization algorithm. It is derived as an unconstrained minimization problem.

The diffusion term in this model is considered to be convex for all values of U though not strictly so. However, the reactive term is conditionally convex, the condition for convexity being $0 < U < 2U_0$, where U and U_0 follows the earlier definition. Therefore, the model is not strictly convex. Hence an unique solution cannot be derived for this model and also the sequence may not converge in the space of Bounded Variations (BV) provided the image U satisfy the above mentioned condition.

MAP estimator

Consider a multiplicative noise degradation model,

$$U_0 = (KU)n, \tag{3.2.3}$$

where, U is the original input image, n represents noise, K denotes the linear blurring operator and U_0 represents the resultant image after adding noise. Further, assume that n follows a Gamma law with mean 1(Aubert and Aujol, 2008) which is obtained by,

$$h_N(n) = \frac{L^L}{\Gamma(L)} n^{L-1} e^{-Ln} 1\{n \geq 0\}. \tag{3.2.4}$$

After computation the following equation is obtained,

$$h_{F|U}(U_0|U) = \frac{L^L}{U^l \Gamma(L)} U_0^{L-1} e^{(-LU_0)/U}. \tag{3.2.5}$$

It is also assumed that U follows a Gibbs prior,

$$h_U(U) = \frac{1}{C} exp^{(-\lambda \phi(U))}, \tag{3.2.6}$$

where, ϕ represents a non-negative given function, and C denores a normalizing constant. The aim is to maximize $p(U|F)$, where, U represents the original image, F represents the noisy image p denotes the probability. Maximizing $p(U|F)$ leads to classical MAP estimator.

From the Bayes rule,

$$p(U|F) = \frac{p(F|U)p(U)}{p(F)}. \tag{3.2.7}$$

Maximizing $p(U|F)$ amounts to minimizing the log-likelihood,

$$-log(p(U|F)) = -log(p(F|U)) - log(p(U)) + log(p(F)). \tag{3.2.8}$$

The given image is discretized and the following functional is used to restore the given image from gamma corrupted noise.

$$\int \left(log(KU) + \frac{U_0}{KU} \right) + \frac{\lambda}{2} \int \phi(U) dx. \tag{3.2.9}$$

The corresponding energy minimization problem is given as,

$$\min_U \left\{ E(U) = \int_\Omega |\nabla U| dxdy + \lambda \int_\Omega (log(KU) + \frac{U_0}{KU}) dxdy \right\}. \tag{3.2.10}$$

Its Euler Lagrange equation is,

$$-\nabla . \left(\frac{\nabla U}{|\nabla U|} \right) + \lambda K^* \left(\frac{KU - U_0}{(KU)^2} \right) = 0. \tag{3.2.11}$$

The gradient descent solution for Equation (3.2.10) is given as,

$$u_t = div \left(\frac{\nabla U}{\sqrt{|\nabla U|^2 + \beta}} \right) + \lambda K^* \left(\frac{U_0 - KU}{(KU)^2} \right), \tag{3.2.12}$$

where, $\lambda > 0$ is the Lagrange multiplier and K is the linear blurring operator. By multiplying Equation (3.2.12) by $K^*(U_0 - KU)$, the value of λ is obtained and is given as,

$$\lambda = -\frac{1}{\sigma^2} \int_\Omega \left(\frac{\nabla U}{\sqrt{|\nabla U|^2 + \beta}} \right) . \nabla K^* (U_0 - KU) dxdy. \tag{3.2.13}$$

3.2.1.4 Non Local Total variation

A total variation (TV) regularization under non-local framework (Gilboa and Osher, 2008) for additive Gaussian noise can be stated as,

$$\min_U \{ \| \nabla_{NL} U \| + \lambda \| U - U_0 \|_2^2 \}, \tag{3.2.14}$$

where, U and U_0 represent the original and noisy images respectively, λ is a regularization parameter, and $\|.\|$ represent the TV norm. The above model takes the following form under a gamma distributed noise set-up (derived through a MAP estimator process).

$$\min_U \left\{ \int_\Omega \left(\| \nabla_{NL} KU \| + \lambda \log KU + \frac{U_0}{KU} \right) dxdy \right\}. \tag{3.2.15}$$

Assuming the linearity and spatial invariance of the operator K, we use a linear convolution operation with a kernel k. Therefore, KU is implemented as $k * U$ where $*$ is the convolution operator and k is a linear convolution kernel. A Gaussian kernel is used in

place of k, i.e. $k(x,y) = (1/2 * \pi * \sigma^2) * exp(-(x.^2 + y.^2)/2 * \sigma^2)$. The Non-Local gradient of a function $U : \Omega \longrightarrow \mathbb{R}$, for a pair of points or pixels $(x,y) \in \Omega \times \Omega$ is defined as,

$$\nabla_{NL} U(x,y) = (U(y) - U(x))\sqrt{w(x,y)} : \Omega \times \Omega \longrightarrow \mathbb{R}, \tag{3.2.16}$$

where, $w(x,y)$ defines the weight of edge between x and y. It is assumed that the $w : \Omega \times \Omega \longrightarrow \mathbb{R}$ is symmetric, i.e, $w(x,y) = w(y,x)$.

Therefore norm of the Non-Local gradient is defined as,

$$|\nabla_{NL} U|(x,y) = \sqrt{\int_{\Omega} (U(y) - U(x))^2 w(x,y) dy} : \Omega \longrightarrow \mathbb{R}. \tag{3.2.17}$$

The projection scheme proposed in Chambolle (2004) is adopted for numerically solving the above mentioned functional as most of the explicit schemes converge very slowly. The projection based method converges faster than the explicit scheme. The sample original Lung CT images considered in the present discussion are given in Figure 3.4. The results of each step starting from denoising till identification of sub-solid nodules are discussed using these sample images henceforth.

Denoised images obtained by adopting NLTV method for Gamma distribution are presented in Figure 3.5. It is observed that method adopted removes noise giving acceptable results by preserving fine details, edges and texture, in the data sets used for the present study.

Quality of the denoised images are measured using the metrics called Peak Signal to Noise Ratio (PSNR) and Structural Similarity Measure (SSIM). PSNR is defined as a ratio between signal power and noise power. The SSIM indicates the structure and contrast preservation ability of the model under consideration.

PSNR value is defined using Mean Square Error (MSE) as follows,

$$MSE = \frac{1}{mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (U(x,y) - U_0(x,y))^2, \tag{3.2.18}$$

where, U and U_0 represents the original and noisy image respectively and m,n represents the size of the matrix.

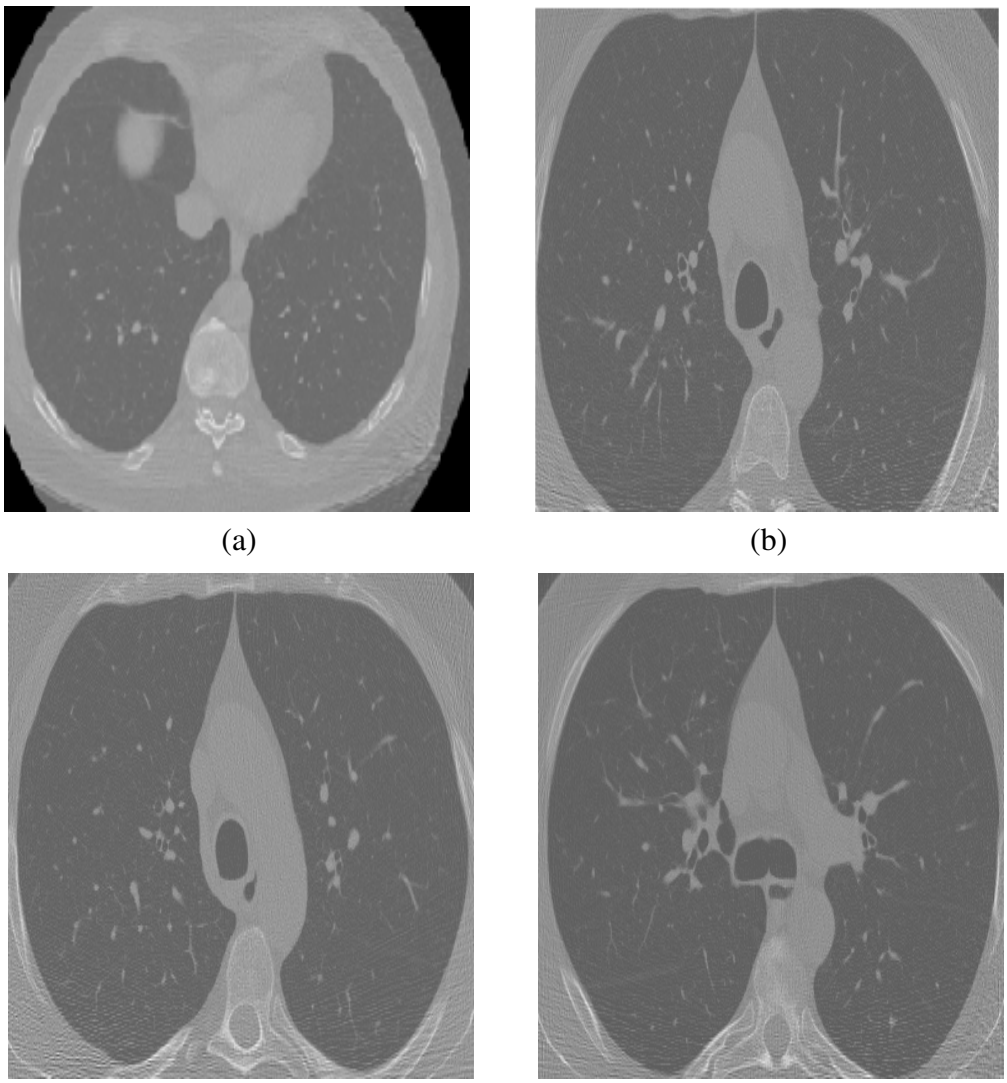


Figure 3.4 (a) and (b) are sample Lung CT image from I-ELCAP database, (c) and (d) are sample Lung CT image from LIDC / IDRI database.

The PSNR (in dB) is defined as,

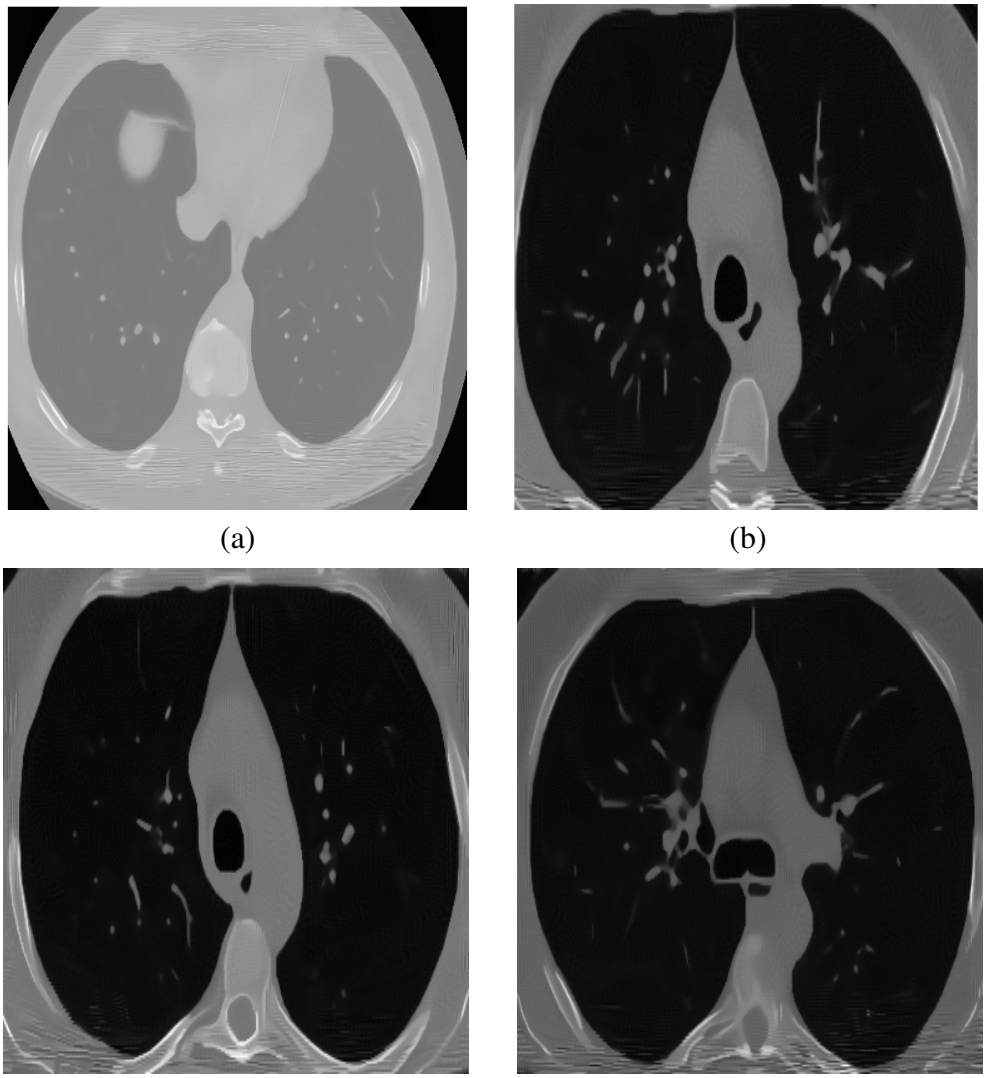
$$PSNR = 10.log_{10}\frac{MAX_U^2}{MSE}, \tag{3.2.19}$$

where MAX_u is the maximum possible pixel value of the image.

SSIM values obtained are in the range 0 and 1, where 0 and 1 indicates the low and ideal structure preservation capabilities, respectively. It is evaluated by partitioning the image into overlapping windows. The measure between two windows $x1$ and $x2$ of size $N \times N$ is

$$SSIM(x1,x2) = \frac{(2\mu_{x1}\mu_{x2}+C_1)(2\sigma_{x1x2}+C_2)}{(\mu_{x1}^2+\mu_{x2}^2+C_1)(\sigma_{x1}^2+\sigma_{x2}^2+C_2)}, \tag{3.2.20}$$

where, μ_{x1} and μ_{x2} represent the average values of windows $x1$ and $x2$ respectively.



(a) (b)

(c) (d)

Figure 3.5 Denoised images (a), (b), (c) and (d) using NLTV

σ_{x1} and σ_{x2} represent the variance of windows $x1$ and $x2$ respectively. σ_{x1x2} is the covariance of $x1$ and $x2$, C_1 and C_2 are constants.

A higher value of PSNR indicates that the noise is removed effectively and a higher value for SSIM denotes the structure preserving capability of the model in general. The PSNR values and SSIM values obtained for the considered sample images are given in Table 3.2.

Table 3.2 Computed PSNR and SSIM values

Images	PSNR value (in dB)	SSIM
Image (a)	23.0	0.83
Image (b)	25.0	0.81
Image (c)	22.5	0.86
Image (d)	22.8	0.79

3.2.2 Segmentation

Segmentation is a process of dividing a given image into groups of pixels called segments. A label is assigned to each pixel such that each pixel having the same label have similar properties. This will help in segregating the required region of interest which in turn helps in identifying the objects present in an image. Region of interest in this model pertains to nodules. Detecting nodules in the preprocessed images is achieved by segmentation process. Segmenting the objects, makes it easier to identify region of interest and locate tumor, lesion and other abnormalities.

Various segmentation algorithms are available for identifying the objects or regions in a given scenario. Widely used ones are thresholding methods, edge detection methods and graph partition approach. Region of interests that need to be identified in this study are the nodules present in lung CT images. These nodules have irregular shape and are small in size. There are some automatic segmentation algorithms developed for identifying the nodules in the given lung CT images (Magdy et al., 2015; Kuhnigk et al., 2006). These algorithms increase the false positives (falsely identified nodule regions). Hence, level set methods that are commonly used for the purpose of segmentation of objects in images that have more irregularities are considered here. The model proposed by Chan and Vese (2001) is widely used for segmenting images with hazy and indistinguishable boundaries. It is an active contour model which is derived from Mumford and Shah (1989) technique and the level set formulation.

Chan-Vese Model

An active contour based model introduced by Chan and Vese (2001) is considered here to segment the region of interest as it is found suitable for segmenting images which

does not have proper edge details. Mumford and Shah (1989) approximate the image U by a piecewise-smooth function f as the solution of the minimization problem. The Mumford-Shah functional is written as,

$$\arg\max_{U,C} \mu \text{Length}(C) + \lambda \int_{\Omega} (f(x) - U(x))^2 dx + \int_{\Omega \setminus C} |\nabla U(x)|^2 dx, \tag{3.2.21}$$

where C is edge set curve and U may be discontinuous. The equation has three terms. First term defines the regularity in C, second term makes U to be closer to $f(x)$ and last term makes sure that U is differentiable on $\Omega \setminus C$. The edge set C is encouraged to select the segmentation boundary by the Mumford-shah approximation model. Chan-Vese model is derived from Mumford-Shah model where an additional term representing the enclosed area is considered.

Further simplification is achieved by allowing U (the image function) to take two values,

$$U(x) = \begin{cases} d_1 & \text{where } x \text{ is inside } C, \\ d_2 & \text{where } x \text{ is outside } C. \end{cases}$$

where C depicts the closed set boundary, the values of U inside and outside C are given by d_1 and d_2 respectively. The level set method is adopted to find the solution for C. The level set function $\phi(.)$ is defined as follows,

$$\begin{cases} \text{On C} : \{\phi(x) = 0\} \\ \text{inside}(C) : \{\phi(x) > 0\} \\ \text{outside}(C) : \{\phi(x) < 0\} \end{cases} \tag{3.2.22}$$

As the topology changes, contour can split and merge simultaneously. This method helps in defining several boundaries simultaneously. Here the initial contour is defined in a checker board shape since it is observed to converge fast and is given as,

$$\phi(x) = sin(\frac{\pi}{5}x_1)sin(\frac{\pi}{5}y). \tag{3.2.23}$$

The energy functional for the formula stated above is given as,

$$E(d_1,d_2,\phi) = \lambda_1 \sum_{i=1}^{n} (U(x)-d_1(x))^2 H(\phi(x)) + \lambda_2 \sum_{i=1}^{n} (U(x)-d_2(x))^2 (1-H(\phi(x))), \tag{3.2.24}$$

where, C is the contour, $\phi(.)$ denotes the level set function of the contour, H is the Heaviside function, $d_1(x_i)$ is the average intensity value inside the curve, and $d_2(x_i)$ is the average intensity value outside the curve.

