

DEEP LEARNING-BASED DECISION SUPPORT SYSTEM FOR LUNG CANCER DETECTION

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

Dodia Shubham Jagdish



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE - 575 025

October, 2023

DEEP LEARNING-BASED DECISION SUPPORT SYSTEM FOR LUNG CANCER DETECTION

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

Dodia Shubham Jagdish

(187125CO006)

Under the guidance of

Prof. Annappa B.

Dept of CSE, NITK



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

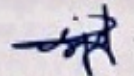
SURATHKAL, MANGALORE - 575 025

October, 2023

DECLARATION

by the Ph.D. Research Scholar

I hereby declare that the Research Thesis entitled **Deep Learning-based Decision Support System for Lung Cancer Detection** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy** in Department of Computer Science and Engineering 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.



Dodia Shubham Jagdish, 187125 187CO006
Department of Computer Science and Engineering

Place: NITK, Surathkal

Date: October 16, 2023

CERTIFICATE

ACKNOWLEDGEMENTS

This is to certify that the Research Thesis entitled **Deep Learning-based Decision Support System for Lung Cancer Detection** submitted by **Dodia Shubham Jagdish** (Register Number: 187125 187CO006) as the record of the research work carried out by him, is accepted as the Research Thesis submission in partial fulfilment of the requirements for the award of degree of **Doctor of Philosophy**.

Dr. Annappa
Professor

Dept. of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal, Post Srinivasnagar
MANGALORE, INDIA - 575 025
annappa@nitk.edu.in

Annappa 18/10/23

Prof. Annappa B.
Research Supervisor

(Signature with Date and Seal)

Chairman
DUGC / DPGC / DRPC
Dept. of Computer Engg.
NITK - Surathkal
Srinivasnagar - 575 025

Baran 18/10/23
Chairman - DRPC

(Signature with Date and Seal)

ACKNOWLEDGEMENTS

Expressing my gratitude is the finishing touch on my dissertation. Research transformed me as a better individual, not only in the technical domain but also on a personal level. Such improvements would not have been possible without the support of the people who have helped me.

First, I would like to express my sincere gratitude to my research supervisor Prof. Annappa B. for the continuous support, guidance and motivation throughout my Ph.D. study and related research. His vast knowledge, encouragement and patience definitely helped me to reach to this height. I thank him for identifying my capabilities and giving this opportunity to pursue research under him. The amount of consideration that he has given during the stages of research paper and thesis writing is impeccable. In fact, I could not have imagined having a better advisor and mentor for my Ph.D study.

Now, I would like to thank my Research Progress Committee (RPAC) members Dr. Biswajit R. Bhowmik and Dr. Anand Kumar M. for their insightful assessment and suggestions to further improve my work. The points raised by them were extremely valuable and helped me to look at my work in a different angle. I wish to show my appreciation to the Head of the Department Dr. Manu Basavaraju, former Head of the Department Dr. Shashidhar G Koolagudi and other faculty members for their continuous support. I thank all the teaching and non-teaching staff of Computer Science Department, NITK for their help during my research period.

I sincerely appreciate Dr. Mahesh A. Padukudru's knowledge of the subject of my research. It has aided me in developing an understanding of realistic medical scenario. The validation he provided for the work has improved my research exposure.

I will always cherish the kind of affection, care and support received from my best friends Spoorthy, Rashmi, and Garima. It would not be possible to even dream about

achieving this goal without their support. Thank you for all the emotional and technical support. Now, I thank my fellow lab mates Vishal, Pravin, Alkha, Khyamling, Nikhil, Pramod, Pradeep, Kallinath, Sneha, Swathi, my room-mate Natarajan and group members Manjunath, Saraswati, Aarabhi, Sachin, Shankar, Naveen for the stimulating discussions, knowledge transfer and team work. I wish to extend my special thanks to all the teachers who taught me in schools and colleges. I would not be in this position to write this without them. In particular, I am grateful to Dr. Damodar Reddy from National Institute of Technology Goa for enlightening me the first glance of research.

Last but not the least, I would like to thank my father Mr. Jagdish A. Dodia, mother Mrs. Sushma J. Dodia, brother Mr. Tushar J. Dodia, and sister-in-law Mrs. Rutuja T. Dodia for their love and guidance throughout my life. My family had to go through a lot of sacrifices for my studies and achievements and words cannot express how hard those were and how much I am grateful for those sacrifices. I would like to express my gratitude towards my in-laws too for their support and motivation during my studies and in my life. Most importantly, I wish to thank my loving and supportive wife Mrs. Trupti S. Dodia who provided unending support and inspiration. I also want to thank my daughter Shanaya, nephew Avneesh, and niece Ishani for helping me laugh and love more despite the stress of research.

Finally, I praise the almighty for giving me the vision and strength to continue with confidence. Thank you God.

Dodia Shubham Jagdish

Place: NITK, Surathkal

Date: October 16, 2023

ABSTRACT

Cancer is a major cause of significant fatal rates and morbidity worldwide. According to the latest World Health Organization (WHO) estimates issued in 2020, cancer disease has the greatest mortality rate, accounting for around 10 million deaths. Over a lifetime, 1 in 18 men and 1 in 46 women are known to develop lung cancer. Accurate identification of lung cancer has been a challenging task for decades. Even though there are techniques to identify lung cancer nodules, it takes enormous efforts from expert radiologists. Therefore, it is very crucial to automate the process of identifying nodules from Computed Tomography (CT) scans. This thesis discusses the methods proposed to perform lung cancer detection, segmentation, and classification using novel deep learning algorithms.

First, the task of detecting lung nodules from CT scan images is performed. The lung nodules are irregular tissue formations that can be as small as 3 mm in diameter. The detection of these lung nodules is a tedious and time-consuming task, as careful examination needs to be carried out by radiologists. The annotations that the radiologists provide must be precise and accurate as well. This can lead to human error. Combining this with computer-aided algorithmic solutions may resolve this issue. However, deploying this real-time environment is another challenge as it needs to be interfaced with these solutions as per doctor's requirements. In this thesis, different deep-learning solutions are used to develop lung cancer nodule detection from CT scan images. The potential nodule candidates are identified by the proposed detection methods.

Second, the task of segmenting the nodule regions from the detected lung nodules is performed. Once the potential nodule candidates are detected, the accurate nodules are to be segmented. One of the main challenges that occur in segmentation of lung nodules is that non-nodules that appear like nodules can be segmented. Therefore, segmentation of lung nodules is essential to avoid misdiagnosis. In this thesis, Artificial Intelligence

(AI)-based methods are proposed to perform accurate lung nodule segmentation tasks from the input CT scans.

Third, the task of classifying a segmented nodule as cancerous or non-cancerous is performed. The tumor/nodule found in the lung/thoracic region can be malignant or benign. The spread rate and re-occurrence of a malignant nodule in the human body are very rapid. Therefore, it is crucial to identify the type of nodule at the earliest. In this thesis, various deep-learning solutions are designed and developed to perform this task.

All the proposed methods in this thesis are evaluated on the publicly available benchmark LUNA16 dataset; their respective results are presented in subsequent chapters and verified by an expert pulmonologist. The proposed models resulted in superior performance in comparison with state-of-the-art techniques. When compared, state-of-the-art techniques had accuracies of 96.9%, 94.97%, and 96.9% for the detection, segmentation, and classification task, respectively. However, the proposed models yielded an accuracy of 98.21% for the detection task, a dice similarity coefficient of 98.0% for the segmentation task, and an accuracy of 98.7% for the classification task. This clearly shows an improvement of 1.31%, 3.03%, and 1.8% for the detection, segmentation, and classification tasks respectively.

Keywords: Lung cancer, Medical imaging, Nodule segmentation, Nodule detection, Nodule classification, Convolutional Neural Network.

CONTENTS

List of Figures	viii
List of Tables	x
List of Abbreviations	xi
1 Introduction	1
1.1 Background	1
1.1.1 Lung Cancer	3
1.1.2 Types of Nodules	3
1.1.3 Stages of Lung Cancer	5
1.1.4 Imaging Modalities	5
1.1.5 Computed Tomography	6
1.1.6 Conventional and Deep Learning based CAD Systems for Lung Cancer	6
1.2 Motivation	7
1.3 Applications	8
1.4 Challenges	8
1.5 Brief Overview of Thesis Contributions	9
1.6 Organization of the Thesis	10
2 Literature Review	11
2.1 Conventional/ Traditional Methods	12
2.2 Deep learning Methods	12
2.3 Overall Findings for Conventional CAD Systems	18
2.3.1 Pre-processing	18
2.3.2 Segmentation	18
2.3.3 Feature Extraction	18

2.3.4	Classification	19
2.4	Overall Findings for Deep Learning-based CAD Systems	26
2.5	Research Gaps	26
2.6	Problem Statement	27
2.7	Dataset used for this Thesis	29
2.8	Summary	30
3	Lung Cancer Detection	31
3.1	Introduction	31
3.2	Proposed Methodology	32
3.2.1	Stage 1: Lung Nodule Detection	33
3.2.1.1	Image Pre-processing	34
3.2.1.2	Nodule Detection: RFR V-Net	34
3.2.2	Stage 2: Data preparation for Nodule Classification	37
3.2.2.1	Image Augmentation	38
3.2.3	Stage 3: Post-processing	38
3.2.4	Stage 4: Lung Nodule Classification	39
3.2.4.1	Nodule Classification: NCNet	40
3.3	Results and Discussions	43
3.4	Summary	51
4	Lung Cancer Nodule Segmentation	53
4.1	Introduction	53
4.2	Proposed Lung Cancer Nodule Segmentation System-I	54
4.2.1	Pre-processing	55
4.2.2	Augmentation	56
4.2.3	Light-weight RefineNet for Nodule Segmentation	58
4.2.4	PSO-based CondenseNet	60
4.2.5	Materials and Methods: Neural Network Configurations	62
4.2.6	Evaluation Metrics	64
4.2.7	Results and Discussion	65
4.2.7.1	Segmentation of lung nodules in the CT scan	65

4.2.7.2	Free-response Receiver Operating Characteristic (FROC) analysis of Lung Nodule Segmentation system	65
4.2.7.3	Performance of Lung Nodule Segmentation system	67
4.2.7.4	Comparison of Lightweight RefineNet with SOTA Nodule Segmentation CAD systems	68
4.2.7.5	Performance evaluation for proposed Lung Nodule Segmentation system on multiple datasets	69
4.2.7.6	Performance comparison of various Optimization Algorithms along with parameters used	69
4.2.7.7	AUC-ROC graph representing various Optimization Algorithms combined with CondenseNet architecture	71
4.2.7.8	Analysis of accuracy and loss for the proposed Classification method	71
4.2.7.9	Performance of Lung Nodule Classification system with existing methods	72
4.2.7.10	Comparison of PSO-based CondenseNet with SOTA Nodule Classification CAD systems	73
4.2.7.11	Comparison of Time and Space Complexity for proposed models	73
4.3	Proposed Lung Cancer Nodule Segmentation System-II	74
4.3.1	Image Enhancement	76
4.3.2	Image Segmentation	77
4.3.3	Bag-of-Visual Words	78
4.3.4	Deep Features	79
4.3.5	Nodule Classification	79
4.3.6	Results and Discussion	79
4.3.6.1	Nodule Segmentation	79
4.3.6.2	Nodule Classification	81
4.4	Summary	84

5 Lung Cancer Nodule Classification	87
5.1 Introduction	87
5.2 Proposed Methodology	88
5.2.1 Image Preprocessing	89
5.2.2 Data Augmentation	91
5.2.3 Level-1 Classification: CBWO-CNN	93
5.2.4 Level-2 Classification: SE-Xception	96
5.3 Results and Discussion	99
5.3.1 Level-1 Classification	99
5.3.2 Level-2 Classification	101
5.4 Summary	104
6 Conclusions and Future Scope	105
6.1 Future Scope	106
References	110
Publications	126

LIST OF FIGURES

1.1 Different types of nodules. Column-wise: 1. Juxta; 2. Well-Circumscribed; 3. Vascularized and 4. Pleural-tail	4
1.2 Conventional machine learning and deep learning based CAD systems for Lung Cancer	7
3.1 Schematic overview of the proposed model for Lung Nodule Detection .	33
3.2 A CT Scan illustrating that pulmonary nodules occupy less than 5% of CT scan	34
3.3 Architecture diagram of proposed RFR V-Net	36
3.4 These are multiple architectures used in this work, where, (a) shows Conventional ResNet, (b) shows working of Squeeze and Expand block in SqueezeNet, and (c) shows proposed NCNet architecture which is a new combination of (a) and (b)	42
3.5 Results achieved for the proposed RFR V-Net.	44
3.6 Nodule and non-nodule images obtained before post-processing ((a) and (b)) and after post-processing ((c) and (d))	45
3.7 Graph representing the accuracy and loss achieved for the proposed NCNet3D model	46
3.8 Confusion matrices obtained for best recognition accuracies	47
3.9 FROC curve comparing the performance of (a) NCNet-Pseudo-3D-56 at 2.4 FPs/scan, and (b) NCNet-3D-56 at 2.3 FPs/scan	48
4.1 Block diagram of the proposed Lightweight CAD system	55
4.2 Schematic architecture of Generative Adversial Network	57

4.3	Architecture of proposed Lightweight RefineNet for Lung Nodule	
	Segmentation	59
4.4	Architecture of proposed PSO-based CondenseNet for Lung Nodule	
	Classification	62
4.5	Results achieved by the proposed Lightweight RefineNet model	66
4.6	FROC curve illustrating the performance of the Lightweight RefineNet model at different FPs/scan	66
4.7	AUC-ROC graph representing various Optimization Algorithms combined with CondenseNet architecture	71
4.8	Graph representing the accuracy and loss achieved	72
4.9	Schematic architecture of the proposed method	75
4.10	Comparison of proposed EFCM model with existing techniques	80
4.11	FROC curve for the performance of proposed EFCM at 2.7 FPs/scan	81
4.12	ROC curves for (a) BoVW, (b) Deep features, (c) BoVW + Deep features	82
4.13	Layer-wise accuracy values for (a) Deep features, (b) BoVW + Deep features	83
5.1	Block diagram of the proposed approach	89
5.2	Nodule, non-nodule, benign, and malignant images after preprocessing of CT scans	92
5.3	Augmented images of a nodule CT scan	92
5.4	Diagrammatic representation of (a) Xception network, (b) Squeeze-and-Excitation block, and (c) proposed SE-Xception model	96
5.5	CBWO-CNN model evaluated on various activation functions on testing data	100
5.6	(a) SE-Net, (b) Xception, and (c) proposed SE-Xception evaluated on various activation functions on testing data for Level-2 Classification	102
5.7	Visual depiction of correctly and mis-classified nodules for Level-2 Classification	103

LIST OF TABLES

1.1 Stages of Lung Cancer	5
2.1 Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved	13
2.1 Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	14
2.1 Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	15
2.1 Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	16
2.1 Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	17
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved	20
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	21
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	22
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	23
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	24
2.2 Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)	25

2.3	Details of LUNA16 dataset	30
3.1	Comparison of the proposed work with other Nodule Detection CAD systems	43
3.2	Comparison of various models on multiple layer configurations with proposed work for 2D, pseudo-3D, and 3D data	49
3.3	Comparison of previous works with the proposed NCNet	50
4.1	Performance of the proposed Light-weight along with existing segmentation methods for different FPs/scan values	67
4.2	Comparison of the proposed Light-weight RefineNet with SOTA Lung Nodule Segmentation CAD systems	68
4.3	Performance evaluation for proposed Lung Nodule Segmentation system	69
4.4	Performance comparison of various Optimization Algorithms used	70
4.5	Performance of the proposed Lung Nodule Classification method	72
4.6	Comparison of the proposed PSO-based CondenseNet with SOTA systems	73
4.7	Time and Space Complexity comparison for proposed models	74
4.8	Comparison of the proposed system with the SOTA systems	81
4.9	Performance of the Lung Nodule Classification for the proposed system	82
4.10	Comparison of the proposed system with the SOTA systems	84
5.1	Performance of CBWO-CNN model	100
5.2	Evaluation of CBWO-CNN	101
5.3	Performance assessed without and with shared network parameters	102
5.4	Comparison of previous works with proposed model	103

LIST OF ABBREVIATIONS

<u>Abbreviations</u>	<u>Expansion</u>
ABC	Artificial Bee Colony
ACC	Accuracy
ACO	Ant-Colony Optimization
AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under Curve
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
BBHE	Boosted Bilateral Histogram Equalization
BoVW	Bag of Visual Words
CAD	Computer-Aided Detection
CBWO-CNN	Cauchy Black Widow Optimization-based Convolutional Neural Network
CDF	Cumulative Distribution Function
CDNN	Convolutional Deep Neural Network
CNN	Convolution Neural Network
CSV	Comma Separated Value
CT	Computed Tomography
DALY	Disability-adjusted life years
DBResNet	Dual branch ResNet
DCGAN	Deep Convolutional Generative Adversarial Network
DFCNet	Deep Fully Connected Network
DICOM	Digital Imaging and Communications in Medicine
DNN	Deep Neural Network
DSC	Dice Similarity Coefficient
EFCM	Elagha initialization-based Fuzzy C-Means clustering
ELU	Exponential Linear Unit
FDG	FluoroDeoxyGlucose
FIS	Fuzzy Inference System
FNR	False Negative Rate
FP	False Positives
FPR	False Positive Rate
FPS/scan	False Positives per scan

Abbreviations	Expansion
FROC	Free-response Receiver Operating Characteristic
GA	Genetic Algorithm
GLCM	Gray Level Co-occurrence Matrix
HE	Histogram Equalization
HU	Hounsfield Unit
IoU	Intersection over Union
I-ELCAP	International Early Lung Cancer Action Program
KNN	K-Nearest Neighbor
LIDC	Lung Image Database Consortium
LNDb	Lung Nodule Database
LUNA	LUNG Nodule Analysis
Mask RCNN	Mask Region-based Convolution Neural Network
MLP	Multi Layer Perceptron
MRI	Magnetic Resonance Imaging
NCNet	Nodule Classification Network
NSCLC	Non-small cell lung cancer
PET	Positron Emission Tomography
PNN	Probabilistic Neural Network
PPV	Positive Prediction Value
PR	Precision
PSO	Particle Swarm Optimization
RCU	Residual Convolution Unit
REC	Recall
ReLU	Rectified Linear Unit
ResNet	Residual Network
RF	Random Forest
RF	Receptive Field
RFR V-Net	Receptive Field Regularized V-Net
ROC	Receiver-Operating Characteristics
ROI	Region of Interest
SDAE	Stacked Denoising Autoencoder
SEN	Sensitivity
SE-Net	Squeeze-and-Excitation Network
SE-Xception	SE-Net and Xception
SGD	Stochastic Gradient Descent
SIFT	Scale-Invariant Feature Transform
SOTA	State-of-the-art
SPE	Specificity
SR	Squeeze Ratio
SVM	Support Vector Machine
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Cancer is considered to be the world's most life-threatening illness. Many types of cancer affect men and women alike. The medical term referring to uncontrolled and irregular growth of cells in any tissue causing lumps, nodules, or masses to develop is known as cancer (Rajan et al. 2019). There are over 100 cancer types that can occur in various tissues such as the breast, skin, lung, colon, prostate, and blood, among others (Tran et al. 2019). In many situations, there are no unique symptoms suggestive of cancer; thus, the diagnosis may be delayed without a high degree of suspicion. The symptoms that can be often seen in cancer patients are non-specific, such as fatigue, weight loss, pain, pattern changes on the skin, bowel or bladder function changes, abnormal bleeding, constant cough or speech changes, fever, lumps, or masses.

Cancer staging is related to the magnitude of cancer spread. It is often determined by a combination of various imaging techniques and relevant tissue biopsy, which is helpful in identifying the type of cancer. Staging is relevant for the prognosis, where it helps caregivers to determine therapy protocols, which may include chemotherapy, immunotherapy, radiation, and surgery either alone or in various combinations. In particular, cancer in the body is more aggressive in later staging, i.e., the higher the stage (generally between 0 and 4), the more critical the patient's condition. These treatments vary according to the stage and type of cancer.

The human body retains cell growth and regenerates whenever necessary.

However, in unusual conditions, cell growth is uncontrolled and creates cell clusters called tumors. Tumors can be categorized into two types, non-cancerous and cancerous. Non-cancerous tumors are benign tumors, and cancerous tumors are malignant tumors. It is possible to treat and remove the benign tumor as the tumor does not invade other tissues or organs. Benign tumors can be treated entirely by surgery. There is a maximum possibility that it will not reoccur; if it does, it will occur in the same place. A malignant tumor can be treated with chemotherapy, radiation therapy, etc. However, these tumors may reoccur and have a high chance of invading other tissues and organs (Sinha 2018).

Lung cancer is the most common type among males and the third most common type among females, as stated by Potghan et al. (2018). Notably, there is a decrease in mortality rate in developed countries as there is an improvement in the detection of the disease in its early stages, diagnosis, and treatment. However, the majority of current research discards micro nodules when training models, which in turn fails to detect early stage cancer. Deploying a Computer-Aided Detection (CAD) system is also a very costly process in developing and underdeveloped nations. Therefore, this work aims in developing a low complex computer-aided lung cancer detection system considering micro-nodules which aids in early diagnosis. Tissue biopsy is an image-guided procedure to identify whether a nodule detected is malignant or benign. Studies have shown that the mortality rate is high for many patients diagnosed in the later stages. For the last two decades, there has been a sustained interest in detecting lung cancer at initial stages. Initially, chest X-ray screening was studied, but it was not successful in early detection as many subjects had interval lung cancer. It also has failed to demonstrate the reduction in the mortality rate (Kvale et al. 2014). Low-dose CT scans have been recently tested to detect lung cancer (Choi et al. 2018; Robles and Harris 2017; Sverzellati et al. 2016). The key questions to be addressed are, how effective screening is for the large-scale at-risk population for early diagnosis of lung cancer and how sensitive and accurate are the image findings for lung cancer diagnostics. Deep learning methods in Artificial intelligence (AI) has captured a high interest in medical image analysis for developing computer-aided detection (CAD)

systems for early diagnosis.

1.1.1 Lung Cancer

The tumor originating in the lung region is called lung cancer (Potghan et al. 2018). Many cancer deaths stem from lung cancer, which is also counted among the life-threatening diseases in the world. It is an increasing cause of both morbidity and mortality. Lung cancer today accounts for one in four cancer deaths (Liu et al. 2018a). It can be characterized as abnormal cell growth in one or both lungs (Aggarwal et al. 2015). Early signs and symptoms in patients make it easy to detect lung cancer. While, for most patients, signs are observed only after it has reached a higher stage making it difficult to detect at the initial stages of lung cancer (Chiang A 2019).

Image processing is a technique adapted to extract vital information from an image, where firstly, the image will be converted into a digital form. Then various operations are performed on that image to get an enhanced image (Bhalerao et al. 2019). Improved image processing methods, including AI algorithms, are the current cutting-edge tools for enhancing prediction for prognosis and treatment, particularly in cancer tumors in various medical applications. Liu et al. (2019) suggests that lung cancer can be detected using image recognition methods, as there are apparent differences between balanced and unhealthy pulmonary images.

1.1.2 Types of Nodules

The development of an irregular circular/oval growth in the lungs called lung nodules is regarded as an early manifestation of lung cancer. The complex lung structure makes diagnosing lung cancer a tedious task (Lavanya and Kannan 2018). Classification of nodules can be made into two types: the cancerous nodules, i.e., the affected nodules, which can also be called malignant nodules, and the non-cancerous nodules, i.e., the type of nodules, also called benign nodules. As explained by Kuruvilla and Gunavathi (2014), radiologists categorize nodules into four different types: Juxtapleural, Well-circumsized, Vascularized, and Pleural-tail. This can be illustrated in Figure 1.1 (Farag et al. 2010).

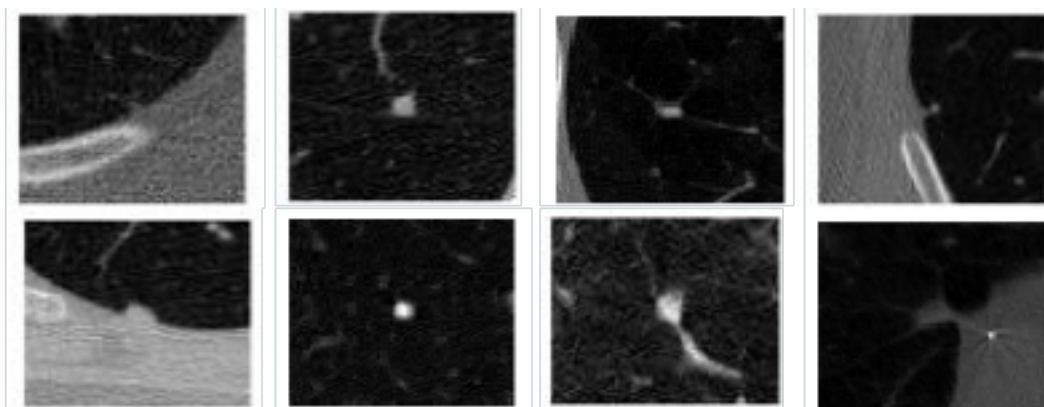


Figure 1.1: Different types of nodules. Column-wise: 1. Juxta; 2. Well-Circumscribed; 3. Vascularized and 4. Pleural-tail

Size is a primary determinant of a nodule, whether benign or malignant. A lung nodule with a size greater than 3 cm has a higher chance of being malignant (Choromańska and Macura 2012). The “doubling time” is even more significant: the time taken for the nodule volume to double. This is a more sensitive and specific feature of malignant lesions than just the size, but it needs serial images ¹. If the doubling time is faster, i.e., less than 30 days, it is likely due to a benign cause such as an infection. A lesion that stays static for two years is considered benign.

A common and effective approach for detecting lung nodules is Computed Tomography (CT). CT has a much higher resolution than a conventional chest X-ray in which nodules usually need 1 cm or larger to be detected in a chest X-ray (Shaziya et al. 2018). The confirmation of the malignancy requires tissue biopsy. Positron Emission Tomography (PET) scan can also help determine if the nodule in the CT image is malignant (Rahman et al. 2019). It should be remembered that benign lesions such as an infection also lead to positive PET scans due to increased uptake of FluoroDeoxyGlucose (FDG) when there is a high metabolism. The confirmation in such situations demands a biopsy from the affected nodules or lymph nodes. According to Liu et al. (2018a), due to the poor quality of CT scans, the complex shapes of lung nodules and the unavailability of annotated information leave no option for current techniques to only measure the nodule center and an important diagnostic

¹To track changes in a diseased area and the results of therapies and treatments, serial images are made up of the sequential acquisitions throughout time of images of the same patient.

criterion, nodule width is often overlooked.

1.1.3 Stages of Lung Cancer

The staging of lung cancer can be done in four levels. These levels are illustrated in Table 1.1. Based on the order of grade, the stages from I to IV are given by [Kulkarni and Panditrao \(2014\)](#).

Table 1.1: Stages of Lung Cancer

Lung Cancer Stages	Parts of the body affected
Stage I	Restricted to the lung
Stage II and III	Restricted to the chest
Stage IV	Extensive metastases outside the chest/thorax or pleural involvement

1.1.4 Imaging Modalities

Various imaging techniques are used to diagnose, classify, and recognize the stages of lung cancer. Some popular techniques are Magnetic Resonance Imaging (MRI), CT, PET, chest radiograph (known as X-ray), etc. These imaging techniques provide local information on the presence of a tumor, metastate of the presence of the disease, and even the lymph nodes in the human body. A combination of imaging techniques, such as PET and CT, is also made to get deeper anatomical information about the region.

CAD devices are used to identify lung nodules automatically, which allows radiologists to detect disease. Lung cancer can be diagnosed in the early stages using CT images. Many CAD systems use CT scanning because the image capture process is quick, does not damage the patient's bones, and is less noisy ([Potghan et al. 2018](#)). Imaging tests like CT scans can recognize lung cancer as it provides a more accurate picture, suggests [Kulkarni and Panditrao \(2014\)](#). CT scans are more powerful in identifying and diagnosing lung cancer than regular chest X-rays. A 3-dimensional image can be reconstructed from CT scans as the acquisition of the images is performed in a continuous manner ([Hollings and Shaw 2002](#)).

1.1.5 Computed Tomography

Scanning CT is the most suitable method to visualize lung cancer (Niranjana and Ponnaivaikko 2017). As per Lakshmanaprabu et al. (2019), CT is best suited for tracking the location of tumors and determining the cancer stage in the body. CT scanning is used to identify lung mass tissue or nodules because it can detect slight irregularities suggesting lung cancer (Aggarwal et al. 2015).

In medical practice, the most powerful and common imaging technique is known to be CT for detecting nodules and for diagnosing with benefits such as cost-effectiveness, high spatial resolution, availability, and non-invasiveness (Gibaldi et al. 2015; Ng and Goh 2010). Compared to PET and MRI imaging techniques, CT is much easier and less costly and can achieve better sensitivity in lung nodule detection than X-ray imaging suggests Bhavanishankar and Sudhamani (2015), and Cieszanowski et al. (2016). The diagnosis of lung cancer at its initial stages has shown significant and encouraging results in CT screening in the form of pulmonary nodules, thereby mitigating mortality rate (Monkam et al. 2019). Therefore, most work in this field has been dedicated to lung nodule identification through CT scans of the thoracic region.

1.1.6 Conventional and Deep Learning based CAD Systems for Lung Cancer

The existing lung cancer CAD systems can be categorized into two types, traditional and deep-learning. Figure 1.2 depicts the block diagrams of the two kinds. The traditional CAD systems, as shown in Figure 1.2(a), consisting of steps such as pre-processing of input CT scans, segmentation of the region of interest (ROI), nodule candidate generation, feature extraction, selection of optimal features, classification of nodules and non-nodules. Traditional CAD systems consist of more steps and are time-intensive. Deep learning-based CAD systems are considerably faster and better performing than traditional CAD systems. The deep learning-based system, as shown in Figure 1.2(b), consists of steps such as nodule candidate generation from CT scan input and deep learning model to classify the nodule and non-nodule classes. This thesis's primary study area is designing and developing deep learning-based CAD

systems for lung cancer.

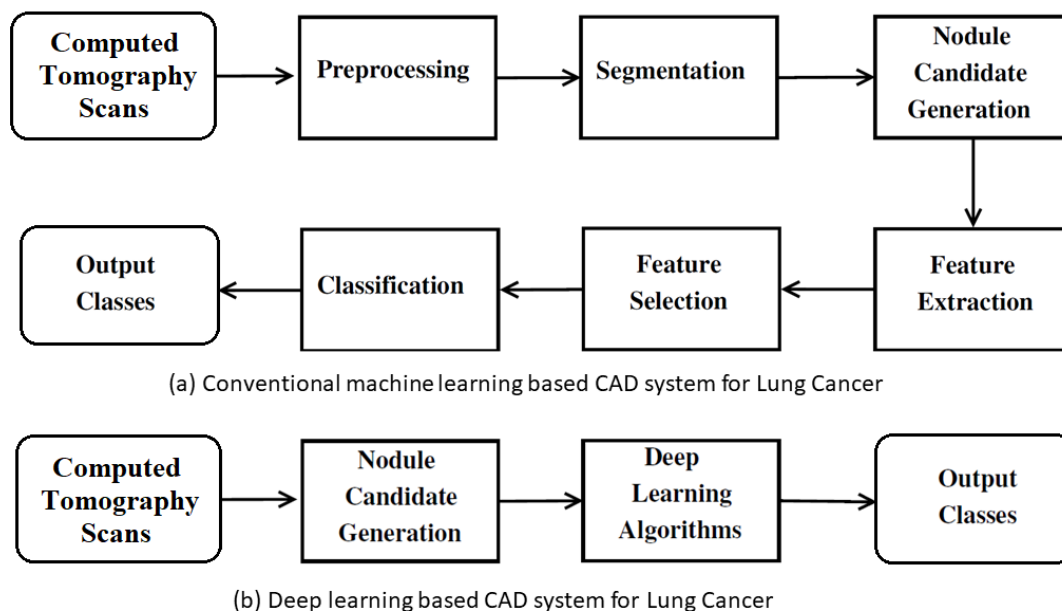


Figure 1.2: Conventional machine learning and deep learning based CAD systems for Lung Cancer

1.2 MOTIVATION

Lung cancer is one of the diseases with a very high mortality rate and is still a major concern. Below are some of the motivations that encouraged to choose this as a research area of this thesis.

1. According to the WHO, cancer accounted for 10 million deaths in 2020, and globally, cancer causes about 1 in 6 deaths. In 2020, incidence rate of lung cancer was 2.21 million and the mortality rate was 1.80 million. Based on incidence and mortality, lung cancer ranked the highest among other cancers worldwide in 2018. These statistics motivated authors to develop a CAD system for lung cancer detection.
2. A reliable CAD system with good performance can help provide radiologists with a second opinion. Also, it can reduce radiologists' interpretation time as numerous CT scans have to be examined by a radiologist on a day-to-day basis. This in turn helps the radiologist to avoid misdiagnosis.

3. About 70% of the patients are detected with lung cancer at their final stages, and only 16% have a chance of surviving in the next five years (Kulkarni and Panditrao 2014). There are greater chances of improving the survival rate of patients of 5 years to 70% if the patients are diagnosed earlier (Baldwin 2015).

1.3 APPLICATIONS

This section discusses some of the important applications for lung cancer-based CAD systems.

1. The main benefit of lung nodule CAD systems is that they give doctors a second opinion while treating patients, enabling them to treat patients more quickly.
2. Once the CAD system for lung cancer detection is successfully developed, the tool can be implemented for other types of cancers as well.

1.4 CHALLENGES

Even though a lot of work is done in detecting lung cancer nodules, there are some challenges to building real-time CAD systems. Some of the challenges are:

1. Acquired data is of poor quality because it contains noise for a range of reasons. Pre-processing is a crucial and difficult step in order to enhance the quality of image.
2. It is difficult to visually identify the difference between a malignant nodule and a normal pulmonary structure. This makes it essential to develop a CAD system.
3. It is a challenging task to increase patients' chances of survival as it is difficult to identify malignant nodules at the initial stage of development.
4. There is a lot of variability in the available datasets as they are recorded by multiple scanning devices, stored in different storage formats, or the modality of the image may differ. It is challenging to build a generalized model.
5. Finding reliable features and integrating them into CAD systems for performing benign and malignant nodules classification is tedious.

6. To reduce the number of false positives/scan (FP/Scan), computation time and improvising the model's sensitivity are also challenging. There is no consistent use of performance metrics to evaluate the system.
7. The radiologists may not give accurate annotations for a huge amount of data as the images are multi-dimensional.
8. Due to the variation in nodule type and size, it is not easy to differentiate between cancerous and non-cancerous nodules.
9. Selection of the optimal set of features for nodule detection and nodule classification is a challenging task and requires much domain knowledge.
10. It is extremely difficult to identify a wider variety of lung cancer nodules with improved sensitivity and less FP/scan. Therefore, it is essential to develop a lung cancer CAD system that is more reliable and robust.
11. Many nodules are non-cancerous and should not be classified as cancerous, leading to inappropriate diagnosis.

1.5 BRIEF OVERVIEW OF THESIS CONTRIBUTIONS

The significant contributions of this thesis include solving three tasks: detection, segmentation, and classification. To address these three challenges, new techniques are proposed. The contributions of this thesis are listed below:

1. Lung Cancer Detection system to identify nodules and non-nodules from raw CT scans.
2. Lung Cancer Nodule Segmentation model to extract ROI from CT scans.
3. Lung Cancer Nodule Classification system on CT scans.

1.6 ORGANIZATION OF THE THESIS

The thesis advances in 6 chapters. An outline of each chapter is given below.

- **Chapter 1 : The Introduction** section covers the discussion on lung cancer, various imaging modalities, the challenges in lung cancer detection, and the need for CAD systems. The chapter ends with a brief overview of research contributions and a thesis outline.
- **Chapter 2 : Literature Review** section mainly consists of a detailed review of the various conventional/traditional and deep-learning CAD systems developed for detecting and classifying lung cancer nodules.
- **Chapter 3 : Lung Cancer Detection** includes the task of identifying nodules and non-nodules from raw CT scans using various machine learning and deep learning techniques.
- **Chapter 4 : Lung Cancer Nodule Segmentation** covers a novel deep learning segmentation algorithm proposed to perform region of interest (i.e., nodule) extraction from the CT scan.
- **Chapter 5 : Lung Cancer Nodule Classification** discusses further categorization of nodules into malignant and benign nodules using novel deep learning architectures.
- **Chapter 6 : Conclusions and Future Scope** chapter summarizes the contributions and findings of this research work with future scope.

CHAPTER 2

LITERATURE REVIEW

In this chapter the literature review of existing lung cancer CAD systems is discussed in detail. The research in detecting pulmonary nodules in the lungs has been a challenging problem for decades. Several methods have been proposed to perform this task in the recent decade. However, not all the systems are reliable enough to deploy in a real-time environment. Some of the deep learning methods have resulted in promising outcomes in terms of evaluation metrics like accuracy, sensitivity, free-response receiver operating characteristic (FROC), and so on. The high performances among the deep learning methods are ensemble Convolution Neural Network (CNN) and deep learning-based segmentation algorithms. The most popular CNN methods resulting in higher performance include combination of two or more conventional CNN models. With this, we can conclude that ensemble methods provide better results as it combines the advantages of two or more networks to learn input features better. This chapter covers a broad range of current works published in lung cancer nodule detection and classification. Also, the overall findings based on the best techniques available in the literature for both traditional and deep learning systems are provided. A set of research gaps evolved from a thorough literature review is listed at the end of the chapter, along with the problem definition and objectives. The review here focuses on two main parts:

1. Conventional/ Traditional Methods
2. Deep learning Methods

2.1 CONVENTIONAL/ TRADITIONAL METHODS

Traditional techniques include machine learning techniques such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN) etc. Table [2.1](#) consists review of research done on lung cancer using traditional machine-learning techniques. The contents of these tables include authors' name, year of publication, dataset(s) that the authors have used, feature extraction techniques, segmentation techniques, classifiers used, performance measures for evaluating the system, advantages and limitations of the work.

2.2 DEEP LEARNING METHODS

Deep learning architectures include CNN, residual networks, LeNet etc. Table [2.2](#) briefly discusses popular deep-learning architectures used to segment and classify lung cancer nodules. Table content includes authors' name, year of publication, the dataset used by authors, methods used, performance measures, advantages, and limitations of the work.