Heaviside function is a discontinous function whose values are zero for negative arguments and one for positive arguments. Average intensities inside and outside the curve $d_1(x_i)$ and $d_2(x_i)$ are defined as,

$$d_1(x) = \frac{\sum_\Omega g_k(x-y)(U(y)H(\phi(y)))}{\sum_\Omega g_k(x-y)H(\phi(y))} \text{ and,} \tag{3.2.25}$$

$$d_2(x) = \frac{\sum_\Omega g_k(x-y)(U(y)(1-H(\phi(y))))}{\sum_\Omega g_k(x-y)(1-H(\phi(y)))}, \tag{3.2.26}$$

where, $U(y)$ is the Intensity value at y, g_k is the Gaussian kernel function, and $g_k(x-y)$ denote the weight assigned to each intensity.

Results of segmentation process carried out using local region based Chan-Vese model are presented in Figure 3.6. It is observed that the method employed herein has identified all the contours in the image (marked in green color). The structure is defined precisely explaining the suitability of the method for segmentation of ROI in lung CT scans. The identified edges are marked in green color and the regions are segmented accordingly. The contours of small regions are also identified which negates the chances of leaving any region in the given lung image. Checker board shape of the initial contours for segmentation helps in identifying the edge-contours in less number of iterations. The approximate number of iterations that the algorithm takes to complete the execution is 160 which in turn reduces the execution time of the algorithm and it amounts to an average of one and half minutes.

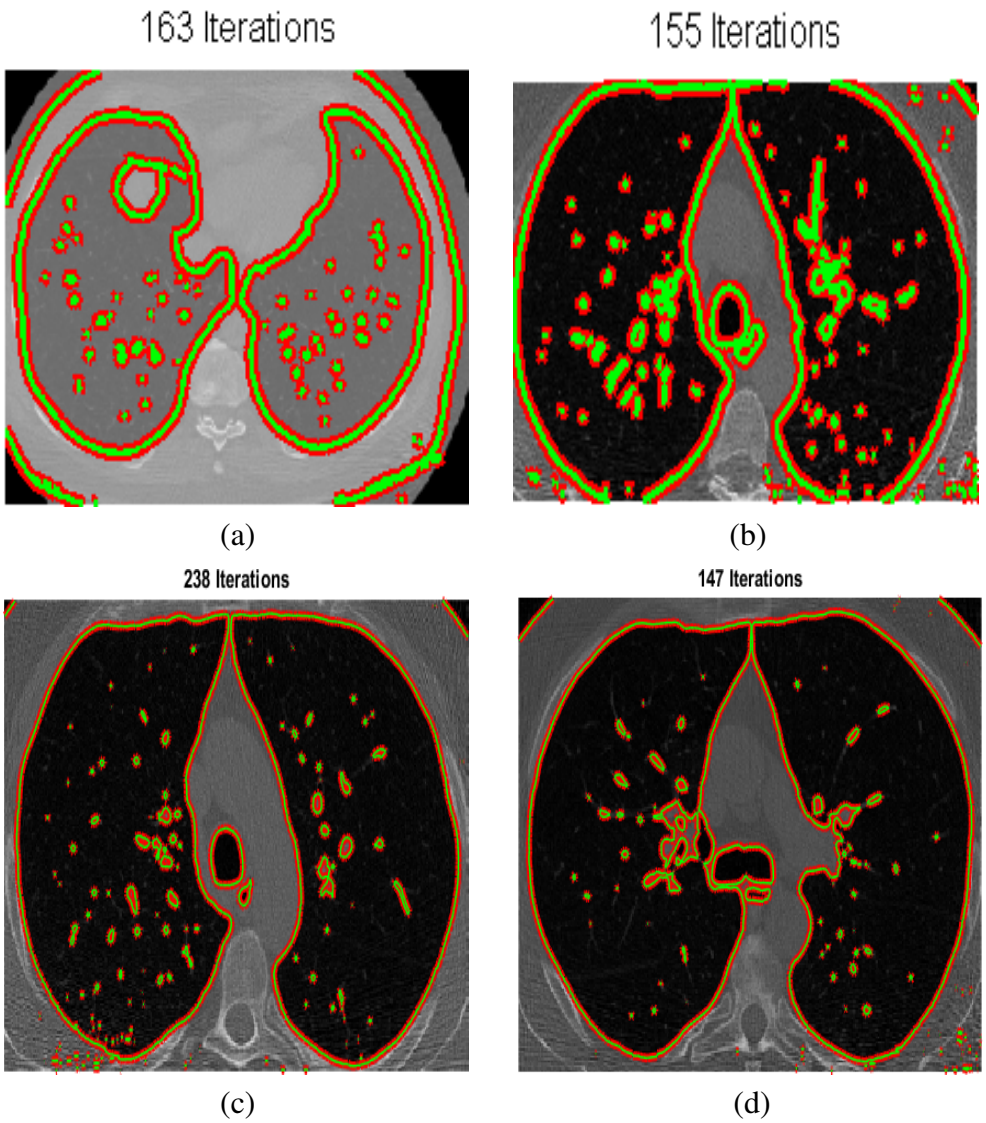


Figure 3.6 Segmented lung images (a), (b), (c) and (d) using Chan-Vese Segmentation method.

3.2.2.1 Segmentation accuracy

The success of developed CAD system depends on the accuracy of segmentation of CT images. It is a crucial step in accurate identification of the nodules. For the purpose, Jaccard Similarity Coefficient, Jaccard Distance and Hausdroff distance are considered. The segmented images and annotated raw images from the data set are used for finding the segmentation accuracy. If both the annotated and segmented images overlap perfectly then segmentation is assumed to be 100% accurate which is an ideal case and is referred to as a perfect segmentation.

Jaccard similarity coefficient:

Similarity and dissimilarity of the given sample is measured with this metric. If R and S are two sets then Jaccard coefficient is expressed as,

$$Jaccard(R,S) = \frac{|R \cap S|}{|R \cup S|}. \tag{3.2.27}$$

Here $|R|$ represents the cardinal of set R.

Jaccard distance:

It finds the dissimilarity between two sets. It is expressed as,

$$d_J(R,S) = 1 - Jaccard(R,S). \tag{3.2.28}$$

Hausdorff Distance:

This metric is commonly used for finding shape similarity. The distance measures should ideally yield zero so as to establish the similarity.

Table 3.3 Segmentation Accuracy Results

Segmented Images	Jaccard Coefficient (percentage)	Jaccard Distance	Hausdorff Distance
Image (a)	97	0.03	0.0031
Image (b)	96	0.04	0.0020
Image (c)	97	0.03	0.0026
Image (d)	98	0.02	0.0014

The three measures namely Jaccard Coefficient, Jaccard distance and Hausdorff distance for the segmented images of lung CT scans and the annotated images are shown in Table 3.3. It can be observed that Jaccard coefficient values for Image 1 and Image 2 are 97% and 96% respectively, indicating a perfect similarity between the segmented image 1 and its annotated raw image. Further, Jaccard distance being less than 5% for both the segmented images, the dissimilarity between the two sets are near zero. The Hausdorff distance measured in the two images are in the order of 0.003 and 0.002. This shows very clearly that there is very high shape similarity between two sets. Thus the results indicate that the segmentation of CT images have been achieved and there is near perfect accuracy in the segmentation process. This shows that the CAD system developed till this stage has yielded reliable results.

3.2.2.2 Morphological Operation

The shape and size of the lung nodules are important morphological features which facilitates the identification of nodule region in CT images. These operations are carried out for establishing more clarity and perfection in the observed lung nodule region.

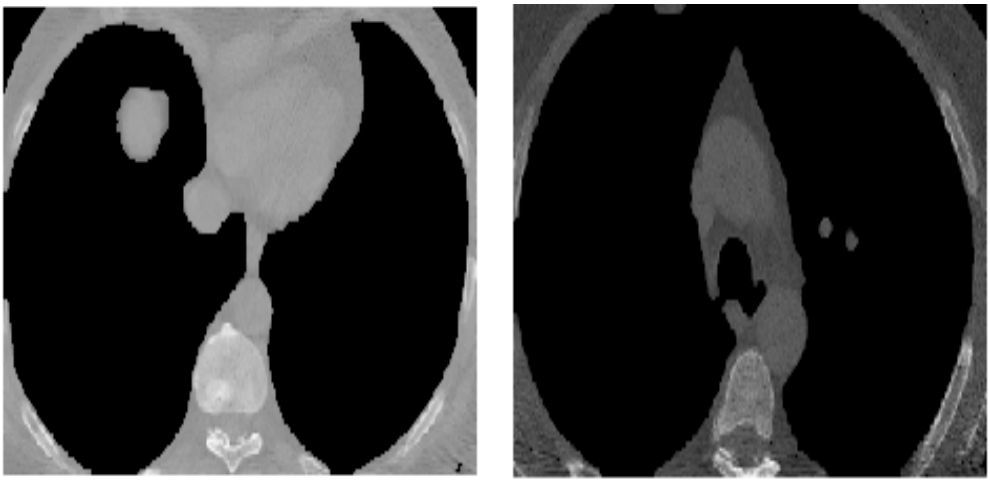
Morphological operation is a set of non linear operations which relates itself to shape features in the image. Most of the imperfections in the image are removed and structures are more resolved by morphological processing. The task of removing these imperfections are handled by two fundamental composite operators. These operators combine similar featured areas and remove small holes present in the image with the help of a structuring element. Structuring elements are the probing masks used to modify the morphology of the objects present in an image. Shape and size are the two characteristics related to it. Structuring element is a small matrix of pixels. This matrix is made up of zeros and ones. The arrangement of the zeros and ones specifies the shape of the structuring element. The dimension of matrix specifies the size of the structuring element.

The two fundamental morphology operators are erosion and dilation. Erosion, erodes away the boundaries of foreground pixels, usually the white pixels. Thus areas of foreground pixels shrink in size, and "holes" within those areas become larger. Dilation enlarges the areas of foreground pixels (i.e. white pixels) at their borders. The areas of foreground pixels thus grow in size, while the background "holes" within them shrink. The composite morphological operators named open and close filters are derived from erosion and dilation operators by combining them. Opening of set A by set B is achieved by first eroding the set A by B, then dilating the resulting set by B. Similarly, the closing of set A by set B is achieved by first dilating set A by B, then eroding the resulting set by B.

Jacobs et al. (2014) observed that a disc structuring element of 4×4 pixel is found to give nodule size of 10-30 mm diameter since the standard nodule size is said to be in the range of 10-30 mm. Accordingly, in this study "open" filter followed by "dilation" operation is adopted to make the segmented areas more clearly visible. Small unwanted

regions are removed by merging them to background and the object boundary is made to look more clear merging the pixels to the foreground area. This increases the visibility of the segmented areas and removes spurious holes and unwanted connections which affect the classification process adversely.

The results obtained are presented in Figure 3.7 and zoomed in portions are shown in Figure 3.8. It is observed that very small regions identified in the segmentation process, are eliminated by making them to merge with the background. This also makes the system less complex by reducing the number of regions to be processed in further process. A few more results are shown in Figure 3.9.



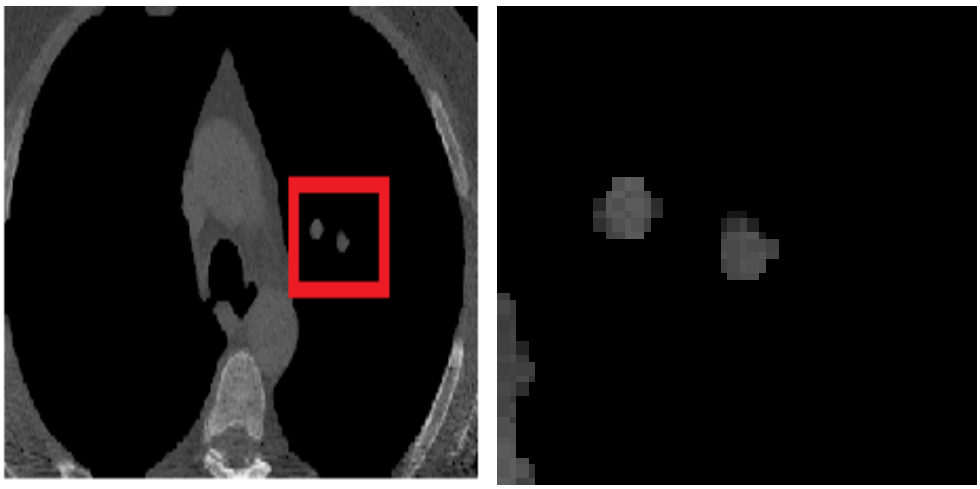
(a) (b)

Figure 3.7 Morphological Operation Results of image (a) and (b)

Similar more zoomed in portions of the images considered are shown in Figure 3.9.

3.2.3 Feature Extraction

The regions identified in segmentation process need to be classified as nodule or non-nodule. In order to do so, appropriate features of nodule region are to be identified and extracted. With the help of these features, classification models predict the class labels. The classifier output determines the possibility of a given structure being a nodule or non-nodule. Generally, features such as intensity, texture, shape and context feature



(a) (b)

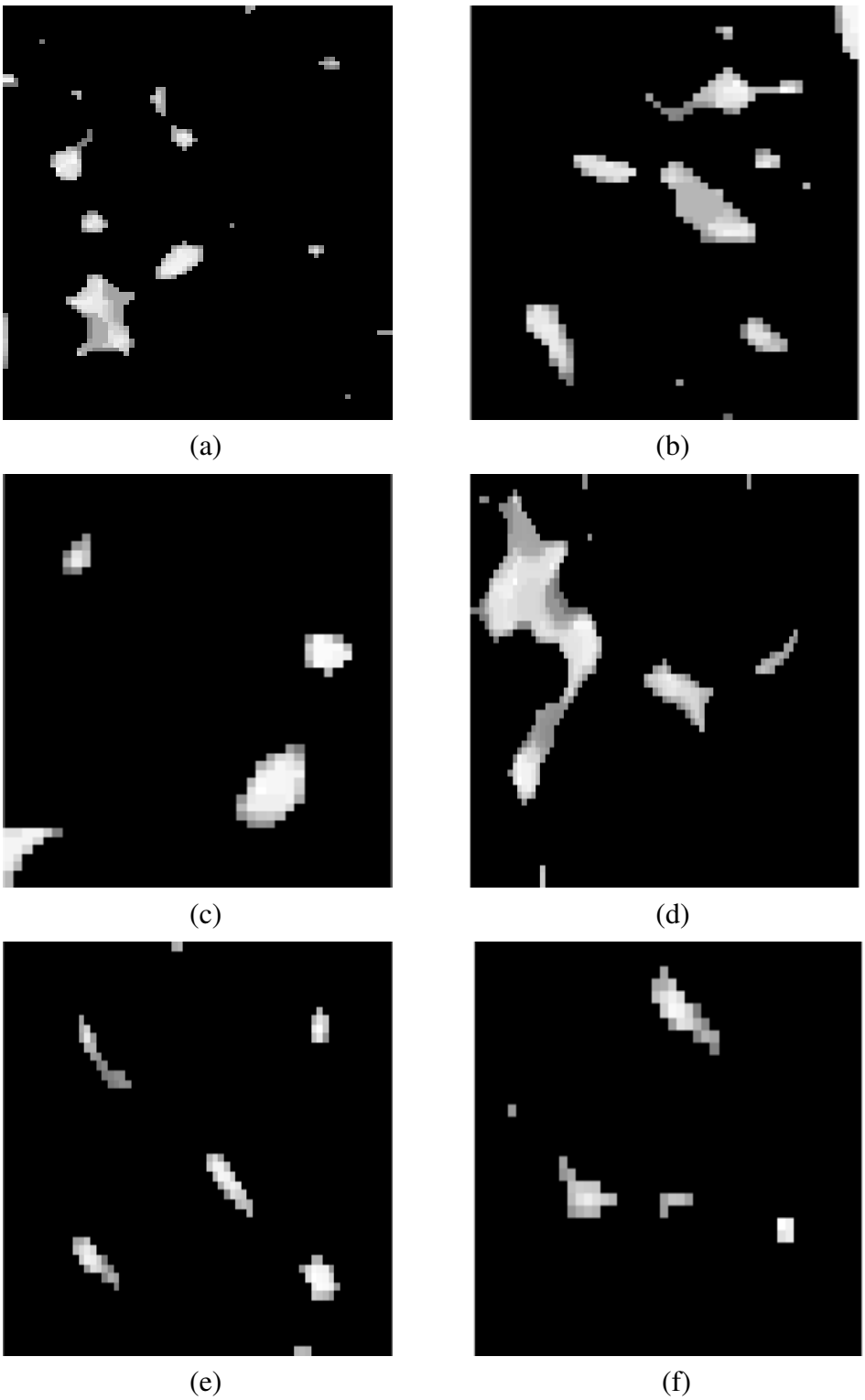
Figure 3.8 (a) Region of interest identified, (b) Zoomed in portion of image (a) showing the region of interest

of an image play a predominant role in classification of segmented ROI (Jacobs et al., 2014). The same have been considered to define the lung nodules. But it is observed that false positives indicating misclassification of nodules and non-nodules are more if only these features are considered. Hence more number of sophisticated features have been incorporated to improve the accuracy of prediction.

(Chen et al., 2012) have observed that first-order and second-order statistical properties give relevant and distinguishable features without involving any computational transformations in general standard images such as Lena, Cameraman etc.. Further, it is also observed that Gray Level Covariance Matrix (GLCM) performs well compared to other statistical measures in homogeneous images (Soh and Tsatsoulis, 1999). Hence, in the present study second order statistical measures are considered as it allows to use pixel neighborhood relationship in an image.

GLCM is created from a gray-scale image. It calculates how often a pixel with gray level value of i occurs whether horizontally, vertically or diagonally to adjacent pixel with value j. $p(i, j)$ is the probability of finding the relationship of (i, j) or (j, i). The sixteen statistical properties generated from GLCM of an image given in Table 3.4 are considered as features for further process.

The details of features used in GLCM are given in Table 3.5. However, it may not



(a) (b) (c) (d) (e) (f)

Figure 3.9 Zoomed in portions of the morphological operation results

Table 3.4 GLCM Features

Autocorrelation	Contrast
Correlation	Dissimilarity
Cluster Prominence	Cluster shade
Energy	Entropy
Homogeneity	Sum average
Maximum Probability	Sum entropy
Sum of Squares (Variance)	Inverse difference
Difference entropy	Difference Variance
Information measure of correlation	Inverse Difference normalized

be necessary to consider all the features for further analysis. Some of them may not be relevant. They may depend on some other feature or may overlap. It is most likely that considering all the features may result in unreliable result. Hence, it is important to choose significant features from an image. For the purpose, Principal Component Analysis (PCA) algorithm has been employed in the current study.

3.2.4 Feature Selection

Principal Component Analysis (PCA) is adopted for selecting the relevant features from already identified feature set. A square symmetric covariance matrix is formed from the features and an Eigen value decomposition is performed. Let A be a $m \times n$ matrix whose columns are formed by the feature vectors and n denotes the dimensionality of the features and m the number. A square symmetric covariance matrix is obtained by multiplying A (centered by subtracting out the mean) with its transpose.

$$PCA = B = (A - \mu)(A - \mu)^T \tag{3.2.29}$$

where μ is the mean vector whose entries are the mean of the feature vectors. The matrix B is a square symmetric covariance matrix, whose Eigen values are real and positive. Moreover, the Eigen vectors corresponding to distinct Eigen values are linearly independent. Therefore, n such eigen vectors span a n dimensional space, hence they are basis vectors.

Table 3.5 Computation of Gray Level Covariance Matrix (GLCM) Features

Feature identified	Formula
Contrast	$\sum_{x=0}^{m}\sum_{j=0}^{n}(i-j)^2 p(i,j)$
Homogeneity	$\sum_{i,j}^{n}\frac{p(i,j)}{1+\lvert i-j\rvert}$
Area	$\sum_{i,j} f(i,j)$
Mean	$\mu_x = \sum_i \sum_j i p(i,j)$
	$\mu_y = \sum_j \sum_j j p(i,j)$
Sum Average	$\sum_i \sum_j (i-\mu) p(i,j)$
Sum of Squares	$\sum_i \sum_j (i-\mu)^2 p(i,j)$
Correlation	$\sum_i \sum_j p(i,j)\frac{(i-\mu_x)(j-\mu_y)}{\sigma_x \sigma_y}$
Angular Second Moment(ASM)	$\sum_i \sum_j p(i,j)^2$
Energy	$\sqrt{ASM}$
Inverse Difference Moment	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Entropy	$-\sum_i \sum_j p(i,j) log(p(i,j))$
Equivalent Diameter	$\frac{\sqrt{4*Area}}{\sqrt{\Pi}}$
Degree of Asymmetry	$\sum_i \sum_j (i-j)^3 p(i,j)$
Kurtosis(Relative Peak)	$\sum_i \sum_j (i-j)^4 p(i,j)$
Perimeter	Distance between neighboring pair of pixel along the border
Eccentricity	$\frac{MinorAxisLength}{MajorAxisLength}$

Eigen decomposition of B is computed as,

$$O, S, V = Eig(B), \tag{3.2.30}$$

where, O represent the Eigen vectors of B(column vectors), S represent the Eigen vectors of B(row vectors), and V represent the Eigen values. The matrix S is a diagonal matrix which contains the Eigen values. The obtained Eigen Values are sorted in ascending order along with the corresponding Eigen vectors O and the vectors corresponding to small Eigen values are discarded as they do not contribute much significantly to the analysis. A set of new basis vectors is formed from the chosen k Eigen vectors corresponding to the highest k Eigen values which creates a space spanned by the principal components. Any input vector is projected on the newly created basis to form a new vector whose dimension is reduced to k.

The Eigen value thus obtained are arranged in the decreasing order and also the corresponding Eigen vectors. The Eigen vectors so formed give a new feature space. The features pertaining to the nodules are given by the first few Eigen values which are significant. The feature vectors or the Eigen vectors are selected by using the dominant Eigen values. To decide upon the selection of the number of vectors, the following method has been used.

The ordered Eigen values are plotted against the corresponding eigen vectors as shown in Fig. 3.10. The number of features to be selected is described by the area under the curve. The number that covers 90% of the area under the curve is chosen empirically to generate the new feature space. Only that many feature vectors are preserved while the others are omitted as they are not much significant for the study.

From the curve presented in Figure 3.10, it can be observed that considering a symmetric Gaussian distribution with standard deviation σ, 3σ values from the origin covers more than ninety percent area under the curve. It clearly shows that identification of nodules can be clearly done with the help of first three Eigen vectors corresponding to the high Eigen values. Accordingly, the first three features are found to play the prominent role in identifying the nodules. Hence these features are taken as input for identifying the nodules in the classification model.

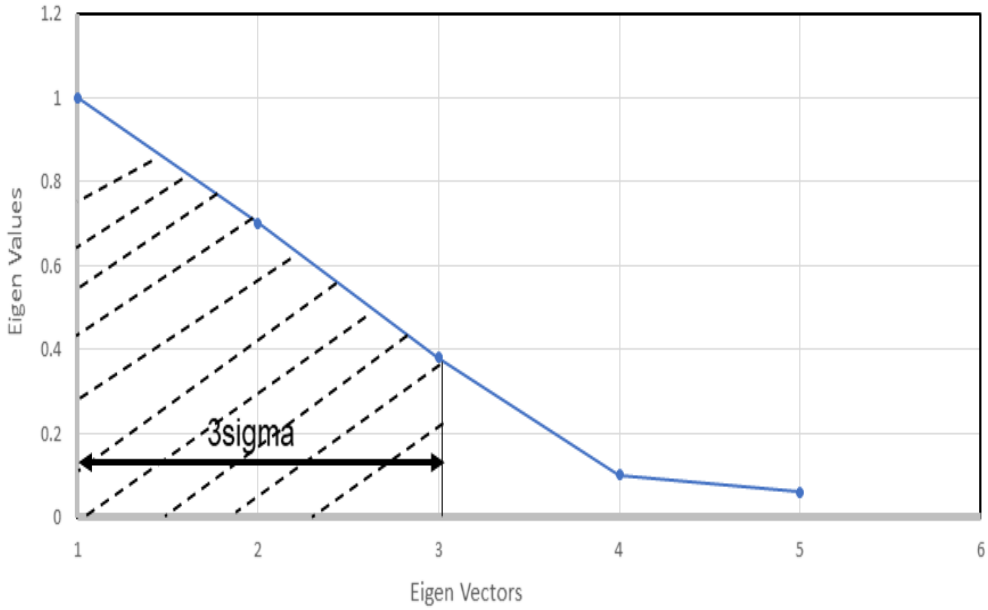


Figure 3.10 Plot of Eigen values versus Eigen vectors

3.2.5 Classification of Nodules and Non-nodules

Selection of relevant features in the nodule region based on their statistical properties form the basis for classification of the nodules. Classification process assigns the class labels for similar kind of data. Classification algorithms are of two types, supervised and unsupervised.

Supervised classifiers predict the output based on already available training samples (sample classes). Classification will be more accurate in this case since the sample classes are used by the algorithm. In case of unsupervised classification, no sample classes are available. These algorithms group the similar pixels together and labels them as a class.

Classes in the present analysis concentrates only on the nodules and non nodules in lung CT images. Hence there are two class labels. Support Vector Machine (SVM), Fuzzy C Means and Random Forest algorithms have been considered for classifying the regions as nodules and non nodules. SVM algorithm and Random Forest algorithms are supervised classifiers which are dependent on training samples. Fuzzy C Means algorithm is an unsupervised algorithm where the similar pixels are grouped using Euclidean distance function.

Support Vector Machine (SVM)

SVM is one of the supervised classification techniques which perform classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels (Guenther and Schonlau, 2016). SVMs are capable for classifying linearly separable data. However, they are made to handle non-linearly separable data by taking the data to higher dimensions in which they are linearly separable. Kernel based SVMs are used for the purpose. Kernels can be linear, polynomial or Radial Basis Function (RBF). Radial basis function kernel is considered in this study for predicting the class labels as nodules and non-nodules.

Radial basis function kernel K for a given data r and each support vector r' is given as,

$$K(r,r') = exp(-\frac{\|r-r'\|^2}{2\sigma^2}) \tag{3.2.31}$$

Since SVM is supervised classification model, it relies on training samples. Training sample is a set of images in which the input data and the required output data are known beforehand. In this case, the nodule and non-nodule regions marked as two class labels are known. Training dataset is a collection of marked nodule and non-nodule regions in the lung CT images. Using these images the algorithm learns and predicts the output in the test data. Test data is set of images given as input to the system for identifying the region of interest.

A random set of 450 images is considered in this study. 150 randomly selected images are used as training data set where the regions of nodule and non-nodule are demarcated. Proving the consistency of the prediction model, is an important step in the study. For this purpose, the total test data considered for analysis is randomly grouped into three sets of 120, 160 and 200 images each with known number of nodules and non-nodules which are given in Table 3.6.

For each set of images four values are obtained as output (Table 3.6). They are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These values are the outcome of the prediction done by the algorithm and together when they are arranged in the form of a matrix they are called confusion matrix. If the sample and the predicted output are positive then it is called as true positive. Similarly, if the sample

Table 3.6 Classification results of SVM

Number of Images	True Positive	True Negative	False Positive	False Negative
120 (40 Nodule,80 Non Nodule)	38	77	3	2
160 (86 Nodule,74 Non Nodule)	82	70	4	4
200 (115 Nodule,85 Non Nodule)	109	80	5	6

considered is negative and output predicted is negative then it is true negative. But, if the sample considered is positive and the predicted output is negative then it is termed as false negative. Similarity, if the sample is negative and predicted output is positive it's called as false positive. Performance of the system is evaluated by using these four parameters. An accurate system should always have less number of false positives and false negatives (false predictions). Hence, these values are used to calculate the accuracy and reliability of the system.

From predicted output of SVM algorithm given in the Table 3.6, it is observed that irrespective of the groups considered, the false prediction is minimum i.e. false positives and false negatives in all the case is in the order of 0.05%. This clearly depicts that number of images considered does not form the criteria for the analysis.

The efficiency of the system is tested by computing the measures such as accuracy, precision, recall, specificity, error rate, f-measure etc. for images in the training set.

Accuracy is a ratio of correctly predicted samples to the total available samples. Precision value indicates correctly identified samples from the actual labeled samples. Hence, a high precision value indicates low false predictions which in-turn makes the system reliable.

The existence of actual labeling for the predicted samples is measured by the parameter recall. Hence, the recall value should be high and it indicates the system sensitivity. Specificity is a measure which identifies the true negatives as true negatives itself. For an accurate system the value of specificity should be high. Error rate defines the frequency of error occurred by the system while making the prediction. F-measure is the weighted average of both precision and recall measures. If there exists an uneven class distribution then f-measure is used to get a balance between precision and recall metrics.

The expressions for computing these parameters are given in Table 3.7. Table 3.8 presents the computed values for these parameters.

Table 3.7 Accuracy Measures

Accuracy= $\frac{TP+TN}{(TP+TN+FP+FN)}$	Precision (P)= $\frac{TP}{(TP+FP)}$
Recall(Sensitivity) (R)=$\frac{TP}{(TP+FN)}$	Specificity=$\frac{TN}{(TN+FP)}$
Error Rate= $\frac{FP+FN}{(TP+TN+FP+FN)}$	F-Measure= $\frac{2*(P*R)}{(P+R)}$

where, TP denotes the True Positives, TN denote the True Negatives, FP and FN denotes False Positives and False Negatives respectively.

The results of performance analysis for SVM algorithm computed using the values obtained in Table 3.6 are given in Table 3.8. It can be observed from the Table 3.8 that average accuracy is in the order of 95% for all the three image sets considered. It clearly shows that 95% of the total samples considered have shown the correct prediction in all the three sets. Average specificity value is in the order of 94% which indicates that 94% of the true negatives are correctly identified as true negative itself. Recall value is in the order of 95% which clearly shows that the sensitivity of the system is high.

The average precision value is 94%. This shows that false prediction is only in the order of 6% showing high reliability of the system. Prediction error is less than 5% which indicates the system is reliable and less prone to errors. The weighted average of both precision and recall measures is again 94%.

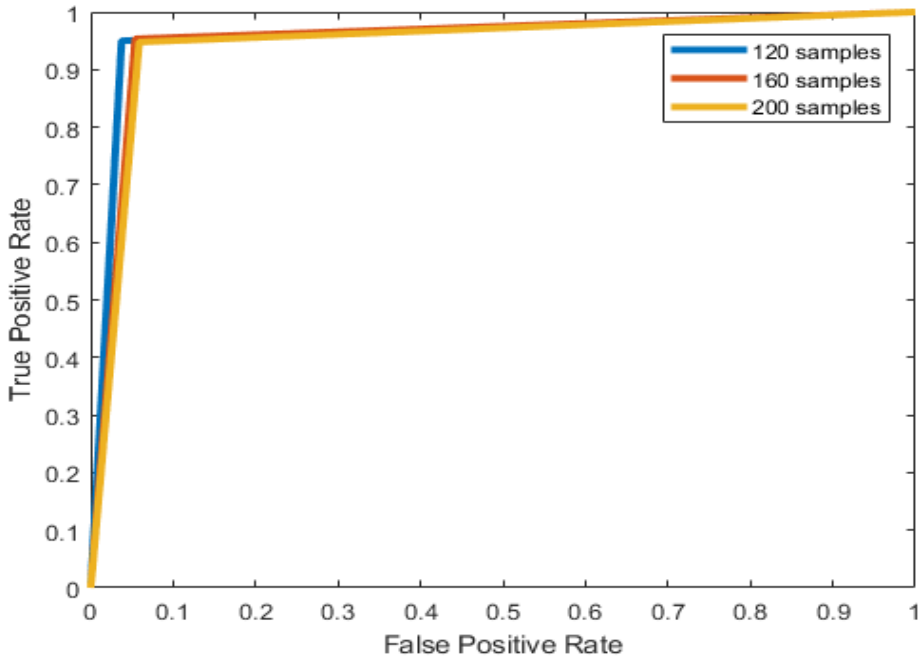
Thus the performance analysis of the SVM classification clearly indicates that the system developed is reliable and efficient. K-cross validation is done for 200 (115 nodules, 85 non-nodules) images with k-value set to 10.

Further, the system performance can also be verified by plotting Receiver-Operating Characteristics (ROC) curve. It is a graphical plot of true positive rate versus false positive rate, the relationship between the sensitivity and specificity. The Area Under the ROC Curve (AUC) depicts the overall measure of specificity and sensitivity. The

Table 3.8 Performance analysis of SVM classification method

Number of Images	Accuracy (percent)	Precision	Recall	Specificity	Error-Rate	F-measure
120 (40 Nodule, 80 Non Nodule)	95	0.92	0.95	0.96	0.04	0.93
160 (86 Nodule,74 Non Nodule)	95	0.95	0.95	0.94	0.05	0.95
200 (115 Nodule,85 Non Nodule)	94	0.95	0.94	0.94	0.05	0.94

value of this can be from 0 to 1. If the AUC is 1 then the classification is perfect (ideal case). If the AUC is nearing one then the system is said to perform well. In this study, ROC curve for all the three randomly selected sample sets (60,120,200 images) are plotted and given in Figure 3.11.



(a)

Figure 3.11 ROC curve for SVM algorithm

It can be observed that the curves are nearing one indicating that the performance of the system is near perfect. It can also be seen that number of samples considered will not effect the performance of the system since the ROC curves for all the three sets are nearing one.

Fuzzy C Means:

It's an unsupervised learning classification technique where each piece of data can belong to more than one cluster. Each data is assigned to a class such a way that items in the same class are similar to each other. Clusters are identified using the similarity measures such as distance, connectivity and intensity based on the data or application. A number of clusters are chosen and each data point is assigned a membership grade randomly. Membership grades defines the degree to which the data considered belong to that cluster. Eventually data points are assigned to one of the clusters. This method is repeated until the algorithm converges (Zheng et al., 2015). Mathematically, this algorithm aims to minimize the following function,

$$arg\min_{s} \sum_{x=1}^{N} \sum_{y=1}^{c} W_{xy}^{m} \|s_x - c_y\|^2, \tag{3.2.32}$$

where, s_x is the xth point in the measure data, c_y is the center of the cluster, $\|s_x - c_y\|^2$ Euclidean norm expressing the similarity between the data and center of the cluster, m is an integer greater than 1, and W_{xy} is the membership degree assigned to s_x in the cluster c_y.

Each time, the membership grade W_{xy} and cluster centers c_y are updated as follows,

$$W_{xy} = \frac{1}{\sum_{k=1}^{c} (\frac{\|s_x - c_y\|}{\|s_x - c_y\|})^{\frac{2}{m-1}}} \tag{3.2.33}$$

and

$$c_y = \frac{\sum_{x=1}^{N} W_{xy}^{m} \cdot s_x}{\sum_{x=1}^{N} W_{xy}^{m}} \tag{3.2.34}$$

Fuzzy C-Means algorithm is considered for classification of images in all the three sets. The results of this unsupervised classification is given in Table 3.9.

The number of false predictions including both false positive and false negatives is again less with an average of 8 samples. The performance parameters computed using the values from the Table 3.9 are given in Table 3.10.

Table 3.9 Classification results of Fuzzy C-Means

Number of Images	True Positive	True Negative	False Positive	False Negative
120 (40 Nodule,80 Non Nodule)	36	76	4	4
160 (86 Nodule,74 Non Nodule)	81	70	4	5
200 (115 Nodule,85 Non Nodule	108	79	6	7

Table 3.10 Performance analysis of Fuzzy C-Means algorithm

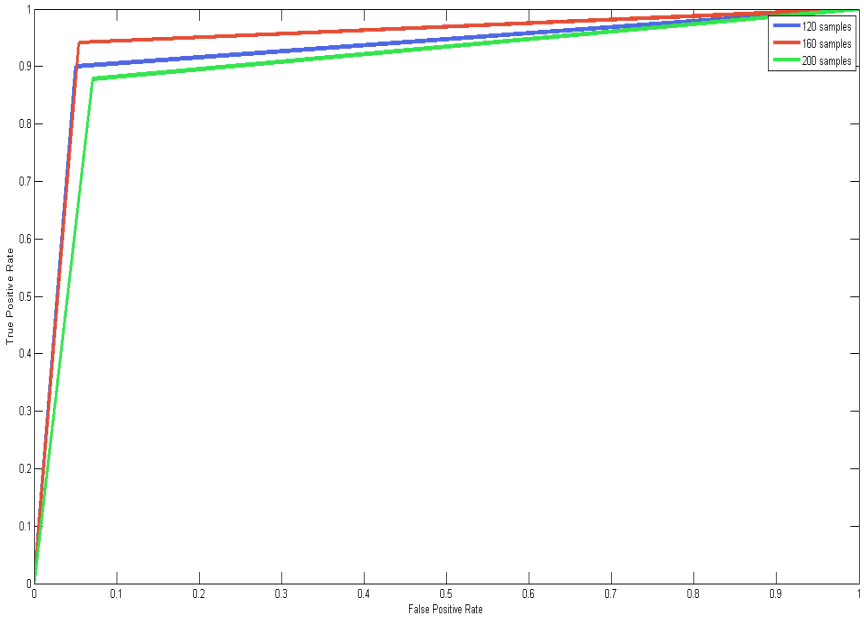
Number of Images	Accuracy (percent)	Precision	Recall	Specificity	Error-Rate	F-measure
120 (40 Nodule,80 Non Nodule)	93	0.90	0.90	0.95	0.06	0.9
160 (86 Nodule,74 Non Nodule)	94	0.94	0.94	0.94	0.05	0.94
200 (115 Nodule,85 Non Nodule)	93	0.94	0.93	0.92	0.06	0. 93

The various parameters considered for performance analysis such as accuracy, precision, recall and specificity give a value higher than 93%. The error rate is found to be less than 6%.