Table 2.1: Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved

Authors	Dataset(s) used	Techniques used				Performance Measures	Remarks	
		Pre-processing	Segmentation	Feature Extraction	Classification		Advantages	Limitations
Potghan et al. (2018)	NSCLC, LIDC	N/A*	Lung Volume Segmentation using local thresholding method and K-Mean Clustering	Gray Level Co-occurrence Matrix (GLCM), Contrast, dissimilarity, Homogeneity, Angular Second Moment (ASM), Energy, Mean and Standard deviation	KNN, Multilayer perceptron (MLP)	Accuracy=98.30% (KNN), Accuracy=98.31% (MLP)	Results are commendable	Data imbalance is not handled as well as it is not preprocessed
Javaid et al. (2016)	LIDC and Private data from Mayo Hospital Lahore, Pakistan	Contrast stretching (normalization)	Intensity thresholding and morphological operations	Intensity, geometric and statistical features	SVM	Sensitivity=91.65% and FPs/case=3.19 and Accuracy=96.22%	Proposed system is assessed with different types of nodules ranging from small to large, including complex ones like juxtavascular and juxtaleural nodules	Proposed CAD system is not efficient in the detection of ground glass nodules (GGO) with low intensity values
Aggarwal et al. (2015)	Cancer Imaging Archive database	Median filter	Thresholding	Area, perimeter, roundness, solidity, eccentricity, equivalent diameter, centroid and convex-area	Linear discriminant analysis (LDA)	Accuracy=84%, Specificity=53.33%	This work provides prior classification of nodules and normal anatomy structure effectively	Data used for training and testing is only 90 and 150 images respectively. Also, Specificity is quite less
Kuruville and Gunavathi (2014)	LIDC	N/A	Morphological operations	Mean, standard deviation, skewness, kurtosis and Fifth and sixth central moment	Feed forward and feed forward back propagation neural networks	Accuracy=93.3%, Specificity=100%, Sensitivity=91.4% and Mean square error=0.998	Two new training functions are proposed in this paper	Sensitivity is 82% with 2 FPs/scan, but when increased to 91.4%, also increases 30 FPs/scan
Gong et al. (2018b)	LUNA16 and ANODE09	optimal threshold for lung parenchyma segmentation by applying OTSU threshold algorithm	3D tensor filtering and local shape feature analysis	Spherical features, intensity features, texture features and location features in two-dimensional and three-dimension	Random Forest (RF), J48 Decision Tree, SVM, Logistic Regression	Sensitivity=84.62% with False positive/scan=2.8	A novel approach is proposed for lung nodule detection by incorporating a 3D tensor filtering algorithm with the use of local image feature analysis	Performance of the proposed method is less

*N/A-Not Available

Table 2.1: Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used				Performance Measures	Remarks	
		Pre-processing	Segmentation	Feature Extraction	Classification		Advantages	Limitations
Li et al. (2018)	LIDC and General Hospital of Guangzhou Military Command	Gabor filter	Improved random walk method	Intensity, geometric and texture features	RF	Sensitivity=92% and Receiver-operating-characteristic curve=0.95 (LIDC), MeanSensitivity= 85% and MeanSpecificity= 82% (Private hospital data)	Proposed approach is tested on 2 datasets and an improved RF classifier is suggested	Performance of the system reported on the private hospital data is quite less, which makes the model less reliable
Farag et al. (2017)	LIDC	Scan filtering	Lung tissue segmentation	Gabor filter, Local binary pattern (LBP) and signed distance fused with LBP	SVM and KNN	Obtained AUC-ROC=0.99 and average f1-score= 0.975	Higher overall AUCs and f1-scores are obtained	Hundreds (e.g., LBP) or thousands (e.g., Gabor) of features are calculated, which slowed down the learning process
Wu et al. (2016)	LIDC	Gaussian noise and median filter	Threshold-based segmentation	Mean, variance, morphological, invariant moments, area, diameter, long and short axis, circularity, compactness, contrast, correlation, angular second moment and homogeneity	Particle Swarm Optimization-based Relevance Vector Machine (RVM)	Accuracy=79.4, Sensitivity=72.7%, False Negative Rate=27.3%	Overall scheme of the PSO-based SVM classification model is well designed	Classification accuracy is still not good enough
Rajan et al. (2019)	Lung cancer data	N/A*	N/A	N/A	Multi-class neural networks	Accuracy=100%	Development of a 3D technique for segmenting lung nodules	Performance of the system is comparatively less
Oğul et al. (2015)	JSRT	Re sampling and region of interest extraction	Gray-level transform, intensity based blob detection algorithm, Lindeberg's multi-scale blob detection scheme	Position, texture, intensity, Gaussian, detector and gradient features	SVM	Sensitivity=80% with False positive/scan=6.4	Five different set of nodule detection technique is applied and the best one is chosen	Proposed approach uses a large number of features and the false positive/scan is very high
He et al. (2014)	LIDC	N/A	Threshold-based segmentation	Texture and morphological based	ANN	Accuracy=78.9%	Stage wise classification is performed	Results are given only on 4 images

*N/A-Not Available

Table 2.1: Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used				Performance Measures	Remarks	
		Pre-processing	Segmentation	Feature Extraction	Classification		Advantages	Limitations
Cao et al. (2020)	LIDC	Spatial and frequency domain are used for image contrast enhancement	Thresholding	Weber local descriptor (WLD)	SVM	Specificity=99.15% Sensitivity=98.73% and Accuracy=98.94%	A novel and effective pulmonary nodule detection framework is proposed	Only 84 cases with 105 nodules were considered which makes model less reliable
Dolejsi et al. (2009)	Lung TIME, LIDC and ANODE09	Multi-scale filtering using a Gaussian filter	Thresholding and binary image processing	Intensity, volume, shape and other measurement	Asymmetric AdaBoost classifier	Sensitivity=78.68% and 4.61 FPs/slice (ANODE09) Sensitivity=89.62% and 12.03 FPs/slice (LIDC) Sensitivity=94.03% and 5.46 FPs/slice (Lung TIME)	Proposed approach is tested on 3 datasets	False positives per slice for LIDC is more
de Carvalho Filho et al. (2017)	LIDC	N/A*	N/A	Texture features using phylogenetic diversity	SVM and Genetic Algorithm (GA)	Accuracy=92.52%, Sensitivity=93.1% and Specificity=92.26%	The performance of the proposed approach is good and best model is selected using GA	No segmentation is performed, however only larger nodule bound is chosen as given by up to 4 experts
Manikandan and Bharathi (2011)	N/A	Wiener filter	Region based segmentation	Diameter, shape and intensity	Fuzzy Inference System (FIS)	Sensitivity=92.3%	Use of FIS in this domain inspired researches to explore further	Results are tested on only 50 cases
de Carvalho Filho et al. (2014)	LIDC	Quadratic enhancement for selective contrast increase, Gaussian Filter and median filter	Quality threshold (QT) algorithm	Shape and texture based features	SVM and Genetic algorithm	Sensitivity=85.91% Specificity=97.70% Accuracy=97.55% with a FPR of 1.82 per exam and 0.008 per slice	The performance of the proposed approach is good	Only scans out of 1012 scans are considered and also the sensitivity of the proposed model is less
Sweetlin et al. (2018)	Private (300 CT scans)	N/A	Otsu's thresholding	Texture, shape and run length based features	SVM and Naive Bayes (NB)	Accuracy=94.36%	A positive impact of choosing relevant features on the performance of the classifier is shown	Only 305 nodules were considered which makes model less robust. Also, noise removal was not performed

*N/A-Not Available

Table 2.1: Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used				Performance Measures	Remarks	
		Pre-processing	Segmentation	Feature Extraction	Classification		Advantages	Limitations
Gong et al. (2018a)	Private data from Shanghai Pulmonary Hospital and NSCLC-Radiomics	Multi-scale 3D dot filtering	A hybrid 3D semi-automatic segmentation	Texture, histogram and shape based	Support vector machine, Naive Bayes classifier and Linear discriminant analysis	AUC values are 0.88 and 0.99 for detecting early and advanced stage nodules	A positive trend between CADx performance and cancer progression stage is presented	Dataset is small, with 243 cases and is unbalanced with 76 benign nodules less as compared to malignant nodules
Kulkarni and Panditrao (2014)	LIDC	Image Smoothing, Image Enhancement (Gabor filter)	Marker based watershed	Area, perimeter and eccentricity	SVM	Area= 1328, Perimeter= 162, Eccentricity=0.960	Stage wise classification is performed	Results are given only on 4 images
Liu et al. (2017)	LIDC	N/A*	Two-dimensional OTSU's curve thresholding	22 texture and shape features	Ensemble classifier and RF	Accuracy=93.2%, Sensitivity=92.4%, Specificity=94.8%, AUC=97.6%	A novel pulmonary nodule segmentation method is proposed	FP/Scan of this work is slightly higher than some of other approaches
Aresta et al. (2017)	LIDC	Multi-scale filtering using a Gaussian filter	Intensity based lung volume segmentation	Intensity, Hessian values, gradient, geometry and distance to the pleura	Support vector machine with radial basis function	Sensitivity is 57.4% with 4 FPs/scan	All solid, sub-solid and non-solid nodules are considered as opposed to other methods	Sensitivity is very less
Alilou et al. (2014)	LIDC	N/A	Multiple thresholds followed by morphological opening and 3D region growing algorithm	3D geometrical, 3D intensity-based, 2D geometrical and 2D intensity-based features	Rule-based procedure and SVM	Sensitivity=80.0% with False positive/Scan=3.9	The proposed method is time-efficient	Performance of the proposed method is less
Mehre et al. (2016)	LIDC	N/A	Hard thresholding with attenuation threshold	Geometry, and intensity-based statistical features	SVM	Sensitivity=92.91% and False positive per slice=3	Development of novel method for balancing the data without loss of information	Large number of features are used, instead feature selection methods could have been used
Dandil (2018)	CT Lung Data set from Sincan Nafiz Korez Hospital	Lung volume extraction method (LUVEM)	Self-organizing maps (SOM).	Intensity, shape, texture, energy and combined features	Probabilistic neural network (PNN)	Specificity=94.24%, Sensitivity=97.42% and Accuracy=95.91%	New preprocessing technique is proposed for image enhancement	Results are given only on 38 images

*N/A-Not Available

Table 2.1: Comparison of various Machine Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used				Performance Measures	Remarks	
		Pre-processing	Segmentation	Feature Extraction	Classification		Advantages	Limitations
Liu et al. (2014)	LIDC	Selective enhancement filter	Region growing and thresholding	Mean, intensity value, intensity variance, kurtosis, minimum intensity value, maximum intensity value, size, overlap value, circular shape descriptor value	Conditional random field (CRF)	Sensitivity=89.30%	A novel and effective pulmonary nodule detection framework is proposed	Only 59 nodules from 38 patients are considered for both training and testing, which is less data
Akram et al. (2016)	LIDC	N/A*	Thresholding, initial label masking, background removal, connected component labeling, morphological operators and contour correction	Geometric and Intensity based statistical features	SVM	Accuracy=97.52%, Sensitivity=95.31%, Specificity=99.73%	The performance of the system is good as there is high specificity for the dataset	Proposed approach uses a large number of features of 2D and 3D geometry
Santos et al. (2014)	LIDC	N/A	Thresholds, region growing algorithm and Hessian matrix	Texture features	SVM	Accuracy=88.4%, Sensitivity=90.6% and Specificity=88.5%	The performance of the proposed approach is good and multiple segmentation was applied to segment different regions in the images	Only 28 scans is considered for the comparison of the methods which is not a fair comparison
Yokota et al. (2014)	LIDC	3D line filter	Thresholding	Density and shape based statistical features	ANN	Accuracy=93.0%	Removal of vessel regions has improved the performance of the proposed system	All GCO based features are used, instead feature selection methods could have been used to provide an efficient model
de Sousa Costa et al. (2018)	LIDC	N/A	N/A	Texture features using phylogenetic distance and taxonomic diversity	SVM	Sensitivity=93.42% Specificity=91.21% Accuracy=91.81% and Area under the ROC curve=0.94	A useful tool for expert physicians	No preprocessing done to remove noise

*N/A-Not Available

2.3 OVERALL FINDINGS FOR CONVENTIONAL CAD SYSTEMS

There are multiple systems developed previously for performing lung cancer detection and classification. The best results obtained from various methods are mentioned below.

2.3.1 Pre-processing

This step is considered one of the important steps in building efficient CAD systems. It is necessary to improve the image quality before performing any other step. In the previously proposed CAD systems, the majority of the pre-processing involved applying filters and normalization techniques to the input image. Amongst multiple filtering techniques, the techniques that provided the best performances are contrasted stretching (normalization technique), gabor filter, image contrast operations from spatial and frequency domains, Weiner filter, and 3D line filters. The application of the filters enhances the quality of the input images, which helps in the extraction of better features.

2.3.2 Segmentation

As mentioned earlier, the nodule size can vary widely, and some images can consist of small nodules, which in turn have much background. The background information is considered noise. Therefore, it needs to be removed. To perform this operation, segmentation of only the region of interest is done on the input images. Only the region around the nodule location is cropped and considered for feature extraction. Some of the segmentation techniques that helped in the better performance of the CAD system are local threshold segmentation, intensity threshold segmentation, morphological operations, and OTSU threshold segmentation. These segmentation techniques help identify only the nodule region and remove unwanted background information.

2.3.3 Feature Extraction

The performance of the CAD system depends on the features extracted from the input images. They play a vital role in providing better results. Multiple sets of features can be extracted from the images. Some of the common feature representations that

proved to provide reliable results for lung cancer detection CAD systems are Gray level co-occurrence matrix (GLCM), contrast, dissimilarity, homogeneity, statistical features such as energy, mean, standard deviation, skewness, kurtosis, Fifth and sixth central moment, intensity, geometric, texture features, and Weber Local Descriptor.

2.3.4 Classification

The lung nodule classification is the ultimate goal in lung cancer CAD systems. Various classifiers are used for performing the classification task. However, not all classifiers perform well and result in reliable performance. The classifiers that resulted in better performance for lung nodule classification tasks are K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Multi Layer Perceptron (MLP), Random Forest (RF), Fuzzy Inference System (FIS), Probabilistic Neural Network (PNN), and Artificial Neural Network (ANN). These classifiers resulted in reliable results in classifying lung cancer nodules.

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Hua et al. (2015)	LIDC	Deep belief network and CNN	Sensitivity=73.4%, Specificity=82.2% (DBN) and Sensitivity=73.3%, Sensitivity= 78.7% (CNN)	First research to apply deep learning techniques to the problem of pulmonary nodule classification	The major drawback of the deep learning techniques used in this study lies in the resizing issue of the input images and it resulted in low results
Kasinathan et al. (2019)	LIDC	Multiscale Gaussian distribution to smoothen CT images, Enhanced CNN	Accuracy= 97%, Sensitivity= 89%, Specificity= 91%	Accuracy is improved, and model takes less computation time	Increasing the sensitivity also increased false positives
da Nóbrega et al. (2018)	LIDC	Features extracted through VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNet-V2, DenseNet169, DenseNet201, NASNetMobile, NASNetLarge and classified through Naive Bayes, MLP, SVM, KNN and RF	Accuracy= 88.41%, AUC=93.19%, F-Score=78.83%, TruePositiveRate= 85.38%, and PositivePredicted-Value= 73.48%	Various models are trained and tested in this work It is believed that based on the performance and parameters of the models used, it will help researchers to choose a model for their problem	F-Score and Positive Predicted Value are less and still can be improved
Manikandan and Bharathi (2011)	LIDC	Median filter, Convolution Neural Network	Accuracy= 96%	Performance of the system is good	Data used for training the model is only 70 images and for testing is 30 images, so the reliability of the model is still a question. It is believed that CNN requires more data to train the network
Tran et al. (2019)	LUNA	Novel 15-layer 2D deep convolutional neural network architecture	Accuracy=97.2%, Specificity=97.3% and Sensitivity=96%	A novel architecture which helped in achieving very good results	15 layer architecture needs more computational resources. So, in order to make this model useful to low processing power systems, optimization is required
Hamidian et al. (2017)	LIDC	3D fully convolutional network (FCN)	Sensitivity= 80% and False positives per case= 22.4	A 3-D FCN model is proposed	The sensitivity of the model is less and the false positives per case is very high which has to be improved

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Jakimovski and Davcev (2019)	Image & Data Archive of the University of South Carolina and the Laboratory of Neuro Imaging (LONI) database	Convolutional Deep Neural Network (CDNN) and a regular CDNN	HighestAccuracy=99.62% (CDNN), HighestAccuracy=87.6% (regular CDNN)	The highest result achieved is commendable	The drawback in the model is that the minimal value of certainty is decided manually and is accepted as being satisfactory. Deciding this value by the user, requires lot of domain knowledge, as a single wrong threshold may lead to misclassification
Kim et al. (2016)	In-house data (20 subjects)	Stacked Denoising AutoEncoder (SDAE), linear SVM	Accuracy=95.5%, Sensitivity=94.4%, and AUC=0.987	Data Augmentation is done to balance malignant and benign nodules	Number of subjects are very less to judge the reliability of the model. The model can be further tested on LIDC data
Setio et al. (2016)	LIDC, ANODE09 challenge and DLCST	Multi-View Convolutional Networks	Sensitivity=85.4% at 1 false positive per scan and 90.1% at 4 false positives per scan	3 datasets are used for showing the performance of the model	Performance can still be improved, by increasing the sensitivity and reducing the false positives per scan
Sahu et al. (2018)	LIDC	Lightweight Multi-Section Convolution Neural Network	Accuracy=93.18%	The proposed model is lightweight which makes model to be ported to mobile devices	A cloud based paradigm is needed where a model exploits GPU's capabilities to determine and fetch the salient sections quickly
Kumar et al. (2015)	LIDC	Deep features extracted from an auto encoder, Binary decision tree as a classifier	Accuracy=75.01%, Sensitivity=83.35%, and False positive=0.39/patient	FPR is reduced and the deep features not only take the different conventional semantic features like lobulation, spiculation etc. in to account but they also take into account the association between them	Accuracy is less for the model to be deployed in a hospital
Liu et al. (2019)	Private data (60 patients)	Low-level features such as color moment and texture feature, ResNet for classification	Diagnostic accuracy=100% (experimental group) and 74.29% (control group)	The obtained accuracy for diagnostic group is 100%, which is good	However, to check the robustness of model, it would be good if the performance is evaluated on publicly available large data sets such LIDC

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Rahman et al. (2019)	LIDC	Preprocessed through blurring and thresholding, Classified through MobileNet, VGG- 8 and Inception-v3 deep neural network models	Accuracy=97%, Specificity=97.85% and Sensitivity=96.26%	Model in low processing power computer, which in turn helps doctor to detect nodules and take better decisions	Images from DICOM format are converted to JPEG format using Microsoft Paint for easy processing, however it loses lot of information
Xie et al. (2018)	LIDC	Multi-View Knowledge-based Collaborative (MV-KBC) Deep Learning	AUC=95.70% and Accuracy=91.60%	Results are superior than various SOTA approaches	There are 27 ResNet-50 models embedded in the system. This makes the proposed MVKBC system require a highly computational complex server for training it
Bonavita et al. (2020)	LIDC	CNN	F1-weighted score=11.8%	This was the first attempt to build nodule-malignancy/patient-cancer integrated framework to quantify nodule malignancy	Performance of the model in terms of precision, recall is less
Shakeel et al. (2020)	Cancer Imaging Archive (CIA)	Improved deep neural network and ensemble classifier	Accuracy=96%, Specificity=98%, Precision=97%, Recall=98% and F1-score=98%	Results achieved from the model are good	Model was tested on very few samples
Rehman et al. (2020)	JSRT	CNN	Accuracy = 88%	The model performed better than some of the high complex classification models	The dataset contains only 100 malignant and 54 benign nodules. This makes model less robust.
Sori et al. (2021)	Kaggle Data Science Bowl 2017 challenge (KDSB) and LUNA 16	Denoising first two-path convolutional neural network (DFD-Net)	Accuracy=87.8%	An image denoising technique is adapted in this work to improve the performance of the nodule detection task. The method is less complex in terms of computation time.	The accuracy of the method is too less to be deployed in real-time environment.
Surendar (2021)	LIDC-IDRI	Deep neural network with adaptive sine cosine crow search (DNN-ASCCS)	Accuracy=99.71%	The authors have proposed new segmentation, feature selection and classification methods for lung nodule classification.	The results are only reported for less number images which may or may not give similar performance for larger datasets.

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Zhang et al. (2019)	LUNA16	Texture features, shape features, 3D deep DPN features, Gradient boosting machine (GBM) for classification	AUC=0.9687% and Accuracy=93.78%	Relevant features are considered and it helped in yielding good results	Used network is complex and there is still scope for optimization
Ypsilantis and Montana (2016)	LIDC	Thresholding-based region for segmentation, Recurrent Convolutional Networks (ReCTnet) for classification	Sensitivity=90.5%, False positives per scan=4.5	Results from this work suggests that LSTM layers enable to better synthesize the anatomical information across adjacent slices eventually resulting in improved discrimination ability	Having such a deep architecture should not need segmentation as an extra step, as model becomes more complex
Fu et al. (2017)	LIDC	Thresholding, 2D CNN, SVM Morphological operation, 3D region growing, hand-crafted feature extraction	Sensitivity=89% at False positive=4/scan and Sensitivity=71.6% at False positive=1/scan	Combination of handcrafted features and CNN features from both lung CT images and enhanced images is proved to be a promising method for lung nodule detection	When FP/scan is 1, sensitivity is very less making model less efficient
de Carvalho Filho et al. (2018)	LIDC	Otsu algorithm, new indexes of phylogenetic diversity based on topology, CNN	Accuracy=92.63%, Specificity=93.47%, Sensitivity=90.7% and receiver operating characteristic curve of 0.934	The model performed very well given larger number of samples and achieved good results in terms of accuracy, sensitivity, specificity and AUC	There is still scope for improving the sensitivity of the model by performing some modifications in the CNN architecture
Wang et al. (2018)	LIDC	Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP), Hybrid CNN network model	Receiver operating characteristic curve (AUC) from 0.9441 to 0.9702	The results achieved are good	The model used for the classification is a complex model and is computationally expensive
Nibali et al. (2017)	LIDC	Deep residual learning, curriculum learning, and transfer learning	Accuracy=89.90%, Specificity=88.64%, Sensitivity=91.07%, Precision=89.35 and receiver operating characteristic curve of 0.9459	Data imbalance of benign and malignant nodules is handled by preprocessing the data	Performance can still be improved, by increasing the sensitivity and precision
da Silva et al. (2017)	LIDC	Otsu algorithm, Particle Swarm Optimization, Genetic algorithm, Convolutional Neural Network	Accuracy=94.78%, Specificity=95.14%, Sensitivity=94.66%, Area under curve=0.949	The results achieved by the model are good	Many preprocessing techniques are applied to the CT scans which in turn increases the overhead of the model and may not be robust

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Shen et al. (2015)	LIDC	Multi-scale Convolutional Neural Network (MCNN)	Accuracy=86.4%	Multi-scale patches of scans are used rather than segmented patches	The performance of the model is low and can be improved by using advanced deep learning architectures
Sun et al. (2017)	LIDC	CNN, DBN, and stacked denoising autoencoder (SDAE)	Area under the curve (AUC)= 0.899 ± 0.018	Three different state-of-art deep neural network architectures have been explored	For critical application like lung cancer detection, the performance needs to be higher to be deployed in real-time environment. Therefore, there is still scope for improving the performance of the model.
Liu et al. (2018b)	LIDC	Novel end-to-end deep learning architecture named dense convolutional binary-tree network (DenseBTNet)	Accuracy=88.31% and Area under curve=0.9335	A novel deep learning architecture DenseBTNet is proposed	The performance of the model can be improved by optimizing the network and a center crop operation is performed in the architecture which is computationally expensive and may not work for all the cases
Liao et al. (2019)	LUNA16 and the training set of Data Science Bowl 2017	Gaussian filter, 3-D DNN	Accuracy=81.42%	Volumetric information is fed to the network and a novel 3-D Convolutional neural network is proposed	The accuracy obtained from this model is too low for using in lung cancer CAD system and it makes use of 3-D volumetric features, which increases computational complexity.
Zhao et al. (2018)	LIDC	A hybrid CNN of LeNet and AlexNet	Area under the curve (AUC)= 0.822 to 0.877	The effects of kernel size, learning rate, training batch size, dropout, and weight initialization were investigated on CT images for pulmonary nodule classification using new Agile CNN	Both the used architectures such as LeNet and AlexNet are designed for color images, while the medical images are gray scale images, this results in an inability to make full use of all channels. Due to this, all the three channels of input are homogeneous
Tajbakhsh and Suzuki (2017)	Low-dose thoracic helical CT (LDCT)	massive-training artificial neural networks (MTANNs)	Accuracy=88.06%	A comparison between massive-training artificial neural network and deep convolutional neural network is given	The accuracy of the model is less and can be improved by optimizing the network
Lakshmanaprabu et al. (2019)	ELCAP	Modified Gravitational Search Algorithm applied to train Optimal Deep Neural Network, LDA.	Sensitivity=96.2%, Specificity=94.2%, and Accuracy=94.56%	An innovative approach is proposed and is claimed to have speed, also is simple to operate, with properties like non-invasive and cheap	The use of 70 images for training and 30 images for testing are too less for lung image investigation

Table 2.2: Comparison of various Deep Learning CAD Systems for Lung Cancer based on techniques used and performance achieved (continued)

Authors	Dataset(s) used	Techniques used	Performance Measures	Remarks	
				Advantages	Limitations
Neal Joshua et al. (2021)	LUNA16	3D AlexNet with lightweight architecture	Accuracy=97.17%	The method uses 3D AlexNet with lightweight architecture implying that model complexity is less as compared to other deep learning models	There is a loss of information in the feature map of the CNN layer.
Faruqi et al. (2021)	Medical IoT (MIOT) dataset	LungNet	Accuracy=96.81% FPR = 3.35%	The authors have performed data collection from wearable devices along with latent features extracted from CT scans. The method provides efficient results.	FPR value reported here is very high. Where, FPR is the proportion of negative data instances that were incorrectly reported as positive. This needs to be reduced.
Heuvelmans et al. (2021)	NLST	Lung Cancer Prediction Convolutional Neural Network (LCP-CNN)	AUC=94.5% Sensitivity = 99%	The model is used to identify benign and malignant nodules using deep learning methods. The model displayed high sensitivity rates.	The model has been trained on fewer CT scan images.
Tian et al. (2021)	LIDC	Enhanced Capsule Networks (ECN)	Accuracy= 96.65% Precision = 96.35% Recall = 96.07% F1-score=96.41%	The method uses an optimized fuzzy possibilistic c-ordered mean based algorithm called Converged Search and Rescue (CSAR) and ECN for final diagnosis. The optimization algorithm resulted in promising results.	The model is computationally complex.
Marentakis et al. (2021)	NSCLC radiomics	Long short-term memory (LSTM) + CNN	Accuracy = 74% AUC = 0.78	The authors have experimented on combinatorial methods to achieve better performance to classify lung cancer histology images. The best results are achieved from the combination of LSTM, CNN, and radiomics.	The results can be further improved.
Feng and Jiang (2022)	Private dataset	Mask RCNN and Dual Path Network (DPN)	Accuracy = 97.94% Sensitivity = 98.12% Specificity = 100%	The method has resulted in high performance.	The dataset used in the work consists of CT and MRI images of 45 patients, which is less data to generalize the results.
Kasinathan and Jayakumar (2022)	LIDC-IDRI	Cloud-based Lung Tumor Detector and Stage Classifier (Cloud-LTDSC)	Accuracy = 98.6%	The proposed method uses a cloud-based method to identify lung cancer in CT and PET scans. The performance reported in the work is good.	From 50 images it is not feasible to generalize the method.

2.4 OVERALL FINDINGS FOR DEEP LEARNING-BASED CAD SYSTEMS

The main reason for using deep-learning-based CAD systems is to overcome the issue of using hand-crafted features for performing lung cancer detection and classification. There is no need for performing multiple steps before classification, such as feature extraction, segmentation, or feature selection. However, the input images may be pre-processed before being input to the network. Also, different feature representations or combinations of these feature representations can be provided as input to the deep neural networks. The deep learning models take the images as input to the network, learn meaningful feature representations from the input image and provide classification output.

Best-performing results are obtained from deep learning architectures such as CNN, Enhanced CNN, MobileNet-V2, VGG-8, Inception-V3, Convolutional Deep Neural Network (CDNN), and Stacked Denoising Autoencoder (SDAE). The use of deep-learning-based lung cancer detection and classification CAD systems proved that the results obtained from these models are better in terms of performance and time complexity. As the number of steps is reduced in deep models, the CAD systems are faster and also efficient.

2.5 RESEARCH GAPS

Based on the literature survey, some research gaps are identified and listed below. These are potential future prospects for developing a better and improved CAD system for lung cancer identification.

1. Conventional algorithms work better when suitable hand-crafted features are chosen for the classification task. This task has become simpler with the recent development of deep learning because these techniques do not require hand-crafted features to train the classifier. So, there is a lot of scope to build deep learning based CAD systems for lung cancer detection.
2. In conventional CAD systems, annotations are pretty expensive as it has to be done manually by experienced radiologists. With the help of deep learning,

annotation costs can be reduced by automating the annotation process. This can be done by training a deep learning model with a huge annotation dataset. This model can now be used for unseen CT scans' annotation.

3. There is a need to provide clinically relevant explanations for the features discovered by the different learning algorithms as it makes the CAD system more reliable for real-time usage.
4. Micro nodules (nodules < 3 mm) are usually truncated during the training of the system as it is difficult to detect this set of nodules. These sets of nodules can provide better insight for early detection of lung cancer nodules.
5. Selection of relevant features and optimal classifiers are required to achieve significant results, which can help doctors consider it a second opinion.
6. The system can be designed in such a way that it learns from both clinical records and medical images for multi-modal analysis.
7. Segmentation of pulmonary nodules plays a fundamental role and helps improve the sensitivity of CAD systems. Hence it needs to be further explored.
8. In the literature, many authors have considered only a subset of data which could be a better practice as it may not be a robust model, which in turn affects the system's performance in realistic scenarios. Hence, there is a need to evaluate CAD systems on larger datasets. More experiments on larger datasets will improve the systems' performance and generalizability.

2.6 PROBLEM STATEMENT

Design, analysis and implementation of a deep learning-based decision support system for lung cancer detection, nodule segmentation and classification of nodules into benign and malignant.

Problem Description: The problem is divided into three tasks: detection, segmentation, and classification. The detection task is identifying nodules and non-nodules from a raw CT scan. CT scan usually consists of tissue structures or

organs which look similar to a nodule. This kind of anomaly must be eliminated to find cancer nodules from the CT scans. The detection task is followed by a segmentation task in which a particular nodule region is extracted. Intensity similarities between nodules and vessels make it challenging to perform accurate segmentation. Also, tumor size in a CT scan is very small compared to other organs, and extracting such a small region of interest is difficult. Finally, nodules are classified into malignant and benign nodules. The main motive behind classifying lung cancer nodules into malignant and benign is to improve the lifespan of the patients.

The objectives of the work are:

1. To build an automated lung cancer detection system with reduced false positives to identify nodules and non-nodules from raw CT scans.

Description: One of the significant challenges in lung cancer detection is identifying lung nodules into nodules and non-nodules. This objective aims to propose a reliable CAD system that can differentiate between nodules and non-nodules. Also, to perform detection, data augmentation techniques need to be applied to avoid bias in the data. To solve these problems, novel deep learning approaches are proposed. Deep learning has proved to provide better performance in terms of medical image analysis. The goal is also to design a lung cancer detection system with a reduced false positive rate with assurance and reliability.

2. To develop an efficient deep learning-based segmentation model to extract region of interest from lung CT scans.

Description: Nodule segmentation plays an important role in the classification of lung cancer nodules. In the CT scan, the nodule size is too small compared to the other organs in the scan. It is necessary to identify the region of interest, i.e., the nodule, rather than other tissues/organs in the CT scan. In order to extract the region of interest, an accurate segmentation needs to be performed on the CT scan and crop only the region where the nodule exists. This helps in improving the classification results of the benign and the malignant nodule. This objective aims

to segment the lung nodules from the CT scan images using new deep learning-based segmentation algorithms.

3. To design and develop a novel deep learning framework for classification on CT scans.

Description: One of the important issues that exist in lung cancer classification is classifying non-cancerous (benign) and cancerous (malignant) nodules accurately. The main motive behind classifying lung cancer nodules into malignant and benign is to improve the lifespan of the patients. This objective aims to resolve this issue using various feature extraction and classification methods. In recent trends, deep learning is exhibiting better performance for these tasks. Therefore, the main aim is to improve the accuracy of benign and malignant nodule classification using deep learning architectures.

2.7 DATASET USED FOR THIS THESIS

This section familiarizes the dataset used in this research work for experimental analysis. The dataset utilized for performing lung cancer detection, segmentation, and classification is LUNG Nodule Analysis (LUNA), released in 2016. The LUNA16 dataset is open, where the reference standards and images are made publicly available. LUNA16 dataset is a curated version of the Lung Image Database Consortium-Image Database Resource Initiative (LIDC-IDRI) dataset (McNitt-Gray et al. 2007), which is also a public-access dataset. Four expert radiologists have provided the annotations (Setio et al. 2017a). The main aim of the LUNA16 dataset was to develop large-scale automated nodule detection algorithms. Table 2.3 provides details related to the LUNA dataset.

The dataset includes 888 CT scans taken from the LIDC-IDRI dataset. The dataset excludes CT scans with a slice thickness of more than 2.5 mm. There are a total of 1186 positive lung nodules in the LUNA16 dataset out of 5,51,065 total candidates. The candidate in two-dimension contains a lesion or lung nodule positioned in the center. The nodule size can vary, so the image has a lot of background for the nodule of a smaller size. The LUNA16 dataset consists of annotations labeled as nodules and

Table 2.3: Details of LUNA16 dataset

Parameters	Details
Dataset	LUNA16 (Setio et al. 2017a)
Date of release	2016
Dataset size (GB)	116
Number of samples	888
Image modality	CT
Image dimension	512x512
Image format	DICOM
Ground truth available	Yes

non-nodules provided in a Comma Separated Value (CSV) file. The details given in the annotation file of the LIDC-IDRI dataset are used further to split the nodule images into benign and malignant nodules (McNitt-Gray et al. 2007). Nodules are classified as benign or malignant depending on a range that four qualified and experienced radiologists determine. The scale is from 1 to 5 (1 being least malignant and five being most malignant).

2.8 SUMMARY

This chapter critically reviewed the current work in detection, segmentation, and classification of lung cancer. Also, the overall findings based on the best techniques available in the literature for both traditional and deep learning systems are provided. The research gaps are listed, and the current research work's problem statement is also included in one section of the chapter. The details of the dataset referred to are also highlighted in this chapter. The proposed methodology for each of the objectives is explained in later chapters.

CHAPTER 3

LUNG CANCER DETECTION

3.1 INTRODUCTION

Lung cancer is regarded as one of the most common types of cancer with the highest mortality rate. The symptoms are visible after it has reached an advanced stage. The diagnosis of lung cancer is performed using nodules. Nodules are abnormal, and irregular tissue formation is observed in the lung region. The size of a nodule is usually as small as 3 mm in diameter. To detect such a small nodule, radiologists must conduct a careful examination. It is a tedious and time-consuming task to provide precise and accurate information about a nodule type. This indirectly may lead to human error. This issue can be resolved by combining CT screening with computer-aided algorithmic solutions. Deploying this in a real-time environment, which involves integrating these solutions with physicians, is another challenge. The solutions must deliver precise and trustworthy outcomes. This chapter presents the method proposed to detect nodules and the results obtained from the proposed approach. In the proposed nodule detection approach, the nodules are detected by using a novel segmentation method and followed by this, the potential nodule candidates are further classified into nodules and non-nodules.

The contributions of this chapter are :

- RFR V-Net: A novel segmentation method is proposed. An investigation of multiple receptive fields (RFs) is carried out for RFR V-Net.

- A post-processing pseudo-coloring method is adapted to improve results of nodule classification.
- NCNet: A new deep learning model is introduced for lung nodule classification. Also, a comparison of both the proposed models with existing CAD systems is provided.
- The results achieved by the proposed model are verified by an expert pulmonologist.

The chapter is organized as follows: Section [3.2](#) discusses the proposed methodology in detail. Section [3.3](#) briefs the results and discussion, and Section [3.4](#) summarizes the proposed method and its significance.

3.2 PROPOSED METHODOLOGY

A schematic overview of the proposed models is shown in Figure [3.1](#). In stage 1, input CT scans are preprocessed using Hounsfield Unit (HU). After preprocessing, they are divided into multiple patches of size 96x96x16, as the computation of raw CT scans is complex and expensive. These scans are fed as an input to the Receptive-Field Regularized V-Net (RFR V-Net) to detect candidate nodules. In stage 2, the data is prepared to identify whether a candidate nodule is a true nodule or an FP candidate. The CT scans are cropped to 32x32x16, considering the center point of the candidate nodules. The reason for reducing the size of the image is that the size of the nodules is very small in the CT scan and can consist of background noise which can lead to misclassification. In stage 3, an image enhancement technique is applied to the 2D CT scan slices to post-process the image. In final stage 4, three sets of inputs are provided to the models: a 2D slice of CT scan (middle slice), a 3D CT cube, and a post-processed 2D CT slice image. These inputs are fed to three deep learning architectures: SqueezeNet, ResNet, and the proposed Nodule Classification Network (NCNet) for the nodule classification task. The details of each stage are provided in the following sub-sections

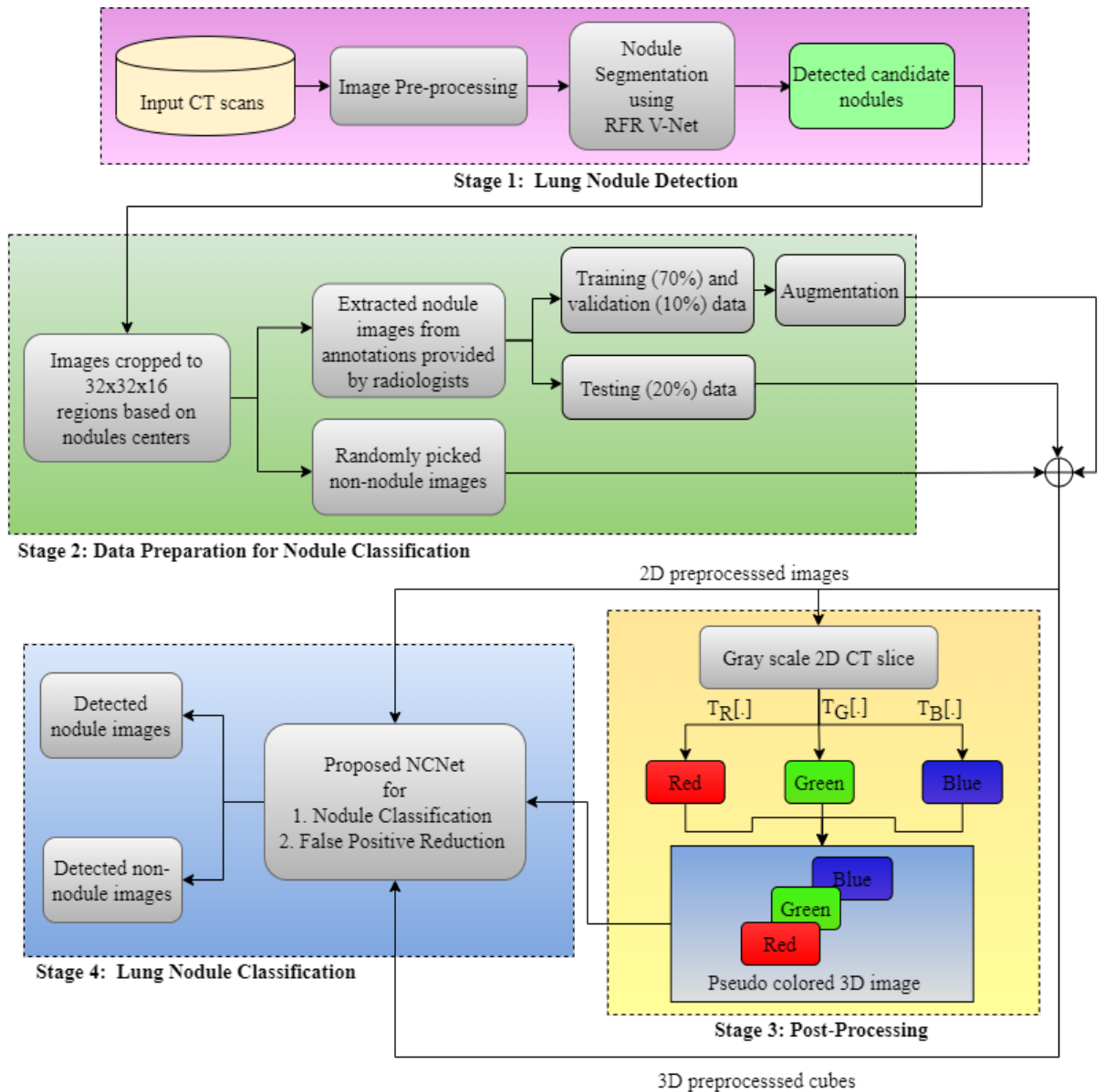


Figure 3.1: Schematic overview of the proposed model for Lung Nodule Detection

3.2.1 Stage 1: Lung Nodule Detection

The first stage of lung nodule classification involves the identification of the candidates in the CT scan. To perform this, the raw CT scan needs to be pre-processed in order to clearly differentiate between the candidate nodules and tissue/organs in the lung. Further, once the CT scans are pre-processed, RFR V-Net algorithm is applied to identify candidate nodules. A detailed explanation of the pre-processing and the proposed RFR V-Net model, is given in the subsection [3.2.1.1](#).

3.2.1.1 Image Pre-processing

The raw CT scan images of the thoracic region are usually larger and for the early stage detection, one major drawback is a region of interest being smaller in size as shown in Figure 3.2. Therefore, efficient pre-processing of these scans are necessary to attain a clearer image and to locate the areas of interest. In this work, pre-processing is performed by measuring the different radio densities of various substances such as soft tissues+, lungs, water, fat, and air. The Hounsfield unit (HU) is used as a relative quantitative measurement of radio density by radiologists to analyze CT scans. These densities are measured using the unit HU. The scans are clipped in the range of -1000 and 400 HU to discard the variations that occur in the images due to the resampling of the images and bone densities. Once the scans are clipped based on the HU, images are normalized before feeding as input to the neural network. After the normalization, the images have a standard mean voxel value of 0 and a variance of 1.

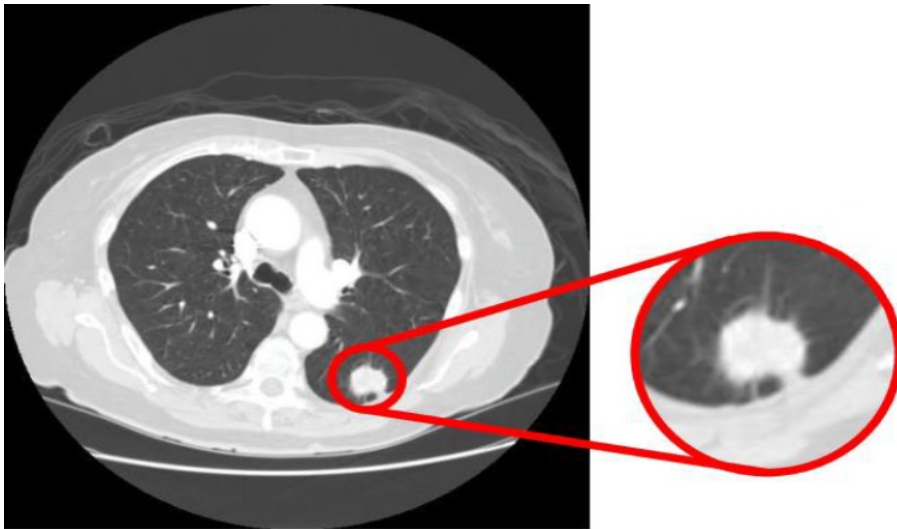


Figure 3.2: A CT Scan illustrating that pulmonary nodules occupy less than 5% of CT scan

3.2.1.2 Nodule Detection: RFR V-Net

The primary nodule detection is performed using the proposed RFR V-Net. The proposed network is a fully CNN which is a modified V-Net architecture that has demonstrated improved performance for medical image segmentation (Milletari et al. 2016). Our key contribution to the V-Net architecture is the regularization of receptive

field in convolutional layers. The V-Net architecture consists of three encoder-decoder block pairs and skip connections apart from input and output blocks as shown in Figure 3.3. The encoder block consists of downsampling convolutions and is also connected to the residual connection of the output. In contrast, the decoder block consists of upsampling deconvolutions. The convolution and deconvolution blocks are investigated for various RFs. Most CNNs use larger RFs which do not generalize well. In computer vision, image sizes vary for each problem. The RF of a convolution layer plays an important role in deciding the feature size and the information passed to the network's succeeding layer. Therefore, the term RF Regularization was coined by Koutini et al. (2020). They have shown that larger RFs often tend to overfit the model and lead to performance degradation of the system. Therefore, a regularization technique is adapted for RFs in order to improve the segmentation performance.