Although these values are less compared to the performance analysis parameters of SVM algorithm, the difference is hardly 2%. Form this, it can be concluded that unsupervised learning classification method can also be used and training data set is not always required for classifying the regions as nodules and non-nodules in given lung CT images. Further it can also be concluded that whenever training data set is not available the unsupervised classification can also be used very effectively for classification of lung CT scans.

Further, performance of system is measured by plotting Receiver Operating Characteristics (ROC) curve for Fuzzy C-Means algorithm which is shown in the Figure 3.12. It can be observed that curves are nearing one showing performance of the system is better.

Classification results show good acceptability by adopting both the algorithms since, area under the curve is nearing unity. Also, accuracy plots are drawn for the purpose of comparison. Figure 3.13 shows the plot of the parameter values of two classification

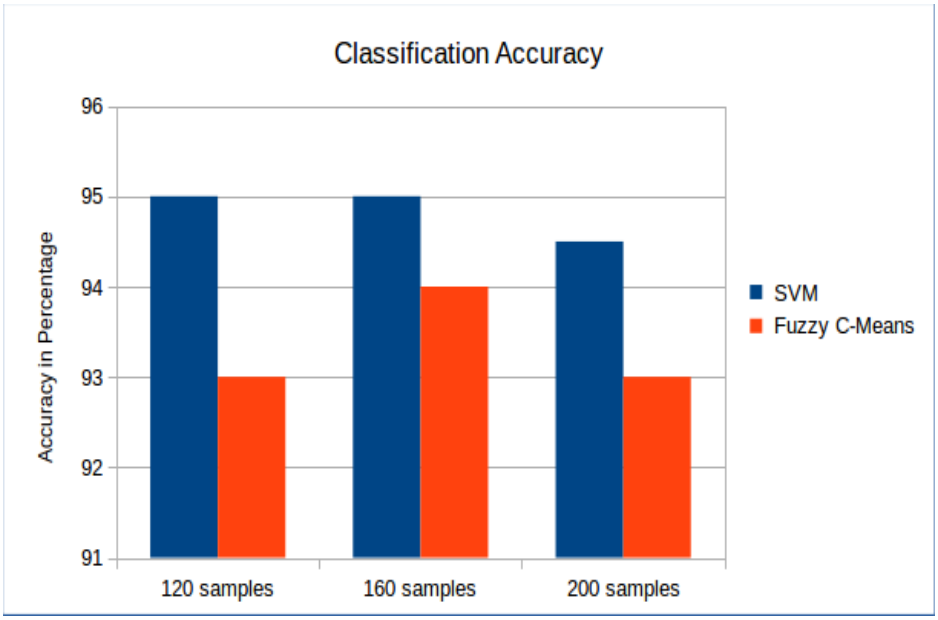


(a)

Figure 3.12 ROC curve for Fuzzy C-Means algorithm

methods versus the number of samples. The plot in Figure 3.13 shows consistent results for both the methods with a slight increase in case of supervised classification.

Random Forest Algorithm: It is an ensemble learning method and consists of two stages. They are Random forest creation and Random forest prediction. Random forest creation is the training phase of the system which generates multiple decision trees. Randomly *n* features are selected from total of *m* features each time. Among the randomly selected *n* features a best split is identified and accordingly the tree is built (Breiman, 2001). Random subset of features are selected at each split point in the learning process. This is called feature bagging. This procedure is repeated many times and several *N* trees are obtained which forms the random forest. Best split for attribute selection is determined by using Information Gain or Gini Index. Information Gain determines the information each attribute is holding and is in turn calculated by Entropy measure which is a measure of uncertainty of a variable. Gini Index defines how often a randomly chosen element would be incorrectly identified. Hence, an attribute with lower gini index is always preferred.



(a)

Figure 3.13 Plot showing classification accuracy

Random forest prediction is performed using these randomly created decision trees. The test data (features) is passed through these trees and prediction is stored on the leaf nodes of every tree. The classification of the input is defined based on the votes. High voted predicted target will be the final outcome of the random forest. Output of the system is the count of each class.

Prediction error is evaluated by using out-of-bag error (OOB) estimate which uses bootstrap aggregating for further sampling of the data used for training. OOB error is calculated for individual class average of this estimate (Breiman, 2001). Random forest method of classification consists of a large number of deep trees.

The sixteen GLCM features calculated considering 120 randomly selected images are given as input features to random forest algorithm in the training phase. The features of 80 images are treated as test data. The OOB error of prediction is presented in Figure 3.15. It is observed that as the number of trees increases, the error estimate decreases and thereby better estimates are obtained from out-of-bag predictions. The advantage of Random Forest algorithm is that it prevents over fitting by selecting minimum number of features each time to predict the outcome. The results in the form of confusion matrix obtained by applying the random forest algorithm is shown in Figure 3.14. 80 images

are taken as test data out of which 6 are found to be falsely predicted.

TP	FP
46	2
FN	TN
4	28

Figure 3.14 Confusion Matrix

The performance metric accuracy, specificity and sensitivity are computed using the values of confusion matrix. Accordingly, accuracy of prediction is found to be 92.5%. Specificity and sensitivity of the system is observed to be 93.3% and 92%, respectively.

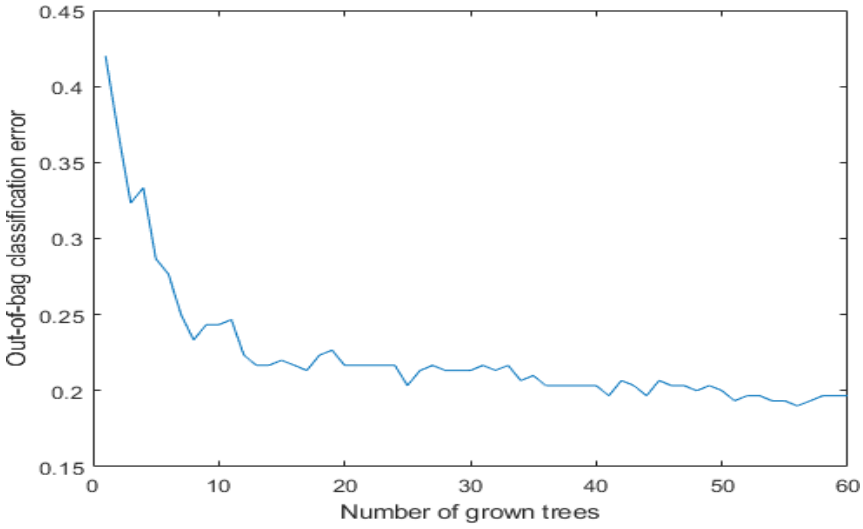


Figure 3.15 Out-of-Bag Error result of random forest classification algorithm

It is observed that average accuracy of SVM algorithm is 0.95, Fuzzy C Means algorithm is 0.93 and Random Forest is 0.92. Similarly, specificity and sensitivity of SVM, Fuzzy C Means algorithms are in the order of 0.94. But, specificity and sensitivity of Random Forest method is 0.92. Error rate presented in all the cases is found to be less than 5 percent. It can be concluded that unsupervised classification using Fuzzy-C method gives consistent results compared to supervised classifications of both SVM and Random Forest method making the CAD system reliable. System is validated by

performing 10 fold cross validation for 200 (115 nodules, 85 non-nodules) images and standard deviation values thus obtained are given in Table 3.11.

Table 3.11 Standard Deviation values of SVM, Fuzzy C-Means and Random Forest Algorithms obtained after performing 10-cross validation

Models	Standard Deviation σ		
	Accuracy	Sensitivity	Specificity
SVM	1.30	1.60	1.50
Fuzzy C-means	1.72	1.85	1.70
Random Forest	1.40	1.74	1.53

Thus the method adopted in Phase I includes removal of noise from the raw data obtained by CT images followed by segmentation and classification of identified regions as nodules and non-nodules. The process considered so far clearly separates nodule and non-nodule regions with percentage of accuracy more than 90% and the error associated is less than 7%. The nodules so determined are further processed for categorizing solid and sub-solid nodules. Thus the nodule region identified in Phase I forms the basic data for further classification of the sub-solid nodules which is the main focus of the present study.

3.3 Phase II

Identification of solid and sub-solid nodules

Detection of sub-solid nodules from already identified nodules needs identification of relevant features prominent and pertaining to sub-solid nodules and extraction of these features. This is followed by reclassification of the features into solid and sub-solid nodules.

The process involved is presented as Phase II in Figure 3.1 which consists of identification and extraction of features followed by classification into solid and sub-solid nodules. Histogram of Gradients (HoG) features is adopted for identification and extraction of relevant features of sub-solid nodules. Classification is achieved by adopting

the supervised learning model, SVM and unsupervised learning model, K-Means algorithm.

3.3.1 Feature selection - HoG Features

Since the focus is on separating the sub-solid nodules from already identified nodules, features prominent in the sub-solid nodules need to be identified. Then separation of solid nodule and sub-solid nodules will have to be carried out on the basis of the identified prominent features. Generally adopted methods for feature selection are once again based on statistical measures. In the process, the information present may or may not be considered (Taşcı and Uğur, 2015). To that extent these methods are limited.

However, in the present analysis, Histogram of Gradient (HoG) method is adopted for feature selection. HoG is a feature descriptor basically used for detecting objects in an image and extracts features from all locations in identified region of interest in an image, which is usually not the case in the other algorithms (Dalal and Triggs, 2005). HoG features are identified and extracted from the already detected nodule regions.

Information of gradients is obtained by dividing the image into small cells of 8x8 pixels and blocks of 16x16 pixels (i.e. 2x2 cells). Each cell has nine gradient orientation bins. The histogram ranges from 0 to 180 degrees (20 degrees per bin). Every pixel in the selected cell vote for a gradient orientation bin. Sometimes the pixel votes selects two bins which is referred to aliasing. Hence votes are bi-linearly interpolated to avoid this confusion. Later, normalization of histograms is accomplished with the help of their energy (regularized L2 norm) across blocks.

An identified cell generally belongs to four blocks since every block has a step size of one cell which defines four different normalized versions of the cell's histogram. These identified histograms are then concatenated to get the descriptor histogram. Descriptor histogram usually contains all the features for that cell. Intensity of larger portion of image called block is calculated. It is used to normalize the local histograms (cell histogram) to improve the accuracy.

Generally used feature extraction descriptors are Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF). SIFT algorithm identifies the main points in the objects present in images which is stored as reference points. When a

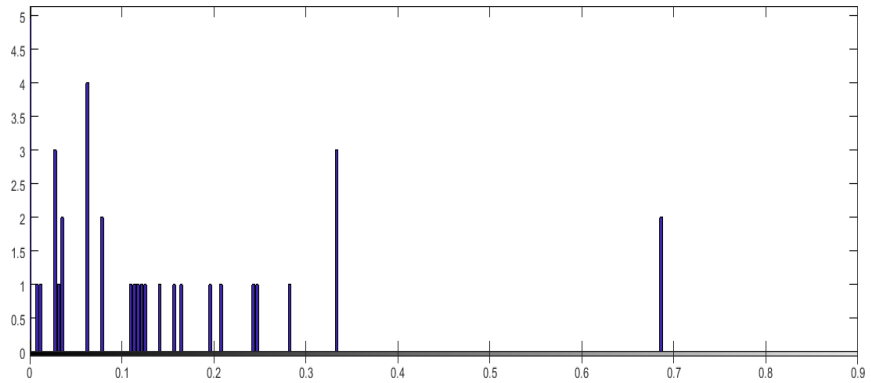
new test data is given as input, SIFT algorithm identifies the object in it by using the reference key points stored in the database. Object, location, scale and orientation of key points and objects in the test image are filtered to find the best suitable match. SURF is an advanced version of SIFT algorithm which works much faster than SIFT algorithm. SURF and SIFT algorithms consider only the key points and store it in database with the help of which they identify the objects in the given image. This is a major drawback since not all the pixel values are considered and hence there are chances of missing out minute variations in the image. This leads to a less accurate system since there will be more number of false predictions by the system.

HoG descriptor has certain advantages compared to SIFT and SURF feature descriptors as it takes into account every pixel present in nodules. Since it operates on local cells, geometric and photometric transformations do not effect the descriptor. All the features are calculated from the entire region. Even minute variations are captured in the histogram because of overlapping concept adopted in the process. Thus HoG feature descriptor is found to be best suited since accuracy achieved by this method is high. HoG features are used as it takes account of texture, both in magnitude and orientation along with statistical measures.

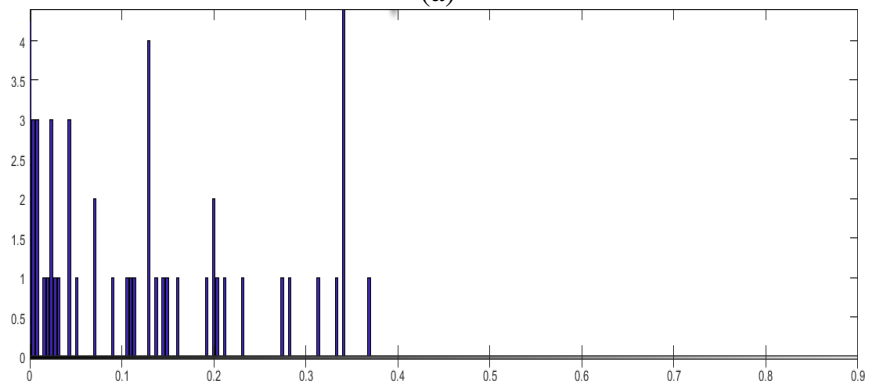
Four histograms of a cell are plotted since a cell can occur in four blocks (Figure3.16). Normalizing these histograms, a combined histogram representing the overall feature vector of that cell is obtained which is shown in Figure 3.17(a). A Visualization of HoG features has been presented in Figure 3.17(b). Gradient and direction of HoG features are shown in Figure 3.17(c).

3.3.2 Classification of Solid and Sub-solid/Part-solid nodules

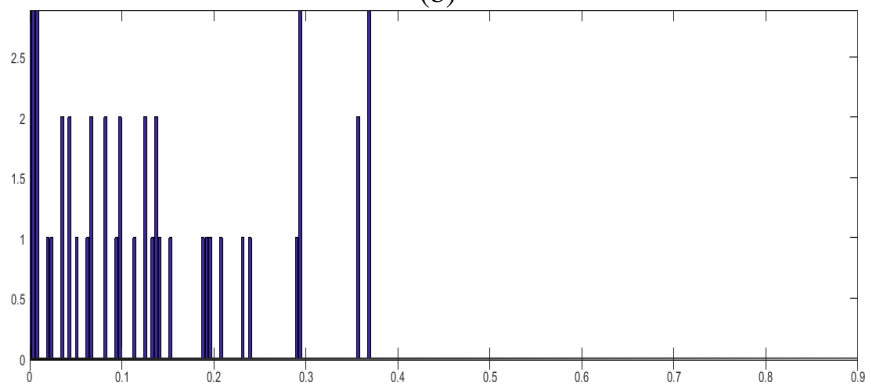
Classification of solid and sub-solid nodules is a two class problem. State-of-art classifiers such as Support Vector Machine and K-Means clustering algorithms are used for predicting the class labels. SVM constructs the hyperplanes with the help of training data set to classify the candidates. It falls into the group of supervised classification. Hence, training data plays a vital role in the classification process. However, unsupervised classifier, K-Means algorithm is modified accordingly to consider the angular



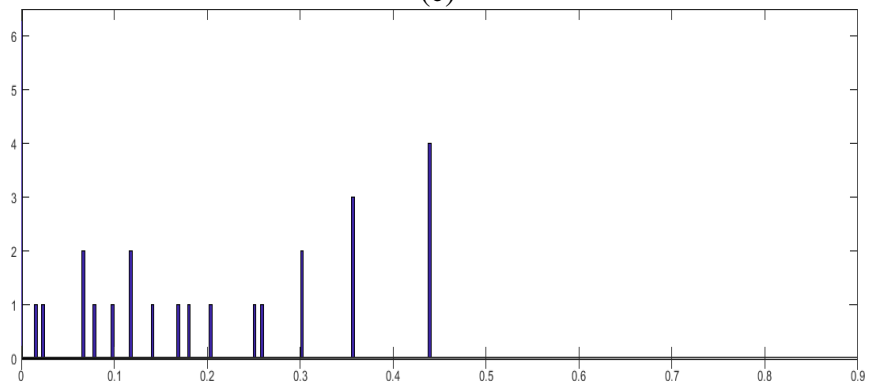
(a)



(b)

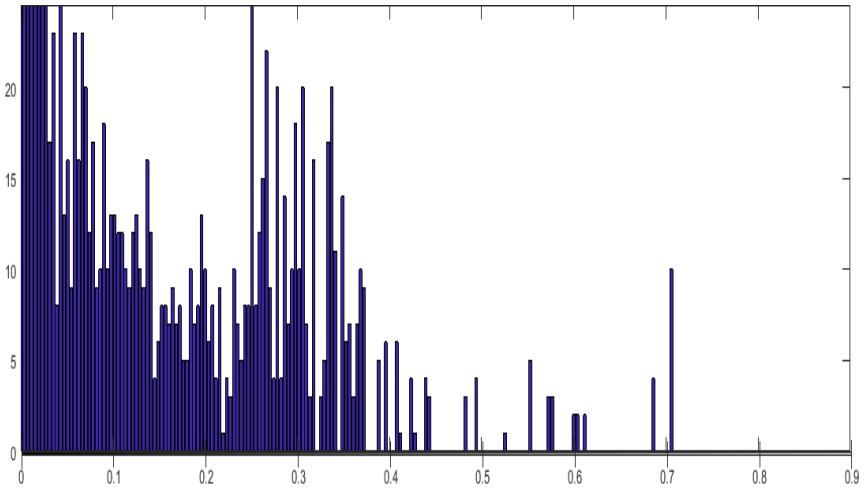


(c)

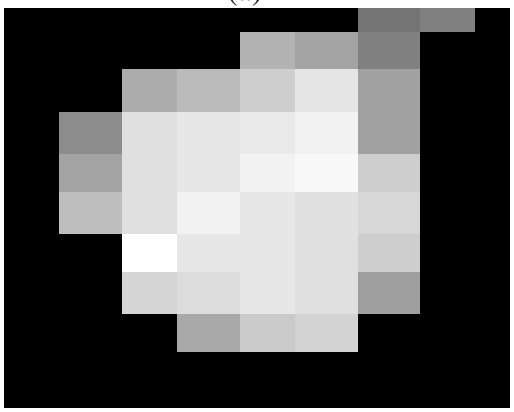


(d)

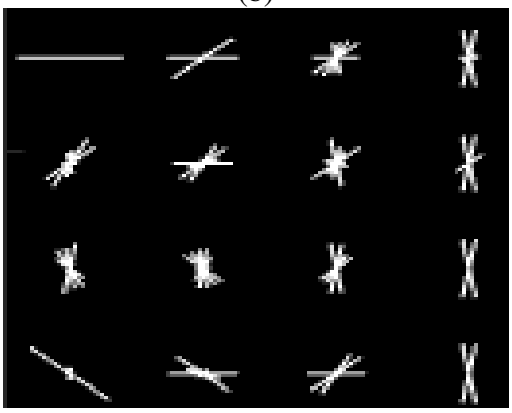
Figure 3.16 (a), (b), (c), (d) Histograms of a cell



(a)



(b)



(c)

Figure 3.17 (a) Normalized histogram of image (b) Visualization of HoG feature (c) Gradient and direction of HoG Features

similarity and angular distance instead of usual Euclidean distance.

$$F(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} (\| x_i - y_j \|)^2, \tag{3.3.1}$$

where, $\| x_i - y_j \|$ is the Euclidean distance between x_i and y_j , c represents the number of clusters, and c_i represents the number of data points in ith cluster. Euclidean distance measure is replaced by angular similarity and angular distance metric obtained from HoG features for classifying the candidate nodules and is given as follows,

$$distance = \frac{\cos^{-1}(similarity)}{\pi}, \tag{3.3.2}$$

$$similarity = 1 - distance. \tag{3.3.3}$$

In the present study 105 nodule regions are identified which form the test data for identification of sub-solid nodules in Phase II of the CAD system.