In a convolutional layer, the RF of a neuron corresponds to the slice of the layer input that impacts the neuron's activation. The convolutional layer's output depends on the activations of the previous convolution layer or network's input. In both cases, it is called the input feature map of a layer. The primary determinant of the size of RF is the filter size of the input feature map. The amount of input feature map affected is directly proportional to the filter's size chosen in the layer. The choice of RF in a layer affects all the RFs chosen in the previous layers of the network. This can be recursively calculated using equation 3.1 (Koutini et al. 2019).

$$\begin{aligned} S_n &= S_{n-1} * s_n \\ RF_n &= RF_{n-1} + k_{n-1} * S_n \end{aligned} \quad (3.1)$$

Where, S_n represents the cumulative stride, RF_n represents the RF of neuron corresponding to layer n , s_n represents the stride and k_n represents filter size of layer n .

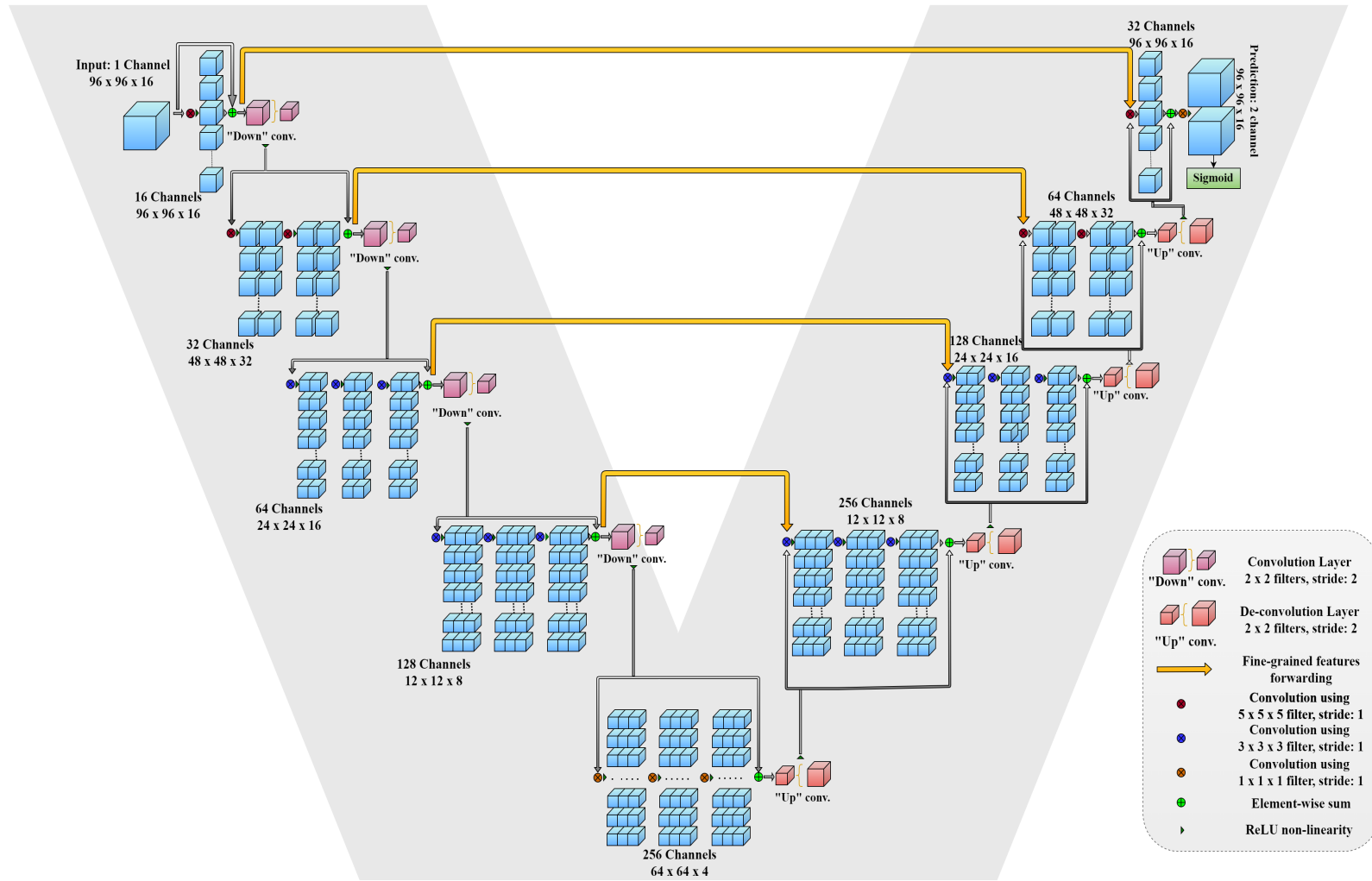


Figure 3.3: Architecture diagram of proposed RFR V-Net

The architecture diagram of the RFR V-Net is shown in Figure 3.3. Due to the large size of the raw CT scan image, the CT scan is divided into patch sizes of 96x96x16 and fed as an input to the RFR V-Net. The model consists of an encoder block with five layers consisting of filter sizes 5x5x5, 5x5x5, 3x3x3, 3x3x3, and 1x1x1, and stride of 1, respectively. As mentioned above, the RF of the network is the last convolution layer, i.e., 1x1x1. The decoder block consists of upsampling deconvolution layers with five filter sizes 1x1x1, 3x3x3, 3x3x3, 5x5x5, 5x5x5, and stride of 1, respectively. Each convolution performed in the network utilizes volumetric kernels having a filter size of $a \times a \times a$, where a represents the number of filters. In this network, the max-pooling operation is not applied; rather a strategy reducing the resolution of the data through different stages along with a compression path is performed. A 2x2x2 voxel wide kernel is applied with a stride of 2, and when only non-overlapping 2x2x2 volume patches are considered, the resulting feature maps' size is reduced by half. This method is inspired by Cicek et al. (2016). The use of pooling operations is discouraged because the main disadvantage of max pooling is that other elements are ignored and only the maximum element from the pooling region is taken into account. The distinguishing features vanish after conducting the maximum pooling operation if the majority of the elements in the pooling area are of high magnitudes. Hence, the pooling operations are replaced by convolution layers instead. Rectified Linear Unit (ReLU) is used as non-linearity in the hidden layers, whereas softmax is used for the final layer of the model. The RFR V-Net model returns images of size 96x96x16, which are potential nodule candidates in the raw CT scan images. The generated nodule candidates may also consist of False Positive (FP) candidate nodules which will be later reduced/removed by an FP reduction technique, and classification of true nodules and non-nodules is performed.

3.2.2 Stage 2: Data preparation for Nodule Classification

In the second stage, the detected candidate nodules are further classified into nodules and non-nodules. This classification is performed using the annotations provided by the radiologists in the dataset. The images are cropped to 32x32x16 regions using the nodule centers for precision. There is an imbalance in the dataset for nodule and non-nodule images. Therefore, to avoid the overfitting of the model, the nodule images are

augmented to balance the two classes. The augmentation is applied to only the training and validation splits of the dataset. The details of the augmentation methods used for nodule images are provided in subsection [3.2.2.1](#).

3.2.2.1 Image Augmentation

Once the nodule candidates are generated, the nodules need to be identified as true nodules and FP candidates. In the LUNA16 dataset, there are only 1186 true nodules and 5,51,065 candidates. Therefore, there is a high chance that some candidates are misdiagnosed as true nodules. In order to reduce this high data imbalance, the number of images of true nodules is increased by performing extensive data augmentation. Also, it strongly avoids overfitting in the deep learning models while the training step. Random image transforms of 90 degrees for zero or more times with 0.5 probability are applied to the images. Transpose of the images is also taken by row and column swap. Affine transform augmentation ([Nalepa et al. 2019](#)) is made uniformly with a translation from 0 to 1 with random scaling from 0.5 probability. A seemingly easy technique called flipping an image (and its annotations) can significantly boost model performance. The images are vertically and horizontally flipped with a probability of 0.7. The images are randomly scaled with a probability of 0.3. Also, the brightness and contrast of the images are varied with a limit of 0.4. Our models discover which clusters of pixels and the connections among those clusters indicate the presence of an item in the frame. However, machine learning models, such as convolutional neural networks, have a tendency to be quite fragile. For example, although our models may memorize a particular arrangement of pixels that describes an object, if that same object is mirrored throughout the picture, our models may find it difficult to identify it. The proposed deep learning model is being fed with more information to learn by generating multiple versions of our images in different orientations without going through the time-consuming process of gathering and labeling more training data.

3.2.3 Stage 3: Post-processing

To enhance the 2D slice from a detected nodule image, the pseudo coloring technique ([Wang et al. 2022](#)) is applied. The grayscale image is converted to a pseudo-colored

image as illustrated in Figure 3.6. Human beings can perceive many different colors but can only distinguish between a few gray-level values (Niranjana and Ponnaivaikko 2017). So, to help the radiologists better identify lung nodules/lesions, the pseudo-coloring contrast image enhancing technique looks promising. To convert a gray scale image into the pseudo-color image, three matrices are required. This is implemented by applying three different transformation functions, i.e., for Red, Green, and Blue colors, on the grayscale image. Combining the output resulting from the transformation functions on the grayscale image provides a pseudo-colored image. The approach of pseudo coloring a grayscale image can be shown by equations 3.2, 3.3, and 3.4.

$$R(x, y) = T_R[f(x, y)] \quad (3.2)$$

$$G(x, y) = T_G[f(x, y)] \quad (3.3)$$

$$B(x, y) = T_B[f(x, y)] \quad (3.4)$$

The intensity values of the primary colors, i.e, Red, Green and Blue at coordinate (x, y) is denoted by $R(x, y)$, $G(x, y)$, and $B(x, y)$. A grayscale image is denoted by $f(x, y)$. The transformation functions for the three primary colors are represented by $T_R[\cdot]$, $T_G[\cdot]$, and $T_B[\cdot]$.

3.2.4 Stage 4: Lung Nodule Classification

The classification of the lung nodules is performed using three sets of image inputs, 2D CT scan slice (middle slice), 3D CT cube, and the post-processed pseudocolor CT image. Three image inputs are separately fed to three deep learning architectures, namely, SqueezeNet, ResNet, and NCNet. The primary objective of this stage is to classify nodule and non-nodule images and to reduce the FPs/Scan. A detailed explanation of the proposed network is provided in the subsection 3.2.4.1.

3.2.4.1 Nodule Classification: NCNet

The proposed model integrates the two best operations of Squeezenet's fire module and ResNet's skip connection outperforming both architectures' performance. ResNet consists of stacking convolutional blocks with a skip connection and an identity function. SqueezeNet consists of a fire module in which two operations, namely, squeeze and expand, are used. The squeeze operation consists of 1x1 convolution blocks rather than 3x3 convolution blocks in order to achieve fewer parameters with better performance. Next, the expand layer performs multi-scale learning and concatenation operations using a convolution block of kernel sizes 1x1 and 3x3. In mathematical terms, the fire module consists of three hyperparameters related to the dimension, i.e., s_{1x1} , e_{1x1} , and e_{3x3} . That is, squeeze operation having convolutional layer kernel size of 1x1 (s_{1x1}), expand layer having kernel size of 1x1 (e_{1x1}), and 3x3 (e_{3x3}). The design space of Squeezenet is defined by meta parameters that control the dimensions in a fire module of the CNN (Iandola et al. 2016). The filters in the expand layer are represented by $base_e$. The number of expand filters in the initial fire module is increased using $incr_e$ with a given frequency $freq_e$. For a fire module i , the number of filters in expand layer is given by the below equation 3.5.

$$e_i = base_e + (incr_e + 1 * \frac{1}{freq_e}) \quad (3.5)$$

The number of filters in the squeeze layer is decided by a metaparameter named Squeeze Ratio (SR). The SR is the range [0,1] which is used by all the fire modules. The number of filters in squeeze layer is given in below equation 3.6.

$$s_{i, 1x1} = SR * (e_{i, 1x1} + e_{i, 3x3}) \quad (3.6)$$

In the ResNet architecture, the residual connections are given to convolution layers. For the proposed NCNet, the residual connections are provided to the fire module. This operation is carried out using identity blocks given in equation 3.7 mimicking ResNet architecture (He et al. 2016). In the equation, x denotes the input vector, and y denotes the output vector. The residual mapping that takes place in the model is represented by

$F(x, W_i)$, where W_i represents weights of i^{th} layer.

$$y = F(x, W_i) + x \quad (3.7)$$

In the above equation, function $F(x, W_i)$ represents a fire module. No extra parameters or computational complexity is introduced in equation [3.7](#). Short-cut connection and element-wise addition are used for performing $F + x$ operation. One important point to be taken care of in performing the above operation is the dimensionality of input x , and F must be the same. In case this condition is not satisfied, linear projection W_s must be performed to match the dimensions. This is illustrated by using the below equation [3.8](#).

$$y = F(x, W_i) + W_s x \quad (3.8)$$

The working of NCNet is demonstrated in Figure [3.4](#). The key contribution of the proposed work is the integration of fire modules instead of conventional convolution blocks with a skip connection. This approach is novel and has not been done as far as our knowledge is concerned. ResNet is a deep architecture and as the depth of the network increases, the generation of training parameters in the network also increases. Therefore, in order to overcome the issue of huge parameters, a fire module is introduced in the network. The proposed NCNet architecture is developed for 2D and 3D CT scan inputs. The model worked exceptionally well for both 2D and 3D CT scans.

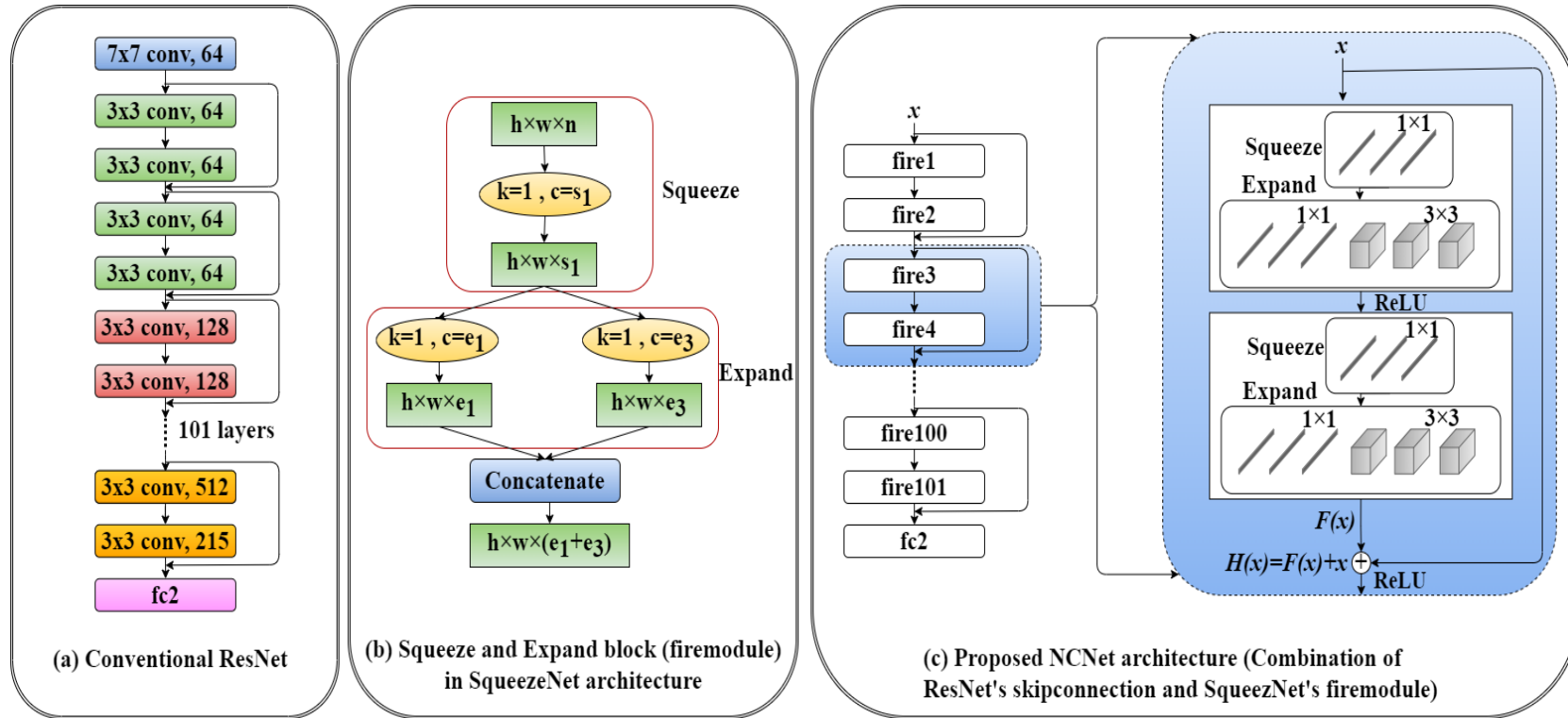


Figure 3.4: These are multiple architectures used in this work, where, (a) shows Conventional ResNet, (b) shows working of Squeeze and Expand block in SqueezeNet, and (c) shows proposed NCNet architecture which is a new combination of (a) and (b)

3.3 RESULTS AND DISCUSSIONS

The performance metrics considered for evaluating the proposed RFR V-Net are Dice Similarity Coefficient (DSC) and Intersection over Union (IoU). DSC is a statistical tool which measures the similarity between two sets of data. IoU measures the percentage of overlap between two objects. DSC and IoU are considered primary evaluation metrics for segmentation task as these metrics are popularly used for analyzing overlapping between the two segmentation results. The computation of the metrics is given in the below equations, where TP, FP, and FN represent True Positive, False Positive, and False Negative, respectively.

$$DSC = \frac{2 * TP}{(TP + FP) + (TP + FN)} \quad (3.9)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (3.10)$$

The results of nodule detection are provided in Table 3.1. The comparison of SOTA methods with the proposed approach for both performance metrics is provided in the table. The existing works that have been considered here for comparison are mainly based on U-Net, V-Net, Mask R-CNN, etc.

Table 3.1: Comparison of the proposed work with other Nodule Detection CAD systems

Ref.	DSC (in %)	Authors	IoU
Cao et al. (2020)	82.74	Aresta et al. (2019)	0.55
Roy et al. (2019)	93.00	Wu et al. (2018)	0.58
Ronneberger et al. (2015)	94.97	Messay et al. (2015)	0.74
Wang et al. (2017)	82.15	Wang et al. (2017)	0.71
Proposed RFR V-Net	95.01	Proposed RFR V-Net	0.83

Table 3.1 shows that the proposed RFR V-Net obtained the best DSC of 95.01% and an IoU of 0.83 compared to existing approaches. However, the second best performing approach was proposed by Ronneberger et al. (2015) with a DSC of 94.97%. The second-best performance for IoU was 0.74 from a technique proposed by Messay et al. (2015). The findings show that the DSC of the proposed method was just improved by 0.04% when compared to the second-best method by Ronneberger et al. (2015).

3. Lung Cancer Detection

However, the proposed model provided a much-improved result in terms of IoU with an improvement of 9% over the second-best method (Messay et al. 2015).

The results obtained for various patients with an image size of 96x96x16 are shown in Figure 3.5, along with the true segmentation mask and the segmentation mask predicted by the proposed RFR V-Net model. From the figure, the mask predicted by the RFR V-Net is almost matching to that of the original segmentation. This shows that the proposed RFR V-Net architecture performed well in segmenting the lung nodules in the CT scan images.

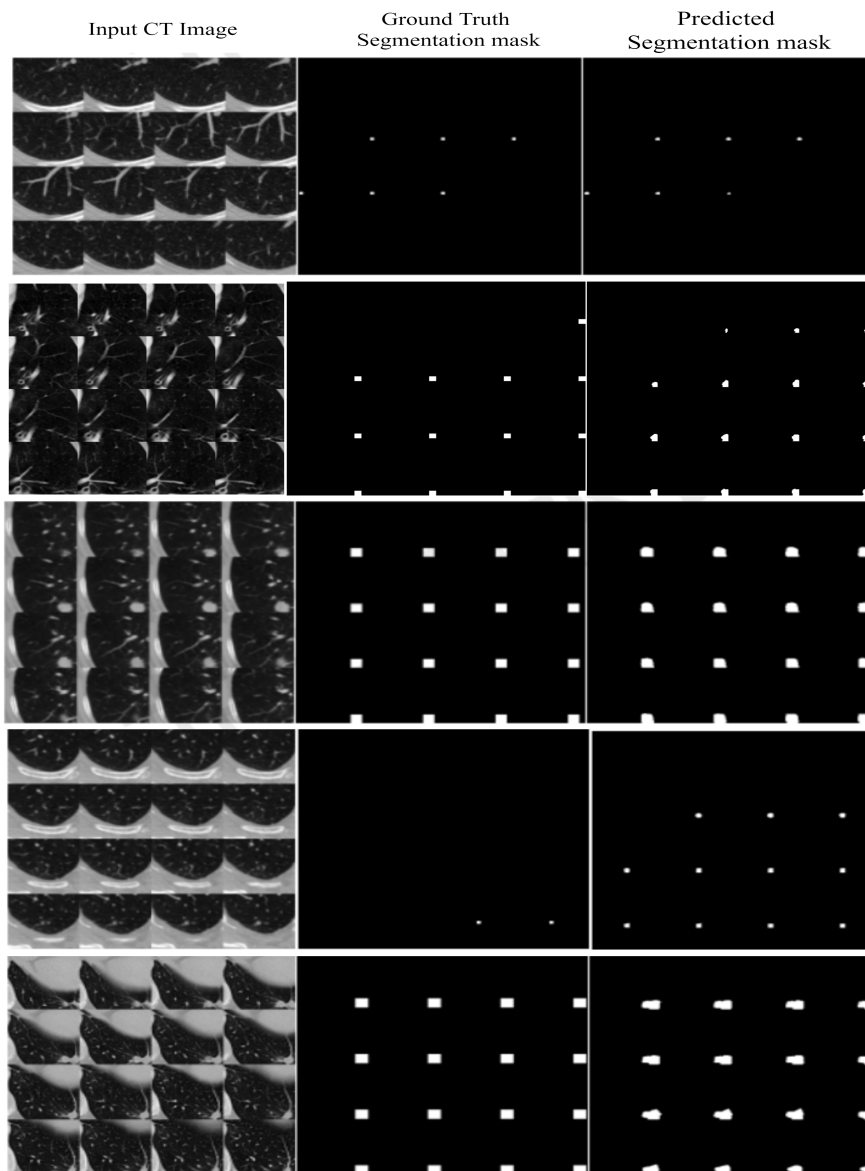


Figure 3.5: Results achieved for the proposed RFR V-Net.

An illustration of the nodule and non-nodule images before the post-processing step is shown in Figure 3.6. The number of nodules and non-nodules present in the dataset after the division is 1186 and 5,51,056, respectively. It can be noted that the data is imbalanced. Because of data skewness, the chances of this data being classified as a non-nodule are higher (Shi et al. 2020). To overcome this issue, the nodule images must be augmented. The image enhancement of CT slices are illustrated in Figure 3.6. Hue, Saturation, Value (HSV) colormap is applied to the grayscale 2D CT slice using the OpenCV tool. The range of values used for Hue is [0,179], Saturation is [0,255], and Value is [0,255]. This provides increased visibility of the CT slice where the lesions are better seen than in the normal CT slice.

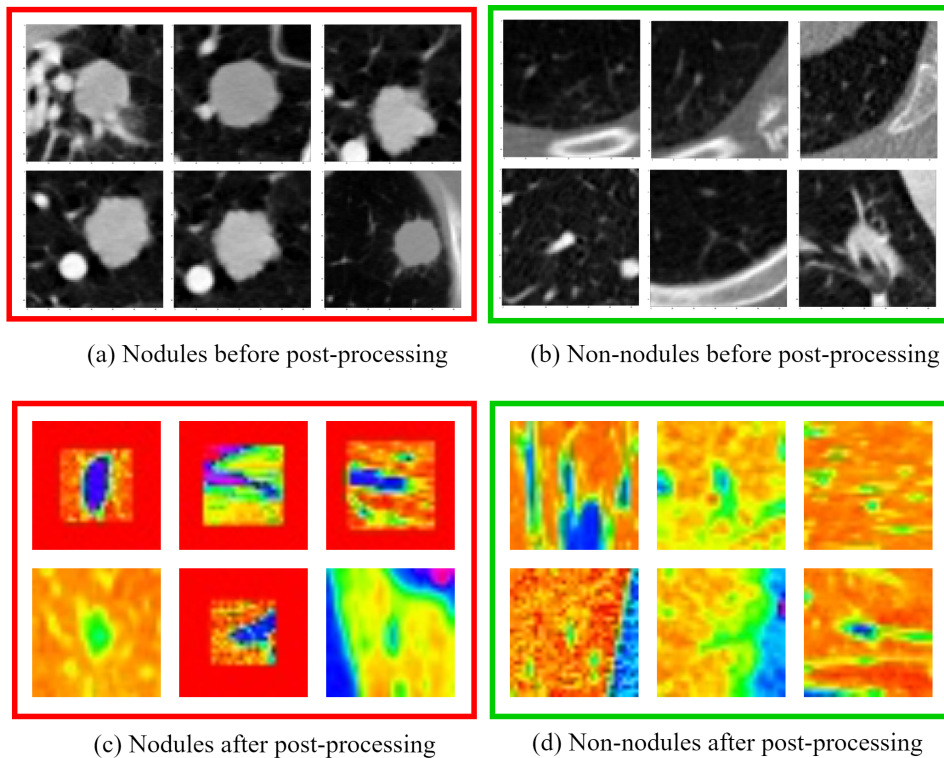


Figure 3.6: Nodule and non-nodule images obtained before post-processing ((a) and (b)) and after post-processing ((c) and (d))

The models used in this work are trained using various parameters mentioned as follows: Optimizer used is Adam, the activation function is ReLU, and the loss function used is Binary cross-entropy. The models are trained with batch sizes of 64 for 2D input images and 32 for 3D input images. The activation used in the hidden layers of all the

3. Lung Cancer Detection

models is ReLU. The reasons for using ReLU in the models are, firstly, ReLU is a non-linear activation function. Secondly, it returns the output value from $\max(0, x)$ where x is the input. Therefore, it removes the negative values and provides a much simpler computation. Thirdly, the ReLU activation function helps to mitigate the vanishing gradient issue. As deep architectures are used in this work, ReLU activation is more suitable. The final decision is made using the sigmoid activation function.

The proposed NCNet 3D was trained and tested for 20, 32, 44, 56, and 110 layers. However, the model provided the best recognition accuracy of 98.21% and the least FPR of 0.0166 for 56 layers. As the literature suggests, more layers give the model more “capacity”. Hence, the proposed model does not provide the best accuracy for 20, 32, or 44 layers. The reason for the decrease in performance for a 110-layered network is due to the overfitting of the model because of the limited training data and complex network architecture. Figure 3.7 demonstrates the training and validation accuracy and loss obtained for the NCNet when trained using 56 layers. The model’s accuracy in any ideal deep-learning architecture should increase with epochs. An observation can be made from the figure that both the training and validation accuracy of the NCNet is rising. This evidently shows that the model is learning discriminative information of the two classes exceptionally. Also, when the number of epochs increases, the model’s loss should decrease and be observed here. Hence, it can be said that the FPR for the model is very minimal.

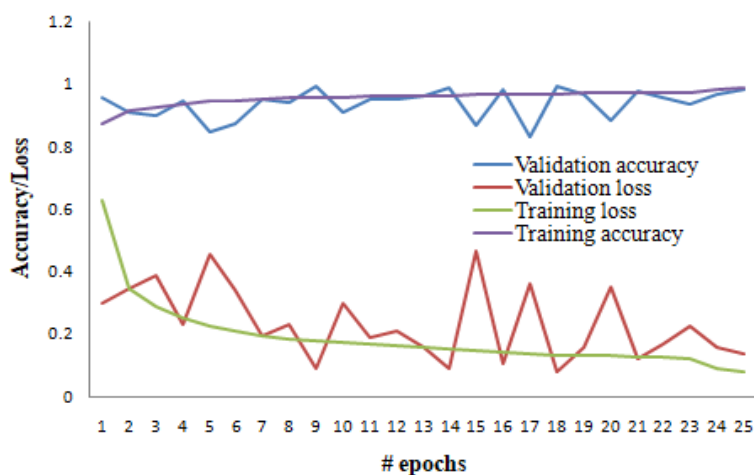


Figure 3.7: Graph representing the accuracy and loss achieved for the proposed NCNet3D model

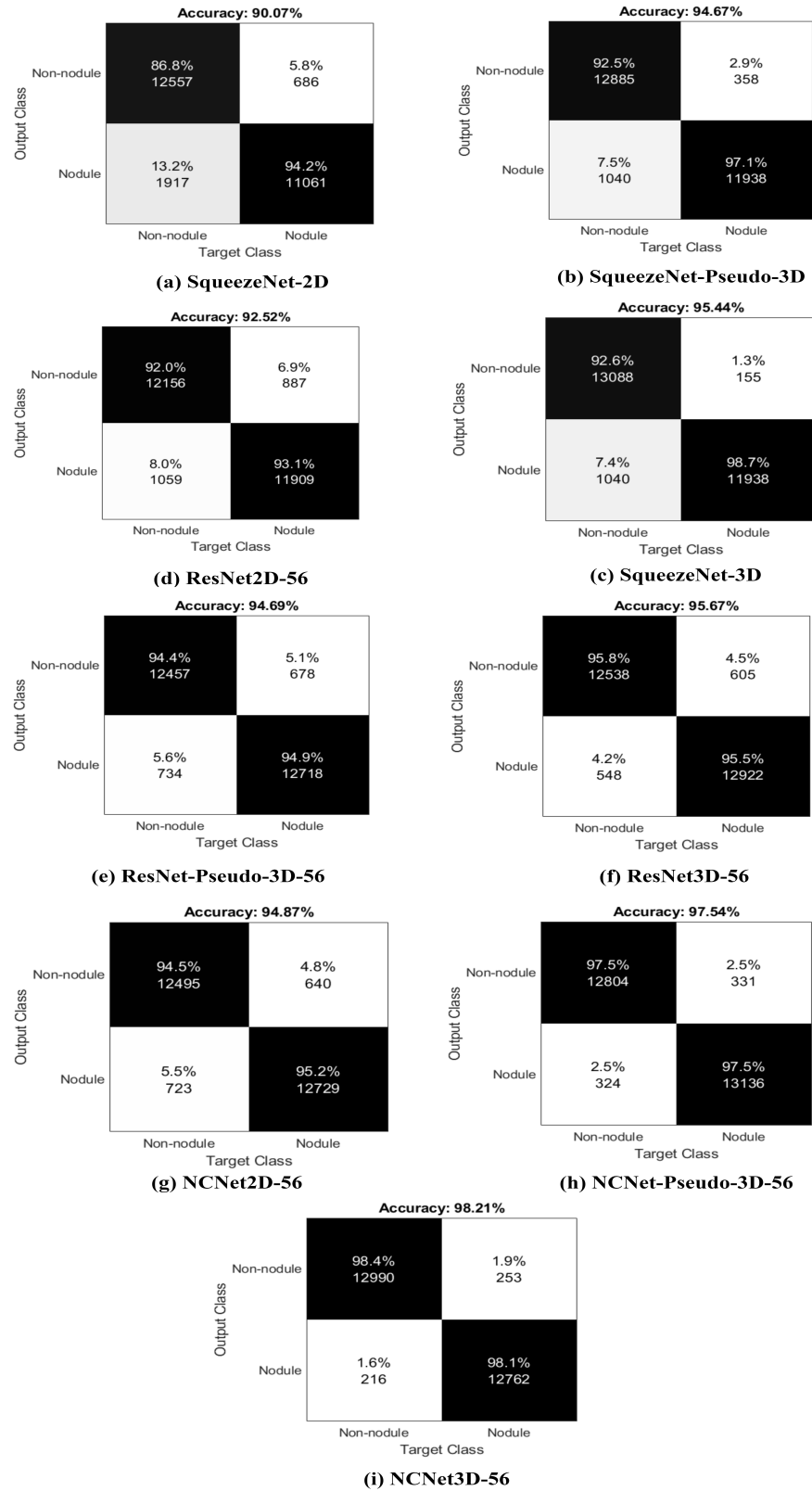


Figure 3.8: Confusion matrices obtained for best recognition accuracies

3. Lung Cancer Detection

Confusion matrices illustrate both the recognition accuracy and the number of images detected as nodules and non-nodules. Figure 3.8 demonstrates the confusion matrices obtained for 2D, pseudo-3D, and 3D models of SqueezeNet, ResNet, and proposed NCNet, respectively. ResNet and proposed NCNet are evaluated by altering the number of layers in the network architectures. The best results were obtained using 56 layers. The number of samples getting detected correctly is comparatively better for 3D models than for 2D models.

The comparison of best-performing models in terms of FPs/Scan is illustrated using the FROC curve in Figure 3.9. The average FPs/Scan value of the NCNet-Pseudo-3D and NCNet-3D models for 56 layers is provided in the FROC curves. The models resulted in 2.4 and 2.3 FPs/scan for NCNet-Pseudo-3D and NCNet-3D, respectively.

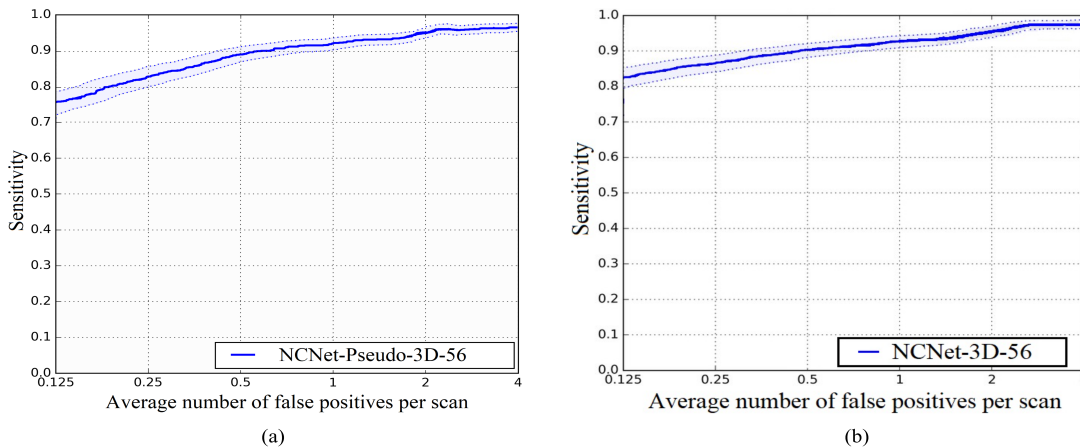


Figure 3.9: FROC curve comparing the performance of (a) NCNet-Pseudo-3D-56 at 2.4 FPs/scan, and (b) NCNet-3D-56 at 2.3 FPs/scan

The performance measures considered for the assessment are accuracy (ACC), specificity (SPE), precision (PRE), sensitivity (SEN), F-score, FPR, and FPs/Scan. Table 3.2 demonstrates the evaluation of the various 2D, pseudo-3D, and 3D models used in this work. The results have been listed for 33 models for 2D, pseudo-3D, and 3D with different layer variations in ResNet and NCNet architectures. From the table, it can be noted that 3D models outperformed the 2D models. This is because the capturing of nodule information in 3D is better than in 2D slices of the CT scans. Volumetric information in 3D data contributes to the improvement of the performance of the system.

Table 3.2: Comparison of various models on multiple layer configurations with proposed work for 2D, pseudo-3D, and 3D data

Models	Dimension	ACC (in %)	SPE (in %)	PR (in %)	SEN (in %)	F-score	FPR	FPs/ Scan
SqueezeNet	2D	90.07	85.22	86.75	94.81	0.9061	0.1477	3.2
SqueezeNet	Pseudo-3D	94.66	91.98	92.53	97.29	0.9485	0.0801	2.6
SqueezeNet	3D	95.44	91.98	92.63	98.82	0.9563	0.0801	2.5
ResNet20	2D	89.56	89.99	90.08	89.14	0.8962	0.1001	3.8
ResNet32	2D	90.54	90.18	90.42	90.88	0.9066	0.0982	3.7
ResNet44	2D	91.61	91.71	91.74	91.51	0.9163	0.0828	3.5
ResNet56	2D	92.52	91.83	91.98	93.19	0.9259	0.0817	3.4
ResNet110	2D	88.93	90.27	89.7	87.54	0.8862	0.0973	4
ResNet20	Pseudo-3D	90.38	90.19	92.16	91.48	0.9012	0.0898	3.5
ResNet32	Pseudo-3D	91.12	90.93	91.67	92.36	0.9152	0.0894	3.3
ResNet44	Pseudo-3D	93.29	90.94	90.70	92.99	0.9258	0.0863	3.2
ResNet56	Pseudo-3D	94.69	94.54	94.43	94.83	0.9464	0.0546	3
ResNet110	Pseudo-3D	87.51	84.01	88.64	86.32	0.8815	0.0815	4.1
ResNet20	3D	91.38	91.21	90.98	91.5	0.9128	0.0879	3.3
ResNet32	3D	92.06	91.42	91.28	92.73	0.92	0.0857	3.2
ResNet44	3D	94.59	94.89	94.7	94.28	0.9449	0.051	2.8
ResNet56	3D	95.56	95.93	95.81	95.39	0.956	0.0407	2.5
ResNet110	3D	89.89	88.15	88.3	91.67	0.8996	0.1184	3.7
NCNet20	2D	91.28	89.05	89.29	93.57	0.9138	0.1094	2.8
NCNet32	2D	92.34	90.27	90.4	94.54	0.9243	0.098	2.6
NCNet44	2D	93.11	91.81	91.84	94.44	0.9313	0.0818	2.5
NCNet56	2D	94.87	94.62	94.53	95.12	0.9483	0.0537	3.1
NCNet110	2D	90.9	90.9	90.71	90.92	0.9082	0.0901	3.6
NCNet20	Pseudo-3D	92.36	91.42	91.89	91.89	0.9651	0.0812	2.5
NCNet32	Pseudo-3D	93.97	92.45	92.85	94.01	0.9467	0.0310	2.9
NCNet44	Pseudo-3D	95.14	94.42	95.92	95.68	0.9543	0.0362	2.7
NCNet56	Pseudo-3D	97.53	97.59	97.53	97.48	0.9751	0.0201	2.4
NCNet110	Pseudo-3D	89.60	87.46	88.91	86.59	0.9278	0.0214	3.4
NCNet20	3D	93.59	92.99	92.92	94.2	0.9356	0.07	3.1
NCNet32	3D	94.49	94.39	94.27	94.6	0.9444	0.0561	3
NCNet44	3D	96.63	96.41	96.34	96.87	0.9661	0.0359	2.7
NCNet56	3D	98.21	98.33	98.36	98.38	0.9823	0.0166	2.3
NCNet110	3D	91.45	91.32	91.14	91.59	0.9137	0.0868	3.3

However, the results obtained for pseudo-color 3D images are approximately near to 3D models. Even though the input size is drastically reduced from 32x32x16 to 32x32x3 in pseudo-color images, the performance achieved from the image enhancement technique proved to work efficiently. The limitation of the proposed

3. Lung Cancer Detection

NCNet is the decrease in performance for a 110-layered network. This is due to the overfitting of the model because of the limited training data and complex network architecture. The issue can be resolved using huge image nodule data to train the model.

Table 3.3: Comparison of previous works with the proposed NCNet

Ref.	Input dimension	Methods used	ACC (in %)	SEN (in %)	SPE (in %)	FPS/Scan
Wang et al. (2019)	2D	Deep Fully Connected Network (DFCNet)	*N/A	90	N/A	15
Xu et al. (2020)	2D	DeepLN	N/A	91.17	N/A	N/A
Kumar et al. (2015)	2D	Deep features from autoencoder	75.01	83.35	N/A	N/A
Hua et al. (2015)	2D	Deep belief Network, CNN	N/A	73.4	82.2	N/A
Proposed NCNet	2D	SqueezeNet-ResNet combination	94.87	95.12	94.62	3.1
Hussein et al. (2017)	3D	Transfer learning and multi-task learning	91.26	N/A	N/A	N/A
Han et al. (2015)	3D	CMixNet, Faster R-CNN	N/A	94	91	N/A
Naqi et al. (2020)	3D	Stacked autoencoders	96.9	95.6	97	2.8
Liao et al. (2019)	3D	3D deep neural network	81.42	N/A	N/A	N/A
Gong et al. (2018b)	3D	Random Forest, J48 Decision Tree, SVM, Logistic	N/A	84.62	N/A	2.8
Proposed NCNet	Pseudo-3D	SqueezeNet-ResNet combination	97.53	97.48	97.59	2.4
Proposed NCNet	3D	SqueezeNet-ResNet combination	98.21	98.38	98.33	2.3

*N/A-Not Available

The performance comparison of the proposed NCNet with other previous works for lung nodule detection on the LUNA16 challenge and the parent dataset LIDC-IDRI is provided in Table 3.3. Deep learning-based architectures are the foundation for all of the techniques that are being compared. Some of the networks used are deep fully connected networks (Wang et al. 2019), deep belief networks, convolutional neural

networks (Hua et al. 2015), Faster R-CNN (Han et al. 2015), etc. Some works have used transfer learning, where pre-trained deep learning models are used for classifying lung cancer nodules (Hua et al. 2015; Hussein et al. 2017). The results achieved from the proposed model outperformed previously proposed CAD systems. The findings show that the accuracy of the proposed method was improved by 1.31% when compared to the second best method by Naqi et al. (2020). The proposed approach also achieved very less FPs/Scan of 2.3, whereas the second best FPs/Scan of 2.8 were obtained by Gong et al. (2018b); Naqi et al. (2020). The main aim of this work was to develop an efficient CAD system that performs lung cancer nodule detection with an higher accuracy rate and less FPs/Scan. The aim is achieved in this work by surpassing the performances for both objectives in comparison to the existing CAD systems.