To establish consistency of the system developed, the identified 105 nodule regions are randomly divided into sets of 60 and 105 images. Each set is classified using SVM and K-Means algorithms. Training data for SVM classifier is taken from annotated images available in the dataset. Accordingly 200 images with 110 solid nodules and 90 sub-solid nodules are considered as the training set. Table 3.12 and Table 3.13 show the results of SVM classification and its performance, respectively.

Table 3.12 Classification results of SVM

Number of Images	True Positive	True Negative	False Positive	False Negative
60 (35 Solid, 25 Sub-solid)	34	23	2	1
105 (65 Solid, 40 Sub-solid)	63	38	2	2

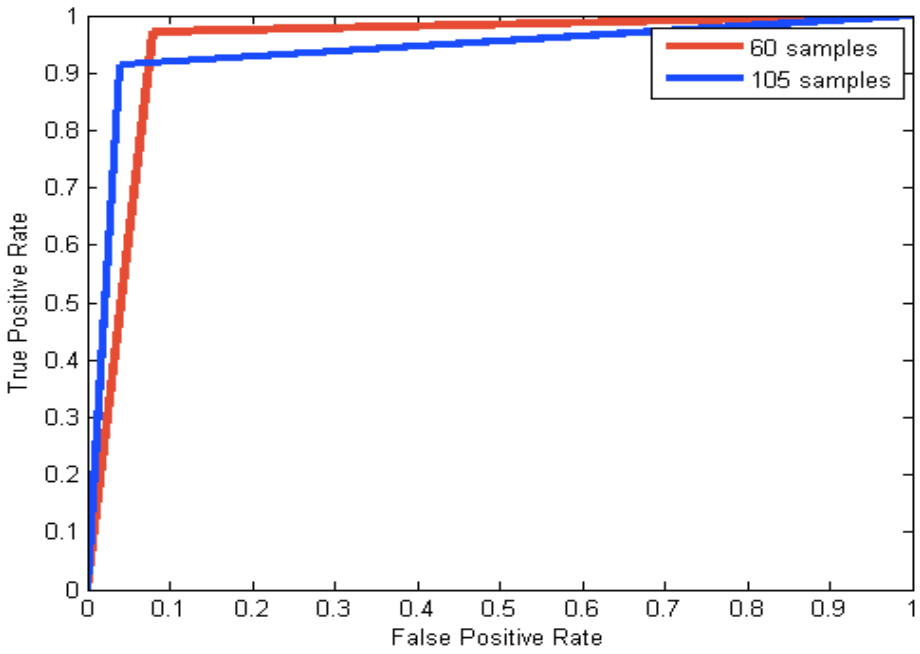
The total count of false predictions are 7 (False positives+false negatives) indicating that the system is reliable. The performance metrics such as accuracy, precision, recall, specificity, error-rate, f-measure are calculated using the values of confusion matrix shown in Table 3.12. The performance metric values are tabulated in Table 3.13.

As observed from the Table 3.13, the accuracy measure is on an average 95%, specificity and sensitivity values are on an average 95% and 96%, respectively. From these

Table 3.13 Performance analysis of SVM algorithm

Number of Images	Accuracy (percent)	Precision	Recall	Specificity	Error-Rate	F-measure
60 (35 Solid,25 Sub-solid)	95.00	0.97	0.94	0.95	0.05	0.95
105 (65 Solid, 40 Sub-solid)	96.10	0.96	0.96	0.95	0.038	0.96

findings it can be established that system performs well in segregating the solid and sub-solid nodules. Further ROC curve is plotted for verifying the system performance and is shown in Figure 3.18. ROC curve is nearing unity which again states that system is accurate in identifying the solid and sub-solid nodules.



(a)

Figure 3.18 Receiver Operating Characteristics curve plotted for Support Vector Machine algorithm

Similarly, the results of classification carried out by K-means algorithm and its performance results are given in Table 3.14 and Table 3.15, respectively.

Table 3.12 and Table 3.14 clearly show that the number of false predictions are less. In the case of SVM classification, it is only 0.03 percent and in the case of K-

Table 3.14 Classification results of K-Means algorithm

Number of Images	True Positive	True Negative	False Positive	False Negative
60 (35 Solid, 25 Sub-solid)	32	24	1	3
105 (65 Solid, 40 Sub-solid)	62	37	3	3

Table 3.15 Performance analysis of K-means algorithm

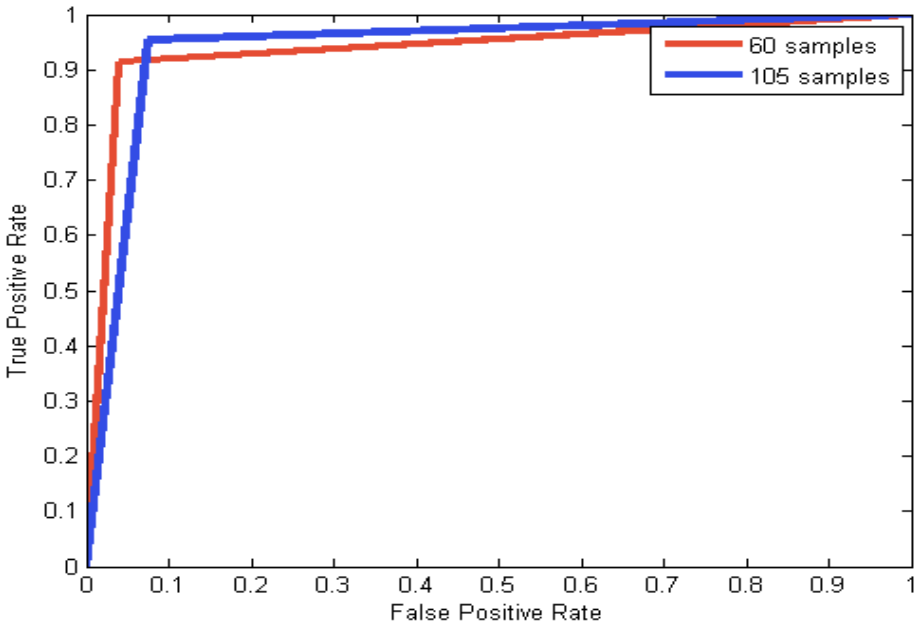
Number of Images	Accuracy (percent)	Precision	Recall	Specificity	Error-Rate	F-measure
60 (35 Solid,25 Sub-solid)	93.30	0.96	0.88	0.95	0.06	0.94
105 (65 Solid, 40 Sub-solid)	94.20	0.95	0.95	0.92	0.05	0.95

means classification it is only 0.04 percent, which shows that the results of classification algorithms are consistent. The overall accuracy is found to be 95 percent by using SVM classification and it is around 93 percent when K-Means classifier is adopted.

To measure the performance of the system, Receiver Operating Characteristics curves are drawn for the classification method considered here and the same is presented in Figure 3.19. It is observed from these graphs that the curves tends to 1 indicating that classification accuracy is high.

The comprehensive CAD system proposed herein, first identifies the nodules. The process includes consistency and reliability check at both segmentation and classification stages. The nodules so identified are further classified into solid and sub-solid nodules.

The performance of the proposed CAD system is compared with the existing CAD systems proposed by Messay et al. (2010), Jacobs et al. (2014) and Setio et al. (2016). The features used and corresponding accuracy, specificity and sensitivity values are given in 3.16, along with the values obtained for these parameters in the present study. It is clearly establishing that the present CAD system developed gives a much higher values for all the three parameters considered. Thus the present CAD system is better in identifying the sub-solid nodules in lung CT images compared to the state-of-the-art sub-solid nodules identification systems in the literature. Also, the standard deviation



(b)

Figure 3.19 Receiver Operating Characteristics curve plotted for K-Means Algorithm

values computed for these systems are given in Table 3.17.

Table 3.16 CAD Systems Comparisons

CAD Systems	Features used	Accuracy (percent)	Specificity	Sensitivity
Messay et.al(2010)	Size,Shape and Location	85	0.89	0.88
Jacobs Colin et.al.(2014)	Context Features	91	0.90	0.91
Francesco Ciompi et.al. (2016)	Neural network	80	0.50	0.64
Proposed System	Statistical Features and HoG Features	94	0.94	0.93

It can be seen that standard deviation for the three parameters considered in the proposed system is less than the standard deviation values obtained for the other CAD

systems. This test shows that the present system gives reliable results.

Table 3.17 Standard Deviation values for the four systems obtained after performing 10 cross validation test

Models	Standard Deviation σ		
	Accuracy	Sensitivity	Specificity
Messay et.al(2010)	1.90	1.56	1.58
Jacobs Colin et.al.(2014)	1.60	1.32	1.20
Francesco Ciompi et.al.(2017)	2.20	1.91	1.94
Proposed System	0.96	1.31	0.98

3.4 Summary

The proposed CAD system analyzes the lung CT images for identifying the sub-solid nodules in two phases. In the first phase it locates the nodule regions and in the second phase it identifies the sub-solid nodules in the already located nodule regions. The system attempts to improve the accuracy and reliability at every stage which ultimately improves the end result.

The first phase of the CAD system consists of preprocessing step for removing the noise by using NLTV method which in turn helps in improving the results of segmentation and morphological operations to identify the region of interest.

The CAD system is completely automated as it is able to detect region of interest in any given lung CT images without human intervention and makes use of Chan-Vese model and morphological operations to segment the nodular structures from the input images. The system further selects the most relevant statistical features based on Eigen values, thereby improving the accuracy of the prediction models in the system such as SVM, Fuzzy C Means, Random Forest compared to the existing models.

The second phase consists of further detecting the sub-solid nodules from nodule regions, for which the HoG features are identified. Deep texture variations are captured by HoG from which the sub-solid nodules are identified with the help of classification models such as SVM, K Means algorithms.

It is also demonstrated that the CAD system gives consistent results for different kinds of CT images from two deferent datasets. The developed CAD system is efficient in identifying the sub-solid nodules in lung CT images. Sensitivity of the CAD system in classifying the nodule and non nodule region (Phase I) using supervised SVM classifier is in the order of 95 percent and that of unsupervised Fuzzy C-Means classifier is 94 percent. Random Forest Method classification is found to give consistent results with reduced out of bag error. The methods adopted for both supervised and unsupervised classification of nodules are found to have error rate less than 6 percent.

SVM method for supervised classification and K-means algorithm adopted for unsupervised classification of solid and sub-solid nodules gives accuracy value in the range of 96 and 94 percent. Supervised and unsupervised classification method adopted in the study give consistent results, thus establishing that prior knowledge of nodule is not essential. The developed CAD system efficiently identifies the sub-solid nodules in lung CT images.

CHAPTER 4

Deep Learning Approach for Identifying the Lung Nodules in CT Images

4.1 Introduction

Machine learning is a part of artificial intelligence that helps the computers to automatically do a defined task. It makes the computer to learn by itself without being programmed explicitly. Machine learning algorithms learn from the observations or look for particular patterns in the input data which helps in decision making for predicting the output (Jordan and Mitchell, 2015).

Machine learning algorithms require structured data. In most of the real time cases the data is non-regular, unstructured and vast. Hence, the deep learning models are developed. It is a part of machine learning except for the fact that the input data could be unstructured. A deep learning architecture has numerous layers of neurons stacked together which does feature learning by itself. It finds applications in various fields like robot navigation, speech recognition, artificial intelligence etc.. (Liu et al., 2017).

CAD system presented in Chapter 3 rely on pipeline image analysis methods and handcrafted features. This includes a series of steps like preprocessing, identifying the region of interest, identifying and extracting features, classification using state-of-the art methods like SVM, Random Forest etc. The performance of each step in these systems depends heavily on the accuracy of the previous step which is a major disadvantage of this system. Heterogeneity of shape and texture of malign (part-solid, liquid) nodules make it difficult to tailor a particular handcrafted feature. Hence, there is a necessity to develop a system that eliminates all these procedures.

A Computer Aided Detection system using deep learning is developed in this context where raw (un-processed) lung Computed Tomography (CT) images are fed as input and appropriate outputs (identified part-solid lung nodules) are obtained. These deep networks learn deep features and identify the required object by using a series of convolution and deconvolution layers. Furthermore, a Conditional Random Field (CRF) framework is used in the present context to make the results of the developed deep learning model more robust and accurate.

4.2 Deep Convolution Neural Network (DCNN) Architecture

Deep learning approach is adopted here for identifying the part-solid nodules as they represent the early stage of lung cancer. The proposed architecture consists of a series of convolution and deconvolution layers as shown in Figure 4.1 along with the encoding and decoding layers. The contracting path extracts the features required to identify the part-solid nodules whereas the expansion path groups similar pixels together so as to identify the part-solid nodules. The contracting path is made of fully connected convolution layers and expansion path is made of fully connected deconvolution layers.

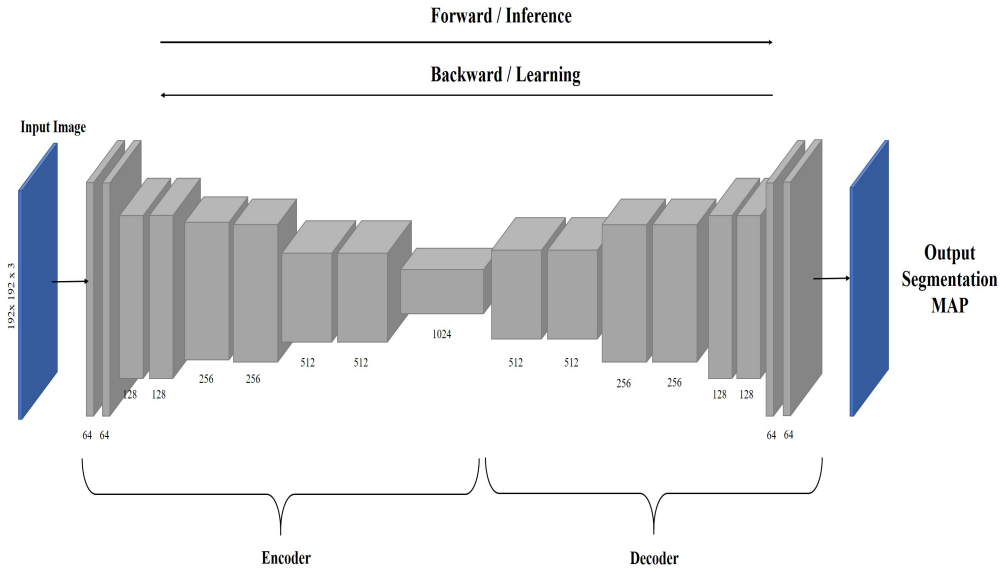


Figure 4.1 Block diagram of Deep Convolution Neural Network (DCNN) Architecture

A set of raw RGB (Red,Green,Blue) CT images with the size $196 \times 196 \times 3$ is given

as input to the DCNN architecture. This input image is passed through the developed architecture which consists of 18 layers. First nine layers work on convolution operation along with activation function called Rectified Linear Unit (ReLU) and the next nine layers work on deconvolution operation to get back the original images. The convolution operation is performed by using a 3×3 kernel. Initially, training weights are initialized using Gaussian distribution. In the later stages, the weights for these convolution filters are learned during the training phase.

Activation function is employed in each layer that helps in the neuron's output decision making process in a neural network. Mapping of input to different outputs is handled by an activation function. Sigmoid, tanh, Rectified Linear Unit (ReLU), softmax are some of the the popular activation functions used in deep learning models. Sigmoid activation function uses sigmoid function to map the values in the range of 0 and 1. Tanh function restricts the output of a neuron to be between -1 and 1. Rectified linear Unit restricts the values to be in between 0 and infinity. Softmax activation function converts the input of a neurons to probabilistic values.

The developed deep learning model performance boosts when, ReLU activation function is employed in each layer to map the output of a convolution operation from zero to infinity which is given by the following equation,

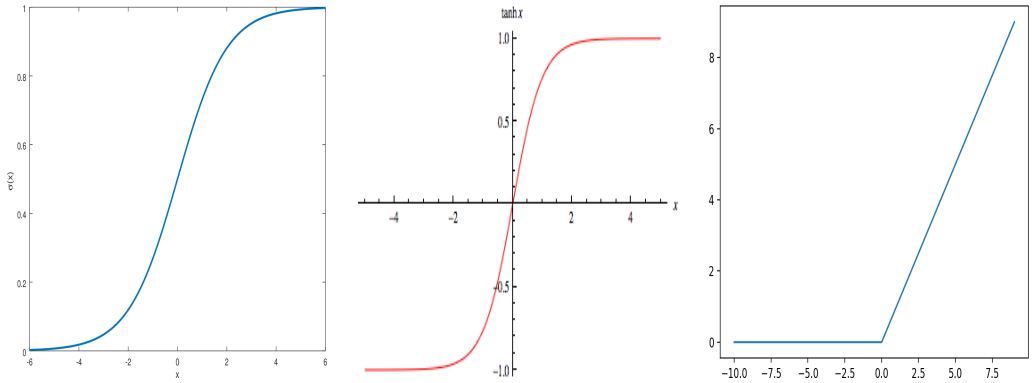
$$Y = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{4.2.1}$$

Activation functions such as sigmoid and tanh are less used since it causes vanishing gradient problem (Agostinelli et al., 2014). Vanishing gradient problem occurs when gradient has a very small value and network finds it difficult to train. When sigmoid or tanh function value is either too high or low, their derivative will be small. This makes the results biased which in turn makes the system less accurate. Hence, Rectified linear Unit activation function is preferred over other activation functions since it always returns the values zero or greater than zero (Agostinelli et al., 2014). The curves of activation functions sigmoid, tanh and ReLU are shown in Figure 4.2.

$$sigmoid(x) = \frac{e^x}{e^x + 1}, \tag{4.2.2}$$

(4.2.3)

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}, \tag{4.2.4}$$



(a) (b) (c)

Figure 4.2 (a) Sigmoid function (b) Tanh function and (c) ReLU function

The images in every layer are padded so that features are preserved for further layers. The step size of the convolution filter is defined by stride function which regulates the movement of filters in a pixel-wise operation in the image. It is considered as stride$(2,2)$ in the proposed architecture. Hence, in every iteration convolution filter slides by two pixels. Dropout layers are adopted in all the layers to neglect the randomly selected features during the training phase so as to avoid over-fitting problem. Localization of nodule regions are obtained through expansion path. Convolution operation with activation function ReLU are applied to all the layers of expansion path. The features extracted by the encoders are given as input to the decoders. The decoding path performs upsampling operation followed by a set of convolution operations. The upsampling operation increases the image size by duplicating the rows and columns. Convolution operations are adopted to obtain the group of similar pixels. Convolution filters learn to group the pixels by utilizing forward and backward propagation algorithm. This operation is repeated multiple times to rescale the image to it's original size. This process eventually achieves the localization of part-solid nodules. The last layer consists of softmax operation. Softmax function converts the input numbers into a vector of probabilities which when added sum up to one. The developed deep learning model has to find the part-solid nodules from the given input images. Hence, the soft-

max function gives the probability of each pixel belonging to part-solid nodule and the target pixel has the highest probability. Softmax function is defined by the following expression,

$$\sigma(q)_j = \frac{e^{q_j}}{\sum_{t=1}^{T} e^{q_t}} \qquad \text{for } j = 1, ..., T \tag{4.2.5}$$

Softmax function is an exponential function where T dimensional vector of q arbitrary values is reduced to T dimensional vector of $\sigma(q)$ real values. In the present context, T corresponds to the number of classes C. The model classifies the pixel either as part-solid nodules or non-nodule regions. Hence, the class label set is defined as $L = \{1, \ldots C\}$ where, the number of classes $C = 2$ i.e. part-solid nodule or non-nodule.

The learning rate determines the change or updation in weight and is an important hyper parameter used in the training phase. These parameters are responsible for identifying the most appropriate weights for a model and making it accurate. During the training phase, the parameters which represent the weights of the convolution filters is tuned in order to identify the part-solid nodules accurately and is handled by back propagation algorithm. This consequently reduces the error. The error is calculated at the output layer as $(Error = y_i - a)$ where, y_i is the predicted output and a is the actual output. This error is minimized using the gradient descent algorithm given by $\frac{\delta Error}{\delta o_j}$, where o_j represents the output of jth neuron. Over the training iterations, parameters are tuned accordingly with the help of loss function and optimizer.

Loss functions are used for calculating the error in prediction in the training phase of the network. Based on the values obtained by loss function the network parameters are tuned appropriately for improving the results in next cycle. Generally used loss functions are cross entropy, binary cross entropy, mean square error (Janocha and Czarnecki, 2017). The selection of loss functions helps in reducing the prediction error and there by attains better convergence. The proposed model employs categorical cross entropy in all the layers which is given as follows,

$$-\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} 1_{y_i \in L_c} log(P_{model[y_i \in L_c]}) \tag{4.2.6}$$

where, there are i observations and C categories. The term $1_{y_i \in L_c}$ is the function of

*i*th observation belonging to the *c*th category. The term $log(P_{model[y_i \in L_c]})$ predicts the probability of *i*th observation belonging to *c*th category. L_c represents the set of class labels.

Loss functions always work with optimizer which are used to tune the parameter such that the loss of the model is reduced. They together describe the protocol to change the parameter (Schneider et al., 2019). Several optimizers exist such as Stochastic Gradient Descent, Adam, RMSprop, Adagrad etc. and the one used in the model developed here is Stochastic Gradient Descent (SGD) optimizer. Using the complete dataset (samples) every time while performing the gradient descent makes the system more complex and time consuming. SGD optimizer uses a single sample from the huge set of samples which makes the system less complex. Hence, a SGD is preferred over the other optimizers (Ruder, 2016). Each time a sample is selected randomly from the dataset. In addition to these, dropout function is applied to avoid over-fitting which randomly masks few neurons before back propagation. Over-fitting is a situation wherein the model performs accurately on training data but fails to perform well on test data. This is generally avoided by the selection of different regularization layers such as batch normalization, drop-out etc.. Drop-out layers remove few connections between layers, thus allowing the model to predict the label with less number of neurons. The masked neurons are not used for updating the parameters.

Pooling layers are utilized to extract features in CNN. The different pooling layers available are Maxpool, average pool, stride etc. Among these, generally maxpool operation is most popular (Agostinelli et al., 2014). It is mostly adopted in the CNN architecture to extract prominent feature from a local region. It extracts the largest value from the considered region of a feature map, there by extracting most important features in the deeper layers. The Average pooling layer computes the average of values present in the neighborhood and replaces the neighborhood pixels by the average value. Stride operation is responsible for controlling the filter convolution around its neighbors. In this context the value is set to 2. Hence convolution kernel shifts by 2 pixels. The application of pooling layers results in the reduction of spatial resolution and loss of detailed information.

The layers of the deep neural network are given in Table 4.1. A detailed procedure of the working of deep neural network is depicted as pseudocode which is given below,

Algorithm 1: DCNN Training Algorithm

Input: K input images
Output: Trained DCNN parameters θ

for *Each image* $m \times n \times t$ **do**
  while *epoch* $s : 1 \rightarrow S$ **do**
    while *Training samples* $j = 1 \rightarrow K$ **do**
      Convolution operation performed to get feature maps
      Calculate Softmax activation function using Equation
      Calculate Error $(Error = y_i - a)$
      Compute Gradient $\frac{\delta Error}{\delta o_j}$ by back propagating the error
      The variable $\theta = \theta - \Delta c_{ij}$ is updated adopting the gradient descent
      $\Delta c_{ij} = -\eta(\frac{\delta Error}{\delta c_{ij}})$ where η is the learning rate
    end
  end
end

4.3 DCNN CAD system

A Computer Aided Detection system using deep learning approach is developed for identifying the part-solid or sub-solid nodules. Raw lung Computed Tomography images are given as input and appropriate output (identified lung part-solid nodules) is obtained. The process consists of localizing the potential region of interest (part-solid nodules) by using the deep learning model which is incorporated within the CRF framework as a unary potential along with pairwise potential. These deep networks learn deep features by itself and identify the required object by using a series of convolution and deconvolution layers. Conditional Random Field framework is used for making the developed deep learning model results more accurate. The reliability of the proposed system is substantiated by analyzing the results. The block diagram of the proposed model is given in Figure 4.3.

Table 4.1 Parameters of DCNN

Encoder				Decoder			
Layer	Type	Input	Output	Layer	Type	Input	Output
	Input	192×192×3		10	Fully connected	12×12×1024	24×24×1024
1	Fully connected	192×192×3	192×192×64	11	Fully connected	24×24×1024	24×24×512
2	Fully connected	192×192×64	96×96×64	12	Fully connected	24×24×512	48×48×512
3	Fully connected	96×96×64	96×96×128	13	Fully connected	48×48×512	48×48×256
4	Fully connected	96×96×128	48×48×128	14	Fully connected	48×48×256	96×96×256
5	Fully connected	48×48×128	48×48×256	15	Fully connected	96×96×256	96×96×128
6	Fully connected	48×48×256	24×24×256	16	Fully connected	96×96×128	192×192×128
7	Fully connected	24×24×256	24×24×512	17	Fully connected	192×192×128	192×192×64
8	Fully connected	24×24×512	12×12×512	18	Softmax	192×192×64	192×192×3
9	Fully connected	12×12×512	12×12×1024				

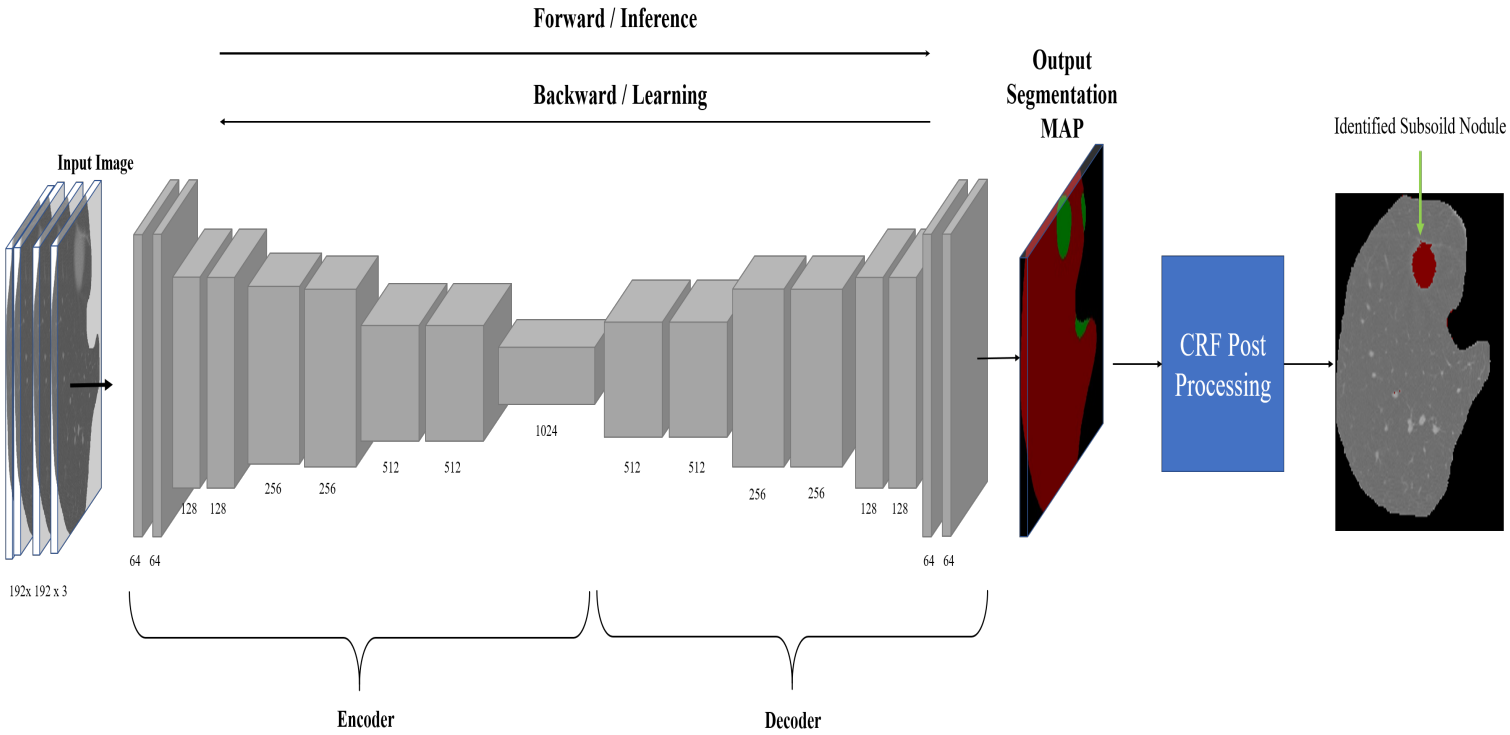


Figure 4.3 Overall system block diagram

4.3.1 Conditional Random Field (CRF) Model

Conditional Random Field is a probabilistic approach for semantic segmentation. A conditional distribution $P(A|B)$ is identified where $A = a_1, a_2, ..., a_N$ represent output random variables and $B = b_1, b_2, ...b_N$ represent the input variables. CRF consists of unary potential and pairwise potential energies. Unary potential energy $\varphi(x_i)$, depends on pixel itself while pairwise potential energy $\lambda(x_i, x_j)$, depends on neighboring pixels. In the present context given an image U of size $m \times n$, the conditional probability of the class labels L (part-solid nodule and non-nodules) is computed as follows,

$$P(L|U) = \frac{1}{Z} exp^{E(U)}, \tag{4.3.1}$$

where,

$Z =$ is the partition functional, and

$E(U) =$ is the Gibbs energy.

Gibbs energy $E(U)$ given as,

$$E(U) = \sum_{x} \varphi(U_x) + \sum_{(x,y)\in\varepsilon} \lambda(U_x, U_y), \tag{4.3.2}$$

where,

$\varepsilon =$ edges connecting the four neighboring pixels in a grid structure.

The result of DCNN is used as unary potential values in CRF model. The pair wise potential considered here is color dependent smoothness term and is given as,

$$W(U)_x = \frac{1}{W_p} \sum_{y\in S} G_{\sigma_s}(\|x-y\|) G_{\sigma_r}(U_x - U_y), \tag{4.3.3}$$

where,

$G_{\sigma_s} =$ Gaussian function which defines the influence of neighboring pixels,

based on spatial distance $\|x-y\|$, and

$G_{\sigma_r} =$ Gaussian function which defines the intensity range between,

the pixels $(U_x - U_y)$.

$$G_\sigma(d) = \frac{1}{2\pi\sigma} exp^{(\frac{-d^2}{2\sigma^2})}, \tag{4.3.4}$$

where,

$d =$ may be spatial distance $\|x-y\|$ or intensity difference $(U_x - U_y)$, and

$W_p =$ normalization factor.

Normalization factor W_p is given as,

$$W_p = \sum_{q \in \varepsilon} G_{\sigma_s}(\|x-y\|) G_{\sigma_r}(U_x - U_y). \tag{4.3.5}$$

Once the parameters are learnt by the CRF model the class labels are inferred using alpha expansion graph cut algorithm (Boykov and Jolly, 2001). The most probable class label is obtained by maximizing the conditional probability defined in Equation 4.3.1, which in turn is achieved by minimizing energy function defined in Equation 4.3.2.

4.4 Results

The sample original Lung CT images considered in the present discussion are given in Figure 4.4.

The current section presents the results obtained for semantic segmentation of nodules present in lung CT images.

From the available dataset 80% of the data is used for training, 10% for validation and 10% for testing. The developed CNN model is trained for 100 epochs and batch size is set to 10 due to memory constraints. Categorical cross entropy is used as the loss function to tune the parameters and initial weights are assigned based on Gaussian distribution.

The DCNN model is trained for 100 epochs. Corresponding accuracy and loss curves for training and validation sets are shown in Figures 4.5 and 4.6, respectively. From the graph it is observed that loss curves reduces non-linearly as the number of epochs increases and the loss is below 0.2 at 20 epochs. The model accuracy for training and validation set is above 90% as shown in Figure 4.5. Drop out layers are included in order to prevent over-fitting of the model which consists of 18 million parameters.

The developed Deep Convolution Neural Network architecture is used to extract

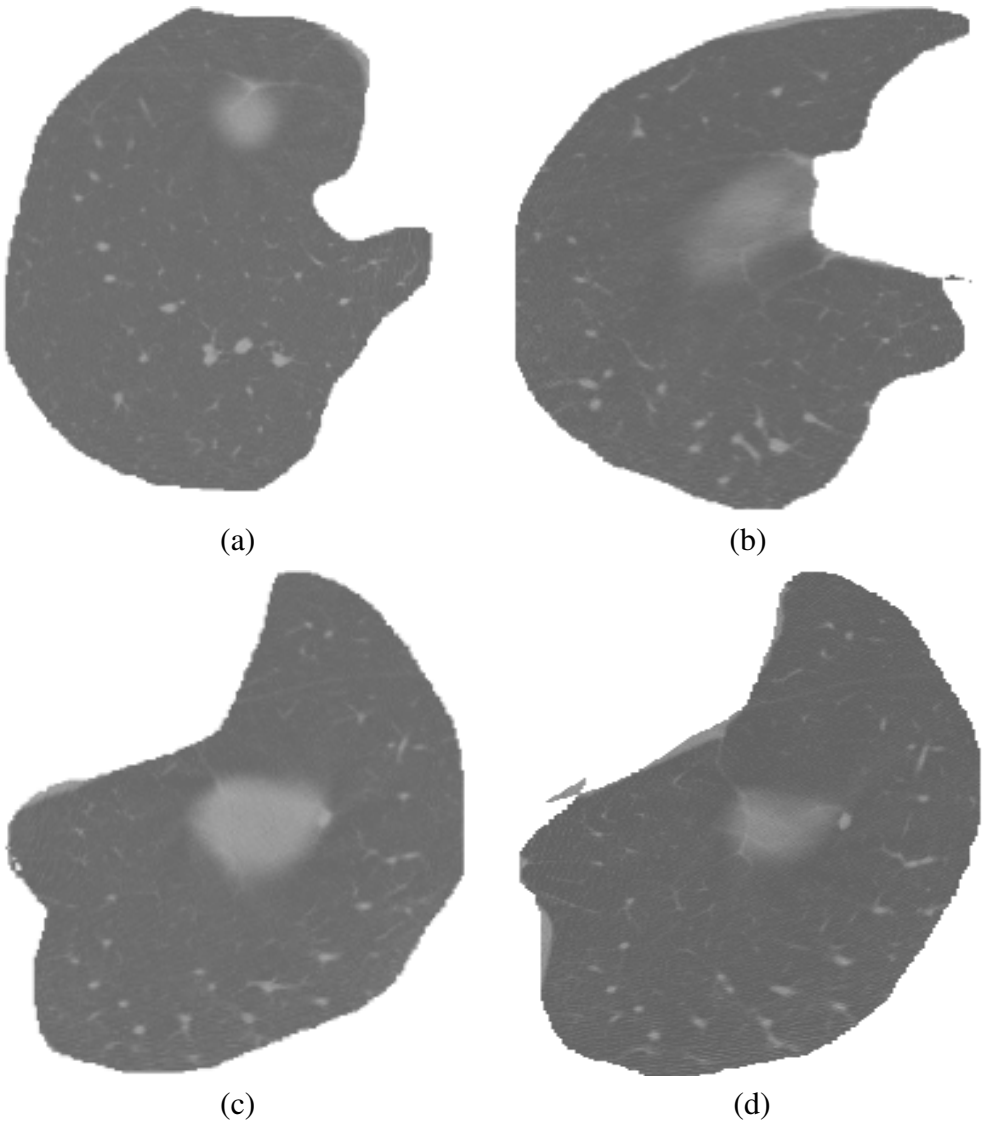


Figure 4.4 (a), (b), (c) and (d) are Sample Lung CT images from LIDC database.

deep features from lung CT images and subsequently, identify part-solid nodules. Performance evaluation of maxpool and stride operations for feature extraction is carried out here. The results obtained by considering both the operations show that, maxpool leads to loss of resolution and information which reduces the accuracy (84%) when compared to the stride operation (88%). Hence stride operation is considered for the analysis.