3.4 SUMMARY

Lung nodule detection using computer-aided algorithmic techniques is challenging, and the research has been carried out for three decades by now. Deep learning algorithms are a recent trend that has displayed drastic performance improvement in medical imaging tasks.

This chapter mainly focused on the proposed novel segmentation algorithm “RFR V-Net” to perform lung nodule detection in raw CT images. The method uses regularized RFs in V-Net architecture encoder and decoder blocks. The introduction of RFs improved the performance of the segmentation method. Another novel architecture named NCNet is proposed to perform lung nodule classification into nodules and non-nodules. The model proposed is a hybrid of SqueezeNet, which is considered as one of the light-weight CNN and results in fewer training parameters. The Global Average Pooling (GAP) used in the SqueezeNet architecture has no parameters, a regularizing effect, and enables various input sizes for the networks. This layer retains an extensive amount of useful localization information instead of fully connected layers, which greatly reduces the model size. Secondly, ResNet, a popular deep learning model for increasing the number of layers in the network without vanishing gradient problem. A pseudo-coloring image enhancement technique

3. Lung Cancer Detection

is adopted in this work. Enhancing the 2D slice input in a CT scan by adding intensity values to the slice and converting a grayscale 2D image to an RGB image proved to perform almost as good as a 3D model.

CHAPTER 4

LUNG CANCER NODULE SEGMENTATION

4.1 INTRODUCTION

Lung cancer is one of the cancers with the highest death rates worldwide. Even though there are techniques to identify lung cancer nodules, it takes enormous effort from expert radiologists. Therefore, it is very crucial to automate the process of identifying nodules from CT scans. This will provide clinicians with a second opinion, making lung cancer nodule diagnosis much easier. CAD systems are used to automate the process of locating lung cancer nodules. Several CAD systems are proposed to perform lung cancer detection and classification. The detection stage is named CADe systems, where the possible nodule candidates are identified. The classification stage is named CADx systems, where the nodule identified by CADe systems are classified into cancerous (malignant) and non-cancerous (benign) nodules. AI-based methods are most commonly used in recent trends to perform lung cancer detection and classification. Identifying the nodule region from a CT scan is quite a tedious task, as in certain cases, the tissue regions present in the scans are considered a nodule. Therefore, a careful examination is required to segment the nodule region from a CT scan input image. This can be achieved by deep learning segmentation approaches. This chapter discusses the two proposed systems for addressing lung nodule segmentation tasks. Followed by this, the nodules are further classified into cancerous and non-cancerous nodules.

The contributions of the two proposed systems are as follows:

1. Contributions of the proposed lung cancer nodule segmentation system-I
 - A novel segmentation method named Light-weight RefineNet to perform lung nodule Segmentation from CT scans.
 - A new lightweight classification model named PSO-based CondenseNet to classify lung nodules.
2. Contribution of the proposed lung cancer nodule segmentation system-II
 - A novel segmentation method named Elagha initialization based Fuzzy C-Means clustering (EFCM) to perform segmentation of nodule regions from a given CT scan.

The chapter is organized as follows: Section [4.2](#) is dedicated for proposed lung cancer nodule segmentation system-I and its results and discussion, Section [4.3](#) discusses proposed lung cancer nodule segmentation system-II in detail along with its results and discussion, and Section [4.4](#) summarizes the proposed method and its significance.

4.2 PROPOSED LUNG CANCER NODULE SEGMENTATION SYSTEM-I

The block diagram of the proposed lung nodule segmentation and classification CAD system is shown in Figure [4.1](#). The input considered for the CAD system is 3D CT scans. The raw CT scans need to be pre-processed as the scans contain background information that is unnecessary for the segmentation or classification task. The pre-processed CT scans are fed to the Deep Convolutional Generative Adversarial Network (DCGAN) network, which performs augmentation tasks. The augmentation step is essential in developing the CAD system as the number of non-nodules is large compared to the nodules in the dataset. Also, when the identified nodules are further classified into benign and malignant nodules, the images are very few in number. Therefore, DCGAN is used for increasing the number of images in both cases. The ROI is segmented from the CT scans using the novel lightweight RefineNet model.

The model uses fewer parameters to perform the segmentation of candidate nodules from the pre-processed CT scans. The detected candidate nodules are further classified into benign and malignant nodules, which is performed using a novel lightweight PSO-based CondenseNet deep learning model. A detailed description of each method is provided in the below subsections.

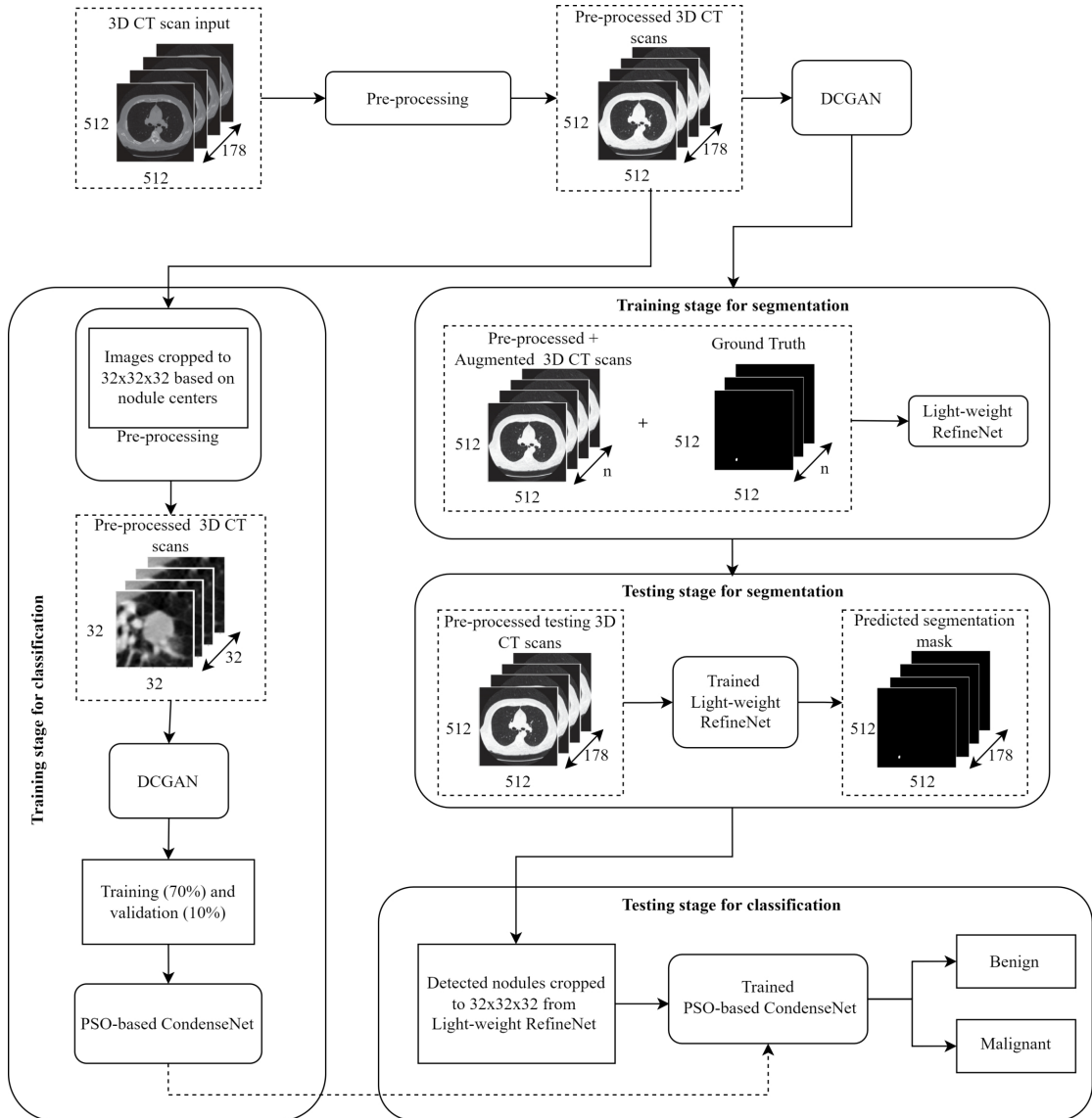


Figure 4.1: Block diagram of the proposed Lightweight CAD system

4.2.1 Pre-processing

Pre-processing step is considered as one of the crucial steps in the lung nodule detection and classification task. The primary step of pre-processing is organizing the images

4. Lung Cancer Nodule Segmentation

provided in the LUNA16 dataset. The CT scans are provided in the digital imaging and communications in medicine (DICOM) file format, a standard data format used to store medical images. These scans are read from python scripts, and two file types are created in this process, namely, .raw and .mhd. The image content is stored in the .raw file type, whereas the image's meta-data, like diagnostics and annotations, is stored in the .mhd file. The CT scans are three-dimension and have approximately 200 slices where each squared slice is 512 pixels (Potghan et al. 2018).

The slices in the dataset are of larger size. Therefore, making it difficult in terms of computation. To overcome this issue, the size of the CT scan is reduced by applying a technique named Hounsfield unit (HU). HU is used as a relative quantitative measurement of radio density by radiologists to analyze CT scans. This popular technique is used to identify various chemical compositions in the CT scans, such as bones, water, air, etc. On the basis of this technique, each image's pixel is converted to HU values. The CT scan pictures are trimmed between -1000 and 400 HU to reduce image resampling variations as well as bone densities. The images are normalized before being fed into the neural network as input. This is done by applying a standard mean voxel value of 0 and a variance of 1. The images are generated based on the annotations provided by the LUNA16 challenge organizers. Two folders are generated, namely, nodules and non-nodules. The number of positive and negative nodules generated are 1186 and 5,47,346, respectively. The images are cropped to 32x32 dimensions to be computationally less expensive.

4.2.2 Augmentation

After pre-processing the nodule and non-nodule images, it can be observed that there is a huge data imbalance in the number of images of the nodule and non-nodule classes. Training the neural network using these images can cause the network to overfit and results in performance reduction. Therefore, Deep Convolutional Generative Adversarial Network (DCGAN) is used to increase the number of images in the nodule classes. DC-GAN is one of the popular techniques used for augmenting images in low-resource data (Kulkarni and Panditrao 2014). GAN architecture mainly

works on the two networks, namely, discriminator (D) and Generator (G), as shown in Figure 4.2. Both networks are trained simultaneously. The role of D is to differentiate between real and fake samples, whereas the role of G is to generate synthesized images as per input samples. The input samples are mapped using a uniform distribution.

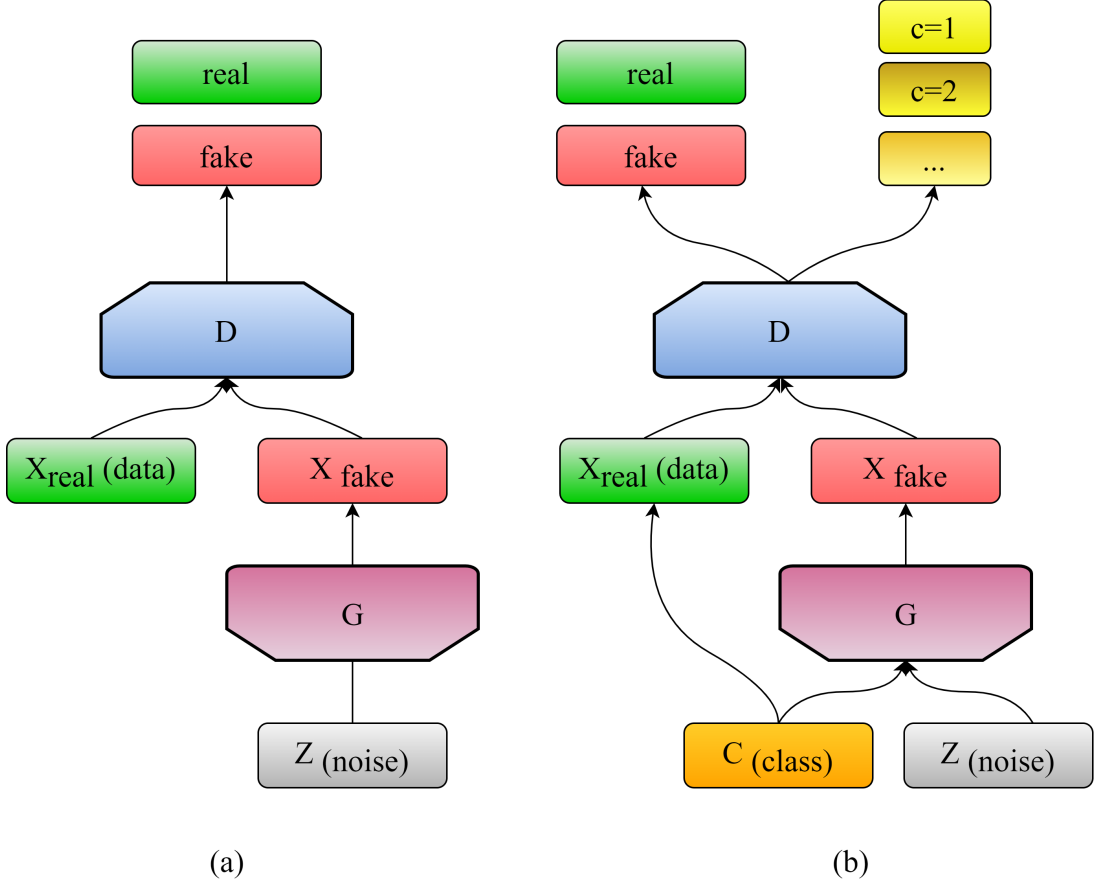


Figure 4.2: Schematic architecture of Generative Adversarial Network

The networks are trained using an optimization function along with the two-player minimax game as shown in equation 4.1:

$$\min_G \max_D \epsilon_{x \sim p_{data}} \log D(x) + \epsilon_{z \sim p_z} [\log (1 - D(G(z)))] \quad (4.1)$$

where, the generator G gets input samples $z(1), \dots, z(m)$, from an uniform distribution p_z . These images are mapped from G_z to image space distribution p_g . The network is trained for the discriminator to maximize $D(x)$ for images with $x \sim p_{data}$ and to minimize $D(x)$ for images with $x \approx p_{data}$. The G network is responsible for

generating images $G(z)$ in order to fool the D network at the time of training network in such a way that $D(G(z)) \sim p_{data}$. Hence, the G network is maximized for $D(G(z))$ and, in the meantime, minimized for $1 - D(G(z))$. During the training of the DC-GAN network, the generator is used to enhance the synthesizing of the real images. In contrast, the discriminator network enhances the ability to differentiate between real and synthesized images.

4.2.3 Light-weight RefineNet for Nodule Segmentation

This work introduces a novel lightweight RefineNet model to segment nodules from CT scans. RefineNet architecture is a deep learning model proposed to perform high-resolution semantic segmentation using a multi-path refinement network by exploiting features at multiple levels of abstraction (Jakimovski and Davcev 2019). The network extracts features by recursively refining low-resolution features with high-resolution features. Figure 4.3 demonstrates the proposed lightweight RefineNet architecture for the nodule segmentation task. The input image fed to the network is a preprocessed CT scan. The output resulting from the network is the predicted mask of the nodule.

The proposed lightweight RefineNet consists of four main components. The first component is the main block which consists of 3 sub-blocks. The first component takes a multi-path input which is the input image with different input dimension sizes. This component consists of three sub-blocks: Residual Convolution Unit (RCU), Multi-Resolution Fusion, and Chained Residual Pooling. The first block RCU is built on the Residual Network (ResNet) architecture backbone, except there is no batch normalization in the RCU sub-block. This block is a type of adaptive convolution.

In the RefineNet model, the RCU block consists of 3x3 convolution blocks with the ReLU activation function. In the proposed lightweight RefineNet model, the conventional convolutional blocks are replaced with depthwise separable convolutions to reduce the network layers' number of additions and multiplications. In the second sub-block, element-wise summation of multi-resolution features is performed. Hence, the name is multi-resolution fusion. This block generates high-resolution feature maps by fusing the inputs from all the paths. Initially, depthwise convolutions are applied for

4.2. Proposed Lung Cancer Nodule Segmentation System-I

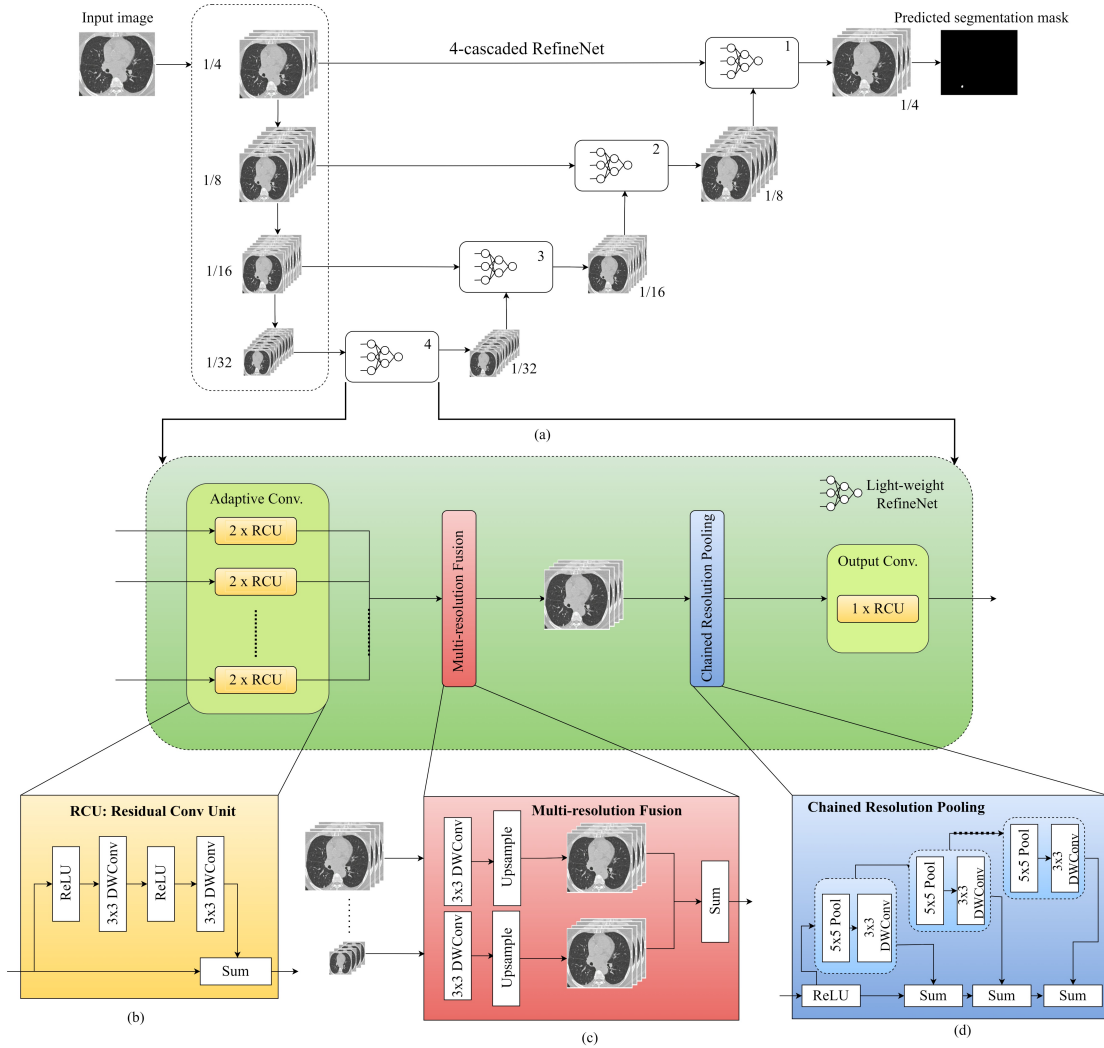


Figure 4.3: Architecture of proposed Lightweight RefineNet for Lung Nodule Segmentation

the inputs to achieve similar feature map sizes and then upsampled to obtain larger feature maps. Again, all the input feature maps are summed, and this layer helps in rescaling the input features for multi-path input. The last sub-block is the chained residual pooling. Here, the background information of the large image region is captured. The input is pooled with different window sizes and then fused using learnable weights to perform this. Here, a chain is formed with multiple pooling blocks where each block consists of one pooling layer followed by a depthwise separable convolution layer. The purpose of forming a chain of pooling layers is that the input of one pooling layer is fed as an input to the next pooling layer causing it to access the features from a large region without using a large pooling window. This

work uses a 4-cascaded Lightweight RefineNet architecture to segment lung nodules from a CT scan image.