A few sample images and corresponding segmentation output of DCNN are shown in Figures 4.8 and 4.9, respectively. It is seen that the model identifies sub-solid nodules with higher accuracy along with a few false positives. These ambiguities occur at the

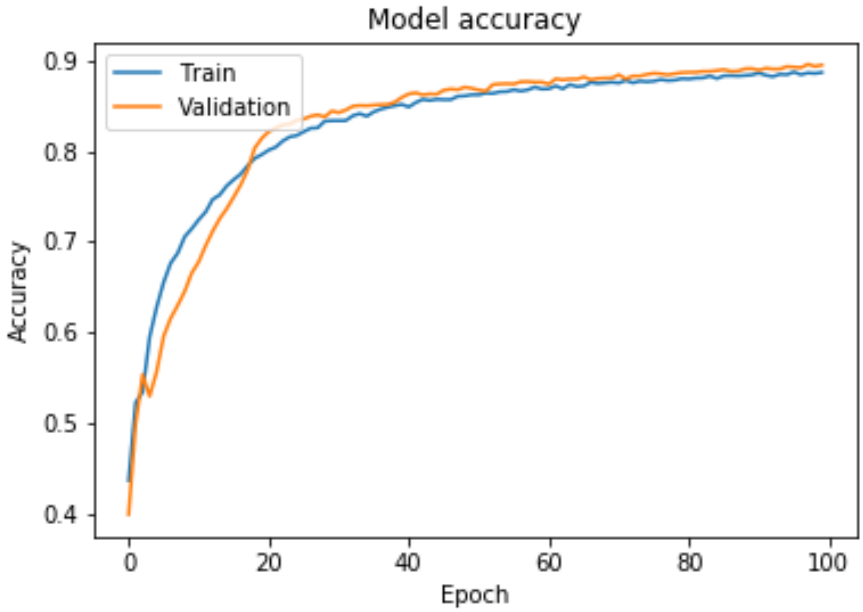


Figure 4.5 Accuracy curve for training and validation set

boundaries of the lungs. Hence post-processing is done to eliminate the false positives and it is accomplished by utilizing the CRF framework. False positives are removed by the usage of the pairwise term which smoothen the segmentation output based on color and spatial distance. The results obtained before and after using the Conditional Random Field algorithm are shown in block diagram 4.7.

The results obtained for the proposed model are depicted in Figures 4.8 and 4.9. In both the figures, original sample images considered are shown in the first row, while second-row depicts the output obtained from Deep Neural Network and third row shows the output obtained by adopting CRF + DCNN. It is seen that the CRF post-processing helps in reducing the occurrence of false positives that are observed in the output obtained directly from deep convolution neural network. The results of CRF + DCNN has improved compared to the application of DCNN alone, by eliminating false positives.

The performance evaluation of the deep learning based CAD system is computed using Mean Intersection over Union (MIoU), Pixel Accuracy (PA), Precision, Recall and F1-score.

The Intersection over Union (IoU) is defined as overlapping of target output and the

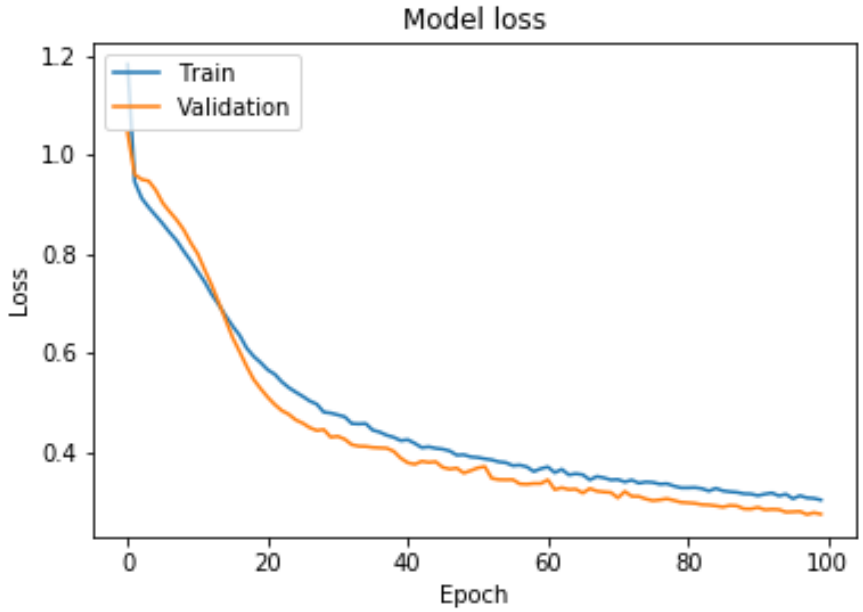


Figure 4.6 Loss curve for training and validation set

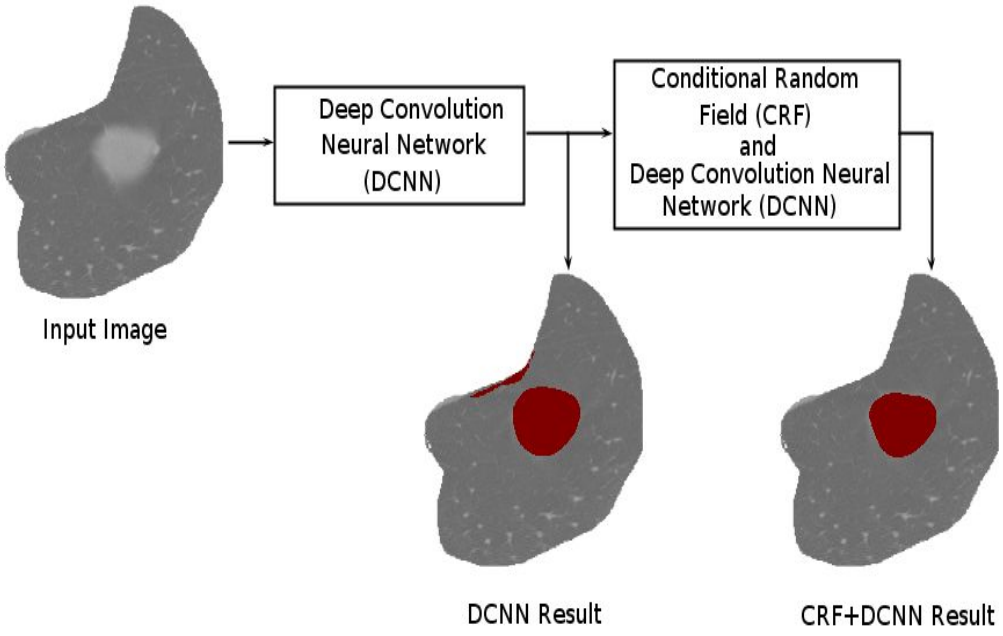


Figure 4.7 Block diagram showing results obtained before applying the CRF algorithm and after adopting the same

predicted output. The individual IoU value is computed for each class which is later averaged for all the classes to obtain a single global value. This value is called as Mean Intersection over Union. The equation for computing the same is given as,

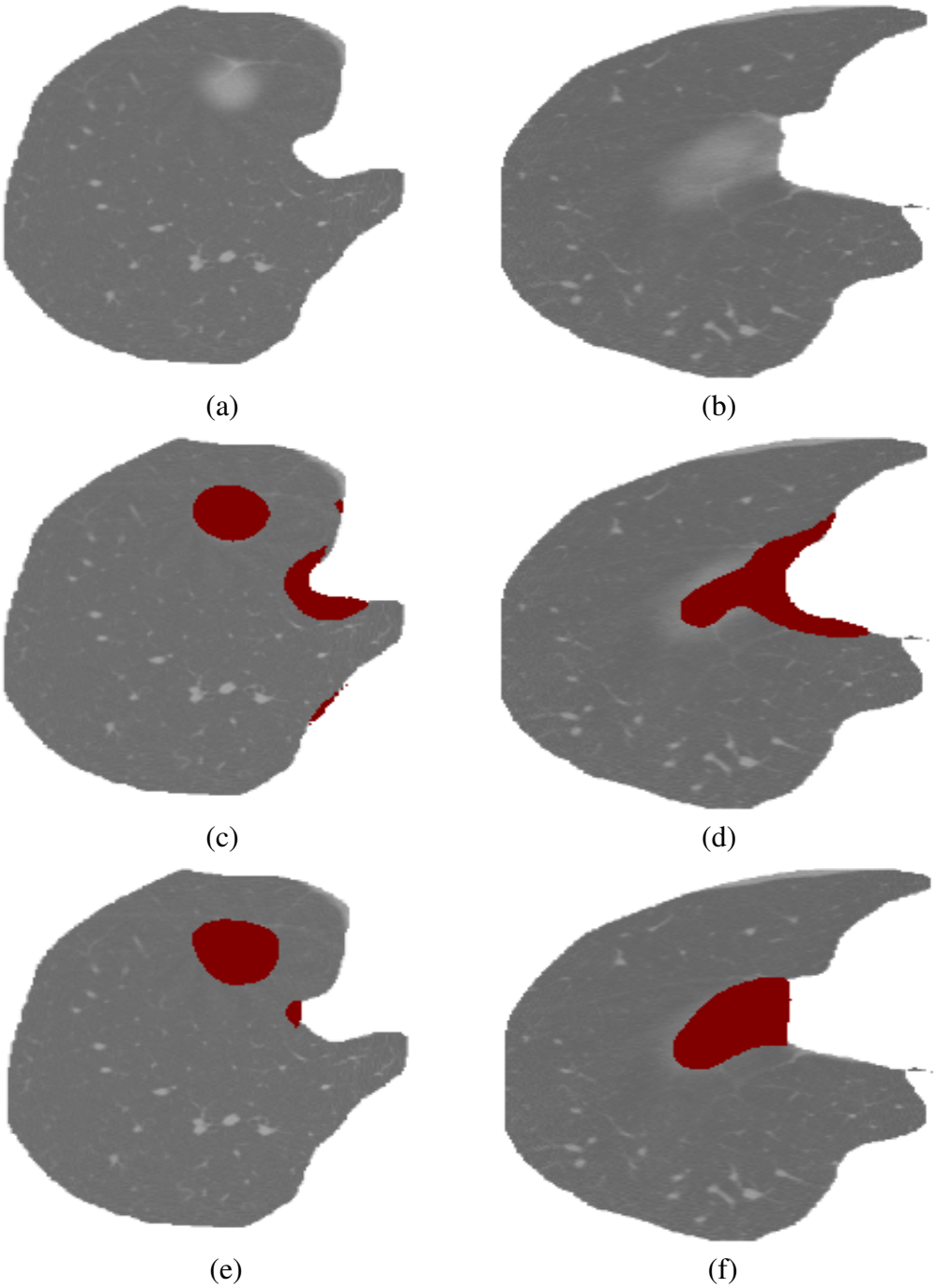


Figure 4.8 (a), (b) Lung Computed Tomography images showing part-solid nodules. (c), (d) Identified Part-solid nodule by the proposed deep learning approach. (e), (f) Results of CRF + DNN

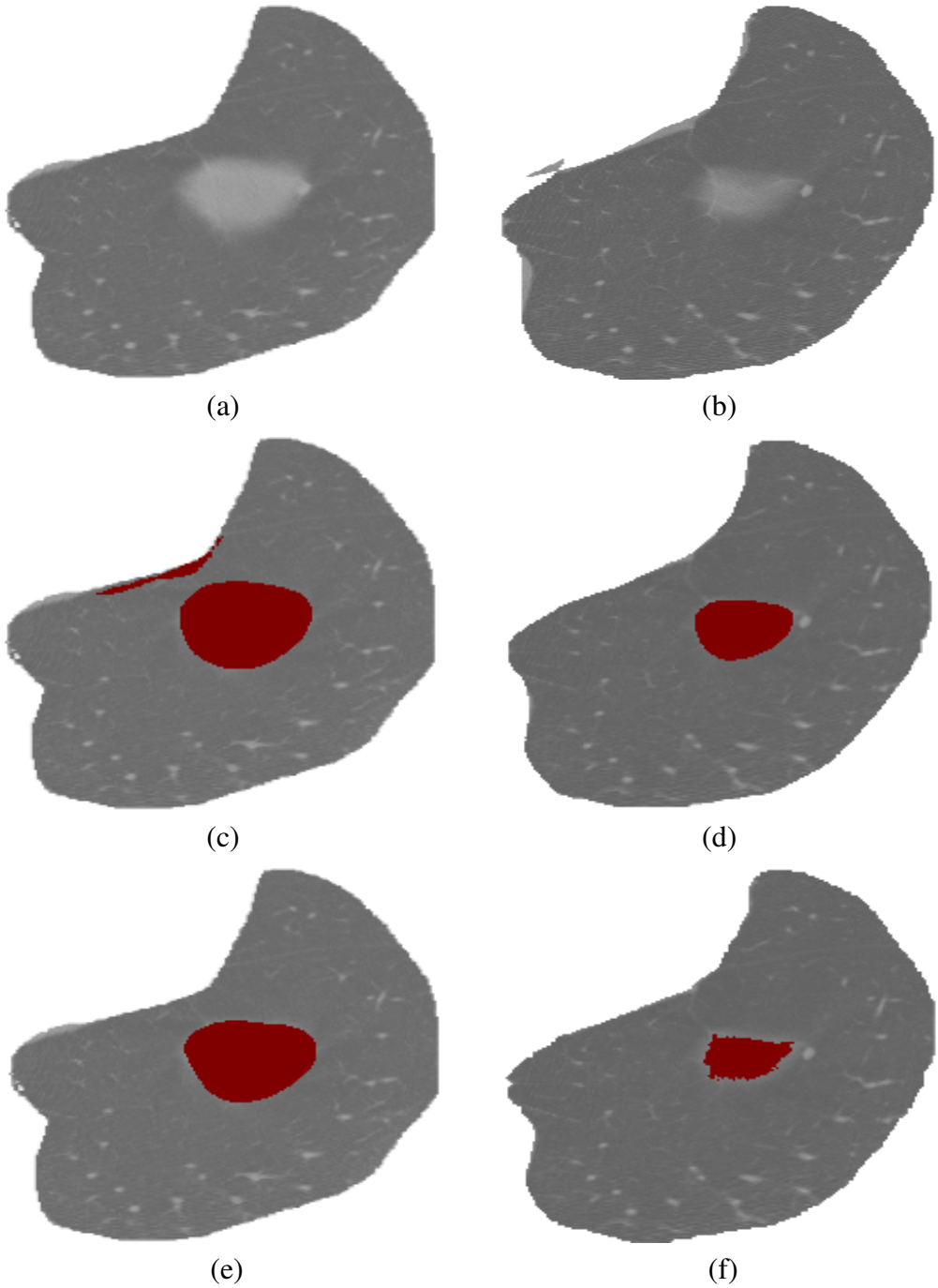


Figure 4.9 (a), (b) Lung Computed Tomography images showing part-solid nodules. (c), (d) Identified Part-solid nodule by the proposed deep learning approach. (e), (f) Results of CRF + DNN

$$MIoU = \frac{\sum_i x_{ii}}{C(\sum_i \sum_{j \neq i} x_{ij} + \sum_j x_{ji} - x_{ii})}, \tag{4.4.1}$$

where, C is the number of classes (2 in this study) and x_{ij} represents the pixels belonging to class i and is predicted as class j. Pixel accuracy is calculated as,

$$PA = \frac{\sum_i x_{ii}}{\sum_i \sum_j x_{ij}}. \tag{4.4.2}$$

Precision, recall and F1-score measures are calculated from confusion matrix as defined in Chapter 3, Table 3.7. The metrics of performance analysis such as accuracy, precision, recall, F1-score, MIoU, pixel accuracy value carried out for the developed model is shown in Table 4.2. It is seen that F1-score of the algorithm is 0.95 with MIoU of 0.911 indicating that the deep learning model when combined with the CRF framework is localizing and detecting the part-solid nodules more accurately (89%). The accuracy level of the algorithm has increased from 83% to 89.48% by applying the CRF algorithm. The Receiver Operating Characteristic curve plotted with false positive rate versus true positive rate for the developed system is given in Figure 4.10. In the ROC curve of the proposed model, the diagonal line represents the random guess and the curve that is nearing the value 1 in y-axis shows that the system is predicting the class labels rather than merely guessing them.

Table 4.2 Performance analysis of the proposed system

Image	MIoU	Pixel Accuracy	Precision	Recall	F1-score
DCNN	0.62	83.00%	0.89	0.78	0.88
DCNN + CRF	0.91	89.48%	0.95	0.95	0.95

Table 4.3 presents the performance of various popular existing methods for semantic segmentation. The authors in Kumar et al. (2015) and Kasinathan et al. (2019) have developed a system to classify the regions as nodule and non-nodule, whereas the proposed system identifies the part-solid nodules in lung CT images with a pixel accuracy of 89% and localization of part-solid nodules with MIoU of 0.911. Hoo-Chang and others Hoo-Chang et al. (2016) have developed two systems for identifying various lung diseases using the existing network architectures such as Googlenet and Alexnet,

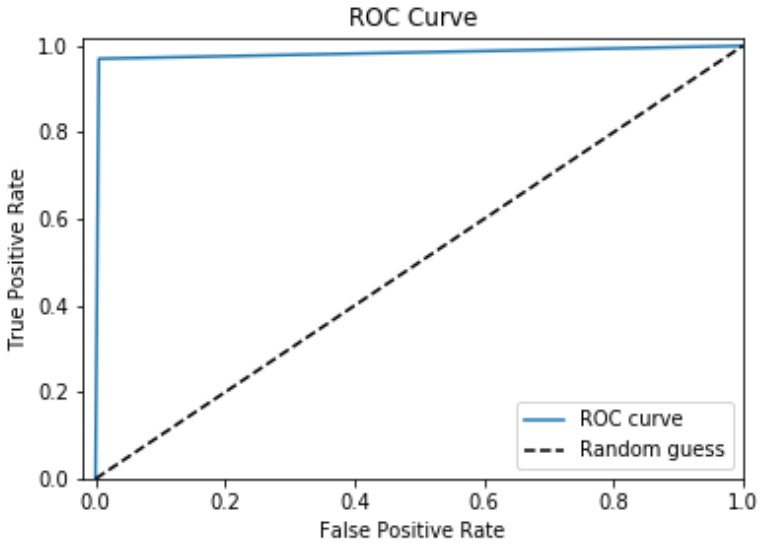


Figure 4.10 ROC curve

which are popularly used for classification, however, they fail to localize the nodules. The proposed model localizes the part-solid nodules which serves as a primary step for further analysis and facilitates the timely treatment of the patients. Decreasing the false positives, increases the precision, which is an inevitable requirement in medical image analysis. The authors of Song et al. (2017) and Kasinathan et al. (2019) have obtained an average of 6 false positives per scan whereas the proposed model uses CRF algorithm for semantic segmentation which reduces the false positives and subsequently increases the overall accuracy of the system under study. Malignancy prediction system using CNN model is developed by Ardila in 2019 Ardila et al. (2019). This system do not specifically localize the part-solid nodules unlike the proposed system.

4.5 Summary

Deep learning approach is adopted for identifying the part-solid nodules as they represent the early stage of lung cancer. A Deep Convolution Neural Network is incorporated within the Conditional Random Field framework to reduce the occurrence of false positives. This also has increased the accuracy of the system from 83% to 89.48%. As the process involved identifies the deep features by itself considering a large number convolution and deconvolution layers, training time required is 45 minutes with i5-7300-HQ

Table 4.3 Comparison of Proposed model with existing models

Method	**Dataset**	**Identification**	**Network architecture**	**Method adopted**	**Accuracy (percent)**
Kumar et al. (2015)	LIDC/IDRI	Lung Nodules	CNN and SVM classifier	Features calculated from CNN	75.01
Hoo-Chang et al. (2016)	LIDC & NELSON	Lung Interstitial Disease	GoogleNet	Segmentation and Classification	57.00
Hoo-Chang et al. (2016)	LIDC & NELSON	Lung Interstitial Disease	AlexNet	Classification	79.00
Song et al.(2017)	LUNA16	Benign and malignant nodules classification	Deep Learning approach	Deep Features calculated	65.00
Kasinathan et al. (2019)	LIDC	Nodule & Non-nodule	Active Contour & AlexNet	Segmentation & classification	97.00
Ardila et al.(2019)	LIDC	malignancy probability & localization	CNN model	Time series data. i.e. prior and current CT images	AUC is 94.40
Proposed Model (CRF+DCNN)	LIDC/IDRI	Part-solid Lung Nodules	A complete Deep Learning Approach	Segmentation and Classification	89.48

at 2.50 GHz processor, 16GB RAM and Nvidia GeForce GTX 1050 whereas average testing time required is 4 seconds per sample.

CHAPTER 5

Conclusion and Future Works

The present study successfully develops two automated CAD systems namely pipeline approach where a series of algorithms are adopted and a deep learning approach which has deep layers of convolution neural networks, for identifying the lung cancer in an early stage when it manifests itself as part solid / sub-solid nodules in lung CT images. The consistency of the two proposed automated CAD systems is experimentally demonstrated using two well known benchmark datasets namely I-ELCAP and LIDC / IDRI database.

The pipeline method adopts a series of algorithms for denoising, segmentation, feature extraction and selection, classification which finally yields results with the desired accuracy.

- The NLTV method adopted for denoising adopts to the noise distribution in the input by appropriately designing the model and subsidizes the noise substantially.

- Use of Chan-Vese model and morphological operations have successfully detected region of interest in the given lung CT images without human intervention. The segmentation method adopted here duly segments the regions even when the edges are not defined properly.

- The proposed system selects the most relevant statistical features based on the Eigen values thereby improving the accuracy of the classification and prediction models such as SVM, Fuzzy C-Means and Random Forest compared to other existing models.

- Use of Histogram of Gradients method considers every pixel and captures deep texture variations, both in magnitude and orientation along with statistical measures thus making identification of sub-solid nodules more accurate and reliable.

- Sensitivity of the CAD system in classifying the nodule and non nodule region in Phase I using supervised SVM classifier is in the order of 95 percent and that of unsupervised Fuzzy C-Means classifier is 94 percent. Random Forest Method classification is found to give consistent results with reduced out of bag error. The methods adopted for both supervised and unsupervised classification of nodules are found to give consistent results with error rate less than 6 percent.

- Sensitivity of CAD system in classifying the solid and sub-solid nodules in Phase II using the supervised SVM classifier is 95% and that of the unsupervised K-Means classifier is 91%. Again the results are consistent with error rate rate less than 6%.

- It can be concluded that the developed system is efficient in identifying lung cancer nodule (sub-solid/part-solid) in lung CT images in its early stages of occurance.

- On an average supervised classification and unsupervised classification models adopted for identifying the sub-solid nodules give accurate results in the range of 96 and 94 percent, respectively, thus establishing that prior knowledge of nodule is not essential.

- The developed CAD system works even if the test data contains less number of images.

- Finally, the developed CAD system efficiently identifies the sub-solid nodules in lung CT images without human intervention.

Deep learning approach adopted for identifying the sub-solid / part-solid nodules that represent the early stage of lung cancer is found to give good results.

- A Deep Convolution Neural Network (DCNN) incorporated within the Conditional Random Field (CRF) framework in order to reduce the occurrence of false positives has increased the accuracy of the system from 83% to 89.48%.

- The developed model eliminates a series of image processing methods for identifying the part-solid nodules in given lung Computed Tomography images which avoids the dependency on handcrafted features.

- The proposed automated system eliminates the need for human intervention and also it is found to give an accuracy of 89.48% with Mean IoU of 0.911 thus making the system robust.

- As the process involved identifies the deep features by itself considering a large number convolution and deconvolution layers, training time required is 45 minutes with i5-7300-HQ at 2.50 GHz processor, 16GB RAM and Nvidia GeForce GTX 1050. The average testing time required per sample is found to be 4 seconds.

- The CAD systems developed in the present study gives much higher accuracy, sensitivity and specificity values compared to similar models developed earlier. Hence the systems developed is more reliable than similar systems available.

Though the pipeline CAD system developed is accurate in identifying the sub-solid nodules in lung CT images, it involves series of algorithms to accomplish the task which makes the system computationally complex. Also, the algorithms incorporated work in a serial manner which is time consuming. Hence, architecture can be improved by parallelizing the algorithms that work on multi-core GPU. The second CAD system that adopts deep learning convolution neural network approach gives accuracy of 89% which may be increased by adopting graph-cut algorithms along with CRF framework. The DCNN architecture may be optimized by adopting other filters along with convolution filter and activation functions.

The developed CAD systems need to be tested by applying on the real-time datasets obtained from hospitals. An automated CAD system for identification of liquid nodules in lung CT images which represent the very initial stage of cancer may be developed since the patients chance of survival can be increased further.

BIBLIOGRAPHY

Agostinelli, F., Hoffman, M., Sadowski, P., and Baldi, P. (2014). Learning activation functions to improve deep neural networks. *arXiv preprint arXiv:1412.6830*.

Alam, F. I., Zhou, J., Liew, A. W.-C., Jia, X., Chanussot, J., and Gao, Y. (2019). Conditional random field and deep feature learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(3):1612–1628.

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature medicine*, 25(6):954.

Armato III, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., et al. (2011). The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931.

Aubert, G. and Aujol, J.-F. (2008). A variational approach to removing multiplicative noise. *SIAM journal on applied mathematics*, 68(4):925–946.

Boykov, Y. Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, volume 1, pages 105–112. IEEE.

Bracco, C., Regge, D., Stasi, M., Gabelloni, M., and Neri, E. (2019). Principles of ct and mr imaging. In *Nuclear Medicine Textbook*, pages 187–198. Springer.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2):89–97.

Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277.

Chen, L., Mao, X., Xue, Y., and Cheng, L. L. (2012). Speech emotion recognition: Features and classification models. *Digital signal processing*, 22(6):1154–1160.

Choi, W.-J. and Choi, T.-S. (2012). Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images. *Information Sciences*, 212:57–78.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE.

De Nunzio, G., Tommasi, E., Agrusti, A., Cataldo, R., De Mitri, I., Favetta, M., Maglio, S., Massafra, A., Quarta, M., Torsello, M., et al. (2011). Automatic lung segmentation in ct images with accurate handling of the hilar region. *Journal of digital imaging*, 24(1):11–27.

Erasmus, J. J., Connolly, J. E., McAdams, H. P., and Roggli, V. L. (2000). Solitary pulmonary nodules: Part i. morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics*, 20(1):43–58.

Gilboa, G. and Osher, S. (2008). Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028.

Gravel, P., Beaudoin, G., and De Guise, J. A. (2004). A method for modeling noise in medical images. *IEEE Transactions on medical imaging*, 23(10):1221–1232.

Guenther, N. and Schonlau, M. (2016). Support vector machines. *The Stata Journal*, 16(4):917–937.

Hansell, D. M., Bankier, A. A., MacMahon, H., McLoud, T. C., Muller, N. L., and Remy, J. (2008). Fleischner society: glossary of terms for thoracic imaging. *Radiology*, 246(3):697–722.

Henschke, C. I., Yankelevitz, D. F., Mirtcheva, R., McGuinness, G., McCauley, D., and Miettinen, O. S. (2002). Ct screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *American Journal of Roentgenology*, 178(5):1053–1057.

Heverhagen, J. T. (2016). Physics of computed tomography scanning. In *Handbook of Neuro-Oncology Neuroimaging*, pages 145–149. Elsevier.

Hoffman, J., Young, S., Noo, F., and McNitt-Gray, M. (2016). Freect_wfbp: A robust, efficient, open-source implementation of weighted filtered backprojection for helical, fan-beam ct. *Medical physics*, 43(3):1411–1420.

Hoo-Chang, S., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., and Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learningte. *IEEE transactions on medical imaging*, 35(5):1285.

Jacobs, C., Sánchez, C. I., Saur, S. C., Twellmann, T., de Jong, P. A., and van Ginneken, B. (2011). Computer-aided detection of ground glass nodules in thoracic ct images using shape, intensity and context features. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 207–214. Springer.

Jacobs, C., van Rikxoort, E. M., Twellmann, T., Scholten, E. T., de Jong, P. A., Kuhnigk, J.-M., Oudkerk, M., de Koning, H. J., Prokop, M., Schaefer-Prokop, C., et al. (2014). Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. *Medical image analysis*, 18(2):374–384.

Janocha, K. and Czarnecki, W. M. (2017). On loss functions for deep neural networks in classification. *arXiv preprint arXiv:1702.05659*.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260.

Kasinathan, G., Jayakumar, S., Gandomi, A. H., Ramachandran, M., Fong, S. J., and Patan, R. (2019). Automated 3-d lung tumor detection and classification by an active contour model and cnn classifier. *Expert Systems with Applications*, 134:112–119.

Kuhnigk, J.-M., Dicken, V., Bornemann, L., Bakai, A., Wormanns, D., Krass, S., and Peitgen, H.-O. (2006). Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic ct scans. *IEEE Transactions on Medical Imaging*, 25(4):417–434.

Kumar, D., Wong, A., and Clausi, D. A. (2015). Lung nodule classification using deep features in ct images. In *Computer and Robot Vision (CRV), 2015 12th Conference on*, pages 133–138. IEEE.

Larici, A. R., Farchione, A., Franchi, P., Ciliberto, M., Cicchetti, G., Calandriello, L., del Ciello, A., and Bonomo, L. (2017). Lung nodules: size still matters. *European Respiratory Review*, 26(146):170025.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., and Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234:11–26.

MacMahon, H., Austin, J. H., Gamsu, G., Herold, C. J., Jett, J. R., Naidich, D. P., Patz Jr, E. F., and Swensen, S. J. (2005). Guidelines for management of small pulmonary nodules detected on ct scans: a statement from the fleischner society. *Radiology*, 237(2):395–400.

Magdy, E., Zayed, N., and Fakhr, M. (2015). Automatic classification of normal and cancer lung ct images using multiscale am-fm features. *Journal of Biomedical Imaging*, 2015:11.

Messay, T., Hardie, R. C., and Rogers, S. K. (2010). A new computationally efficient cad system for pulmonary nodule detection in ct imagery. *Medical image analysis*, 14(3):390–406.

Mohammad, R. M., Muqbil, I., Lowe, L., Yedjou, C., Hsu, H.-Y., Lin, L.-T., Siegelin, M. D., Fimognari, C., Kumar, N. B., Dou, Q. P., et al. (2015). Broad targeting of

resistance to apoptosis in cancer. In *Seminars in cancer biology*, volume 35, pages S78–S103. Elsevier.

Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685.

Murphy, K., van Ginneken, B., Schilham, A. M., De Hoop, B., Gietema, H., and Prokop, M. (2009). A large-scale evaluation of automatic pulmonary nodule detection in chest ct using local image features and k-nearest-neighbour classification. *Medical image analysis*, 13(5):757–770.

Nithila, E. E. and Kumar, S. (2017). Automatic detection of solitary pulmonary nodules using swarm intelligence optimized neural networks on ct images. *Engineering science and technology, an international journal*, 20(3):1192–1202.

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

Schneider, F., Balles, L., and Hennig, P. (2019). Deepobs: A deep learning optimizer benchmark suite. *arXiv preprint arXiv:1903.05499*.

Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S. J., Wille, M. M. W., Naqibullah, M., Sánchez, C. I., and van Ginneken, B. (2016). Pulmonary nodule detection in ct images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*, 35(5):1160–1169.

Shen, W., Zhou, M., Yang, F., Yang, C., and Tian, J. (2015). Multi-scale convolutional neural networks for lung nodule classification. In *International Conference on Information Processing in Medical Imaging*, pages 588–599. Springer.

Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., Zang, Y., and Tian, J. (2017). Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition*, 61:663–673.

Society, A. C. (2016). American cancer society. *Cancer Facts and Figures*.

Soh, L.-K. and Tsatsoulis, C. (1999). Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on geoscience and remote sensing*, 37(2):780–795.

Song, Q., Zhao, L., Luo, X., and Dou, X. (2017). Using deep learning for classification of lung nodules on computed tomography images. *Journal of healthcare engineering*, 2017.

Su, Z., Yang, Z., Xu, Y., Chen, Y., and Yu, Q. (2015). Apoptosis, autophagy, necroptosis, and cancer metastasis. *Molecular cancer*, 14(1):48.

Suárez-Cuenca, J. J., Tahoces, P. G., Souto, M., Lado, M. J., Remy-Jardin, M., Remy, J., and Vidal, J. J. (2009). Application of the iris filter for automatic detection of pulmonary nodules on computed tomography images. *Computers in Biology and Medicine*, 39(10):921–933.

Tan, M., Deklerck, R., Jansen, B., Bister, M., and Cornelis, J. (2011). A novel computer-aided lung nodule detection system for ct images. *Medical physics*, 38(10):5630–5645.

Tao, Y., Lu, L., Dewan, M., Chen, A. Y., Corso, J., Xuan, J., Salganicoff, M., and Krishnan, A. (2009). Multi-level ground glass nodule detection and segmentation in ct lung images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 715–723. Springer.

Taşcı, E. and Uğur, A. (2015). Shape and texture based novel features for automated juxtapleural nodule detection in lung cts. *Journal of medical systems*, 39(5):46.

Teramoto, A., Fujita, H., Yamamuro, O., and Tamaki, T. (2016). Automated detection of pulmonary nodules in pet/ct images: Ensemble false-positive reduction using a convolutional neural network technique. *Medical physics*, 43(6Part1):2821–2827.

Van Ginneken, B., Setio, A. A., Jacobs, C., and Ciompi, F. (2015). Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pages 286–289. IEEE.

Van Rikxoort, E. M., de Hoop, B., Viergever, M. A., Prokop, M., and van Ginneken, B. (2009). Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical physics*, 36(7):2934–2947.

Vlahos, I., Stefanidis, K., Sheard, S., Nair, A., Sayer, C., and Moser, J. (2018). Lung cancer screening: nodule identification and characterization. *Translational Lung Cancer Research*, 7(3).

Willemink, M. J. and Noël, P. B. (2019). The evolution of image reconstruction for ct—from filtered back projection to artificial intelligence. *European radiology*, 29(5):2185–2195.

Zheng, Y., Jeon, B., Xu, D., Wu, Q., and Zhang, H. (2015). Image segmentation by generalized hierarchical fuzzy c-means algorithm. *Journal of Intelligent & Fuzzy Systems*, 28(2):961–973.

Zhou, J., Chang, S., Metaxas, D. N., Zhao, B., Ginsberg, M. S., and Schwartz, L. H. (2006). An automatic method for ground glass opacity nodule detection and segmentation from ct studies. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 3062–3065. IEEE.

PUBLICATIONS

Journal publication

1. Savitha, G., and P. Jidesh. A fully-automated system for identification and classification of subsolid nodules in lung computed tomographic scans. *Biomedical Signal Processing and Control, (Elsevier)* 53 (2019) 1–14 : 101586, https:// doi.org/ 10.1016 / j.bspc.2019.10158, 2019.

2. Savitha G., and P. Jidesh, A Holistic Deep Learning Approach for Identification and Classification of Sub-solid Lung Nodules in Computed Tomographic Scan, *Computers and Electrical Engineering (Elsevier)*, (Accepted for publication), 2020.

Conference Publication

1. G. Savitha and P. Jidesh, (2018). Lung Nodule Identification and Classification from Distorted CT Images for Diagnosis and Detection. *Machine Intelligence and Signal Analysis, AISC, Springer*, Volume 748, 11–23, 2019, https:// doi.org/ 10.1007/ 978-981-13-0923-6-2

BIODATA

Name : Savitha G.

Email : gsavitha24@gmail.com

Date of Birth : 24^{th} July 1989.

Permanent address : Savitha G.,
D/o Prof. S. G. Mayya,
13-4/5, Hamsa,
Sadashiva Temple Road,
Surathkal, Karnataka-575014.

Educational Qualifications :

Degree	**Year**	**Institution / University**
B.E.	2011	Canara Engineering College, Benjanapadavu.
Information Science & Engineering		Visveshwaraiah Technological University.
M.Tech.	2013	NMAMIT, Nitte.
Computer Science & Engineering		Visveshwaraiah Technological University.