4.2.4 PSO-based CondenseNet

In this work, the candidate nodules identified by the Lightweight RefineNet model are further classified into benign and malignant nodules. This classification is performed using a curated version of SOTA lightweight deep learning architecture CondenseNet, which is an improved version of DenseNet (He et al. 2016). The working principle of CondenseNet happens in two stages training, condensing, and optimization. Condensing stage can be one or more. In general terms, this network combines the functionalities of two architectures, ResNext and ShuffleNetv1. From ResNext architecture, group convolutions are used to improve the performance, whereas, in ShuffleNet, the group convolutions' channels are shuffled to improve the performance. In CondenseNet architecture, the input features are grouped together for grouped convolution in the training stage instead of randomly shuffling the channels. The input feature is of dimension $O \times C \times W \times H$, where O represents the number of output channels, C represents the number of input channels, W represents the width, and H represents height. When this input feature is fed to a DenseNet architecture, the 4D tensor is reduced to an $O \times C$ dimension to a matrix D by 1×1 convolutions.

In the training process of CondenseNet architecture, the feature subsets of low importance are screened for each group. The splitting of features into equal-sized groups is performed before the training begins. Therefore for each group, a feature map of O/G size is obtained. In the condensation stage, the importance of y^{th} feature map for the group g is decided based on the averaged absolute value of weights in between them, and it is done across all the outputs in that particular group as given in equation 4.2.

$$\sum_{x=1}^{O/G} |D_{x,y}^g| \quad (4.2)$$

In order to introduce group-level sparsity in the architecture, a popular regularisation technique Group Lasso (Meier et al. 2008) is used. In the condensing stage, the group-

lasso regulariser pushes all the column elements of D^g to zero. This is done because the value in the square root is dominated by the elements with a larger value in the column. This is given in equation [4.3](#).

$$\sum_{g=1}^G \sum_{y=1}^C \sqrt{\sum_{g=1}^{O/G} D_{x,y}^g}^2 \quad (4.3)$$

The number of condensing stages in the architecture is decided using a condensing factor C . If $C=3$, then the number of stages will be $C - 1 = 2$. The weights in the filters are pruned after each condensing stage by using a mask on the filter. By the end of the training, $1/C$ of the weights are left after pruning in each filter group. In the optimization stage, only the strong ones are picked up out of weak and strong feature maps, and the weaker feature maps are discarded. An index layer is introduced in the testing phase to select the best feature and rearrange the groups. Here, the pruned weights are discarded, and the sparsified model is converted into a regular connected model.

In the proposed PSO-based CondenseNet model, the CondenseNet model is curated to improve performance by tuning the model's hyperparameters as given in [Figure 4.4](#). A hyperparameter in machine learning is a parameter whose value is used to regulate the learning process. The choice of hyperparameters can be challenging and depends on factors such as the size of the dataset, hardware's computational power, network size, etc. Therefore, to achieve the best performing model, swarm intelligence algorithm particle swarm optimization (PSO) is used to optimize five hyperparameters in the CondenseNet model, namely, number of layers, loss value, learning rate in Adam optimizer, batch size, and number of epochs.

The optimized value for each hyperparameter is obtained by passing the initial framework parameters to the PSO algorithm. The fitness value of each particle (hyperparameter) is calculated, and the global best and local best values are obtained. These values are further utilized to update the velocity and position of the particle until the termination criteria are obtained. The termination criteria in this network are to achieve hyperparameters with a minimum value. For example, the number of layers in

4. Lung Cancer Nodule Segmentation

the network must be less, and the network’s performance must also be best. This process is iterated till the optimized value is obtained for all the hyperparameters.

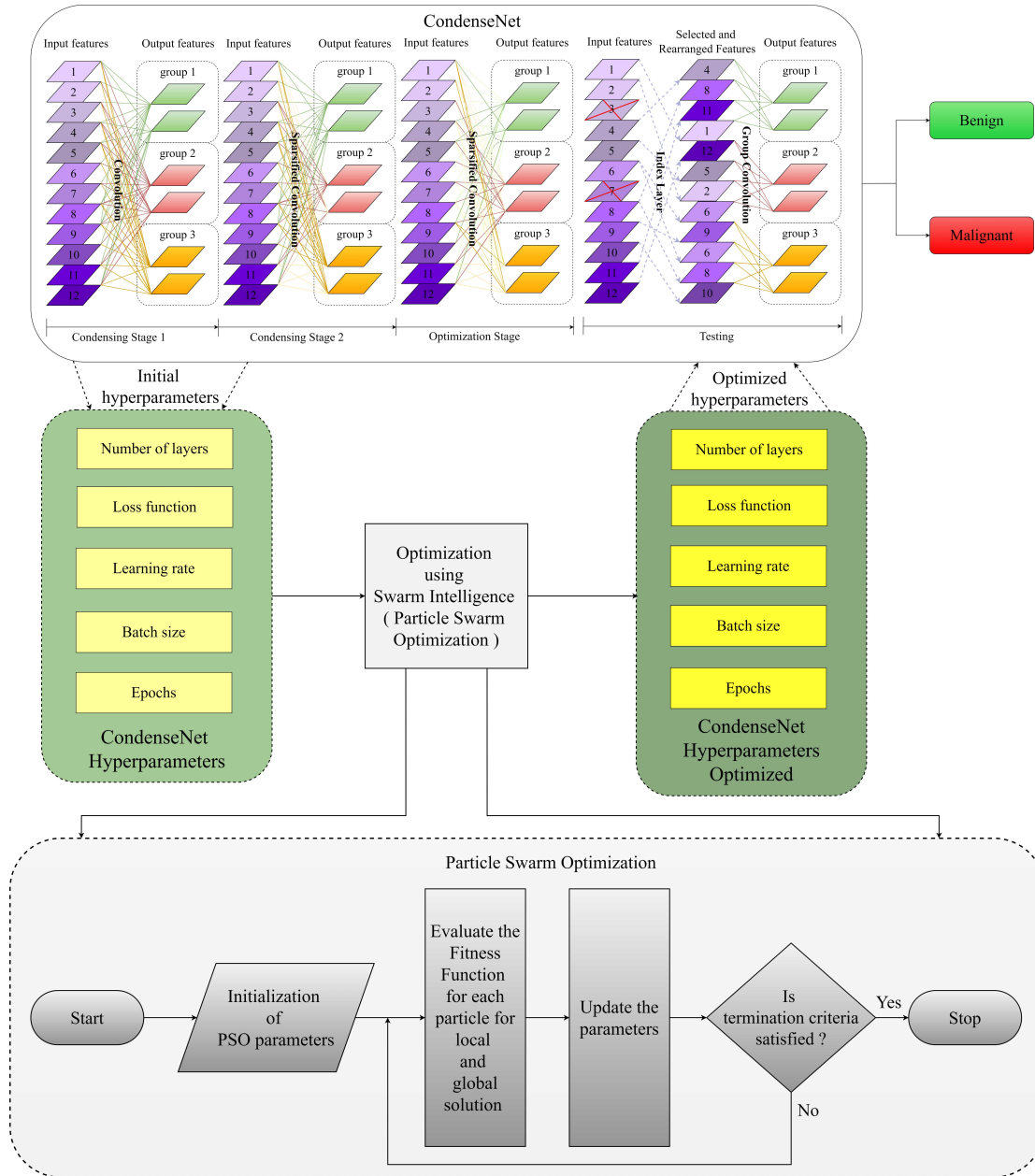


Figure 4.4: Architecture of proposed PSO-based CondenseNet for Lung Nodule Classification

4.2.5 Materials and Methods: Neural Network Configurations

This work uses three different neural networks for three tasks: DCGAN for data augmentation, lightweight RefineNet for nodule segmentation, and PSO-based

CondenseNet for nodule classification. Initially, the parameters for training the neural networks are chosen empirically. After a cross-validation step, each network is further tuned by changing the parameter values until the optimized value is achieved. The parameters that yielded the best results are listed in the subsections below.

1. **DCGAN:** The network consists of a generator model with layers like Dense, batch normalization, upsampling, convolution, and activation. The discriminator model consists of layers similar to the generator model, excluding the upsampling layer. The loss function used is binary cross-entropy, optimizer used is stochastic gradient descent (SGD) with a learning rate set to 0.005. The number of epochs is set to 100 with an early stopping criterion if the loss of the model does not change after 10 epochs. The batch size is set to 128.
2. **Lightweight RefineNet:** The network consists of layers such as depthwise separable convolution, batch normalization, and max pooling. In the multi-resolution block, the upsampling layer is used. The activation function used in the last layer of the model is softmax. The model is trained using binary cross-entropy loss function and Adam optimizer with a learning rate set to 0.001. The number of epochs set is 100 with an early stopping criterion if the loss of the model does not change after 10 epochs.
3. **PSO-based CondenseNet:** The network consists of layers such as convolutions, learned group convolutions, dense, average pooling, and batch normalization. The activation function used in the intermediate layers is rectified linear unit (ReLU). The number of epochs set is 50 with a batch size of 256. The condense factor set is 4, and the group, Lasso Lambda value, is 0. The hyperparameter tuning is performed using PSO, consisting of input parameters such as swarm size set to 18, number of iterations is set to 100, inertia weight is set to 0.7, and cognitive and social parameters are set to 2. The PSO hyperparameters are chosen based on careful analysis and experimentation.

4.2.6 Evaluation Metrics

The evaluation of the proposed system is performed using different performance metrics for segmentation and classification tasks. The segmentation of nodules is measured using metrics such as DSC, IoU, SEN, and Positive Prediction Value (PPV). DSC and IoU are considered primary evaluation metrics, and these metrics are also popularly used for analyzing overlapping between the two segmentation results. The system's robustness is measured using SEN and PPV, which are considered auxiliary evaluation metrics. The computation of the metrics is given in the below equations, where TP, TN, FP, and FN represent True Positive, True Negative, False Positive, and False Negative, respectively.

$$DSC = \frac{2 * TP}{(TP + FP) + (TP + FN)} \quad (4.4)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (4.5)$$

$$SEN = \frac{TP}{TP + FN} \quad (4.6)$$

$$PPV = \frac{TP}{TP + FP} \quad (4.7)$$

The classification of nodules is measured using metrics such as ACC, SEN, SPEC, PR, F1-score, False Negative Rate (FNR), and FPR. F1-score is computed using PR and recall (REC), where the recall metric's equation is similar to SEN. The computation of these metrics is given in below equations:

$$ACC = \frac{TP + FP}{TP + FP + TN + FN} \quad (4.8)$$

$$SPEC = \frac{TN}{FP + TN} \quad (4.9)$$

$$PR = \frac{TP}{FP + TP} \quad (4.10)$$

$$F1 - Score = \frac{2 * PR * REC}{PR + REC} \quad (4.11)$$

$$FNR = \frac{FN}{TP + FN} \quad (4.12)$$

$$FPR = \frac{FP}{FP + TN} \quad (4.13)$$

4.2.7 Results and Discussion

4.2.7.1 Segmentation of lung nodules in the CT scan

Figure 4.5 shows the output obtained for the proposed lightweight RefineNet. The input CT scan consists of lung nodules, and to train the network ground truth mask is fed along with the CT scan. In the testing phase, the proposed method predicts the segmentation output mask. The figure shows that the proposed segmentation method has predicted segmentation masks accurately indicating the system is trained well and working correctly in identifying candidate nodule structures. However, from closer observation in CT scans 1, 2, and 5, some of the tiny nodule regions in the CT scan are not captured in the predicted segmentation mask generated from the proposed model. Therefore, there is still scope of performance improvement in micro-nodule segmentation.

4.2.7.2 Free-response Receiver Operating Characteristic (FROC) analysis of Lung Nodule Segmentation system

The FROC analysis of the proposed Lightweight RefineNet architecture is shown in Figure 4.6. The sensitivity values of the system for different FPs are shown in the figure. It can be observed from the graph that there is an increase in the sensitivity rates, and the model resulted in higher sensitivity values for 2, 4, and 8 FPs. This is a crucial metric as the lung nodule segmentation task is critical, and the number of FPs must be very few for an ideal CAD system.

4. Lung Cancer Nodule Segmentation

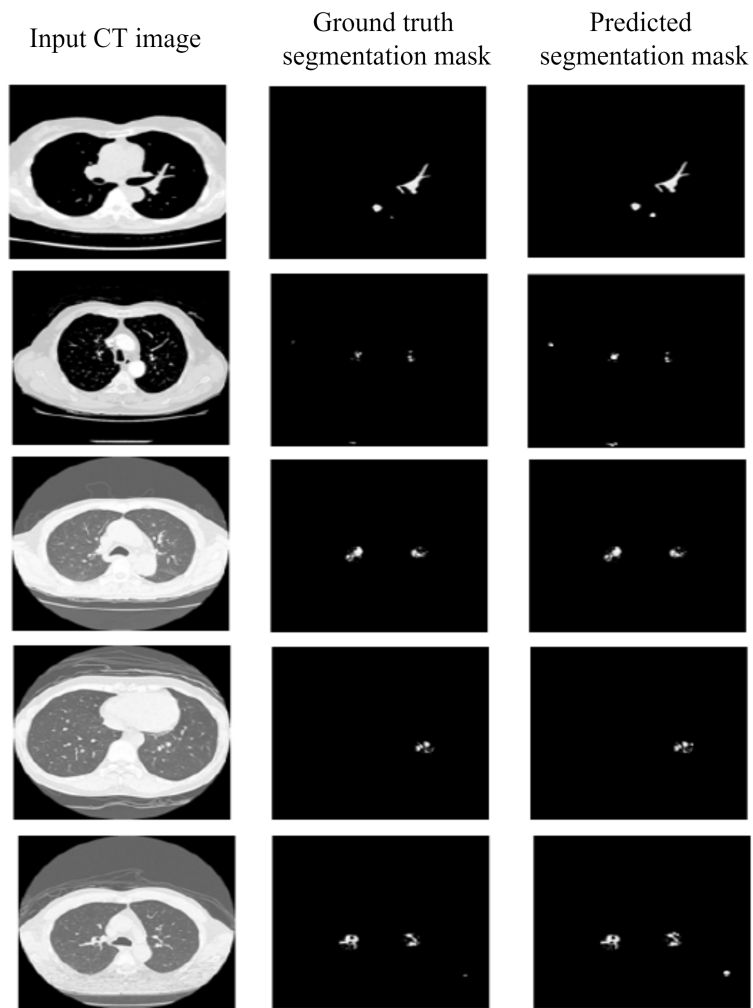


Figure 4.5: Results achieved by the proposed Lightweight RefineNet model

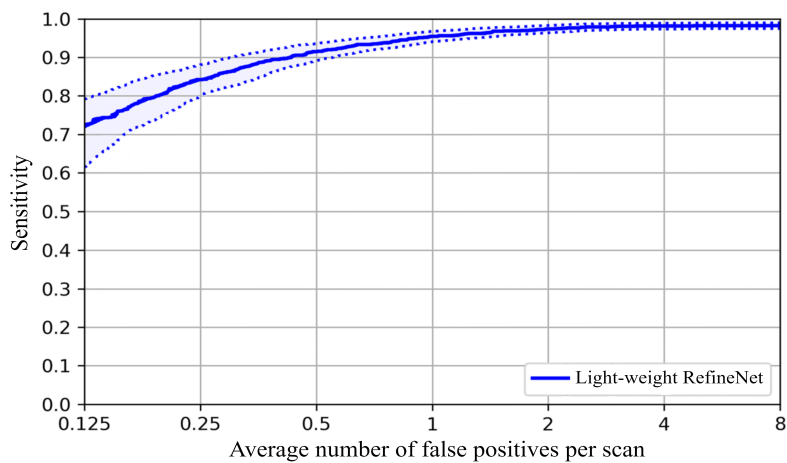


Figure 4.6: FROC curve illustrating the performance of the Lightweight RefineNet model at different FPs/scan

4.2.7.3 Performance of Lung Nodule Segmentation system

The proposed lightweight RefineNet is evaluated in comparison to state-of-the-art (SOTA) segmentation methods such as U-Net, V-Net, Fully Connected Network (FCN) U-Net, Mask RCNN, and RefineNet. The performance metrics considered for the evaluation are DSC, IoU, SEN, and PPV, gave for 0.125, 0.25, 0.5, 1, 2, 4, and 8 FPs/scan. The performance of the methods is presented in Table 4.1. The results of the proposed lightweight RefineNet achieved the best sensitivity rates for 2,4 and 8 FPs/scan. The best performance of the proposed method resulted in 98%, 93.5%, 98.5%, and 98.4% for DSC, IoU, SEN, and PPV for 8 FPs/scan, respectively. The second-best performance was resulted by RefineNet.

Table 4.1: Performance of the proposed Light-weight along with existing segmentation methods for different FPs/scan values

Metrics	Models	FPs/scan						
		0.125	0.256	0.512	1	2	4	8
DSC	U-Net	73.3	74.1	75.9	77.8	78.9	79.5	80.2
	V-Net	88.0	89.4	89.9	90.7	91.5	91.9	92.8
	FCN U-Net	86.7	87.2	88.6	89.1	90.8	91.3	92.1
	Mask RCNN	68.4	68.9	70.3	71.1	71.5	71.8	72.1
	RefineNet	91.1	92.3	93.8	94.9	96.2	96.5	96.8
	Light-weight RefineNet	93.6	94.9	95.8	96.4	97.2	97.6	98.0
IoU	U-Net	76.4	77.7	78.6	79.3	80.9	81.1	81.4
	V-Net	87.9	88.2	89.5	89.9	90.3	90.5	90.9
	FCN U-Net	85.3	86.4	87.8	88.7	89.0	89.3	89.7
	Mask RCNN	69.5	70.3	70.8	71.6	72.0	71.9	72.9
	RefineNet	86.2	87.1	89.3	90.8	91.4	91.5	91.8
	Light-weight RefineNet	88.3	89.0	90.6	91.4	92.1	92.8	93.5
SEN	U-Net	69.5	74.2	75.8	79.1	81.6	83.1	86.3
	V-Net	77.2	80.3	84.9	87.2	89.7	90.4	91.4
	FCN U-Net	78.5	80.6	83.8	84.2	87.6	89.8	91.7
	Mask RCNN	71.3	75.9	78.3	81.6	85.3	86.4	89.4
	RefineNet	71.2	73.5	86.3	88.0	93.1	94.7	95.7
	Light-weight RefineNet	73.6	82.7	95.8	97.2	97.2	97.9	98.5
PPV	U-Net	80.6	82.1	85.9	86.2	87.0	87.3	87.8
	V-Net	82.8	85.4	88.7	90.3	91.8	92.7	93.5
	FCN U-Net	83.9	85.0	86.8	89.2	90.4	91.7	93.7
	Mask RCNN	83.7	84.6	85.9	87.0	88.5	89.1	90.1
	RefineNet	87.0	89.8	91.3	92.6	95.2	95.8	96.1
	Light-weight RefineNet	89.5	91.2	93.6	95.1	96.8	97.6	98.4

4. Lung Cancer Nodule Segmentation

However, one major drawback of other segmentation methods is that they are computationally intensive in terms of time and resource consumption. Therefore, it can be said that the use of Lightweight deep learning methods is beneficial in terms of both computation and improved performance. The proposed model has achieved a DSC of 98% and an IoU of 93.5% with 24 Million parameter and training time of 142 seconds.

4.2.7.4 Comparison of Lightweight RefineNet with SOTA Nodule Segmentation CAD systems

The existing algorithms proposed to perform the segmentation of lung nodules of CT scans are dominated by deep convolution networks. The proposed lightweight RefineNet is compared with the SOTA CAD systems in terms of DSC, IoU, SEN, and PPV, which is presented in Table 4.2.

Table 4.2: Comparison of the proposed Light-weight RefineNet with SOTA Lung Nodule Segmentation CAD systems

Ref.	Methods used	DSC (in %)	IoU (in %)
Cao et al. (2020)	DBResNet	82.74	-
Dolejsi et al. (2009)	Deep ResNet	94.68	-
Ronneberger et al. (2015)	U-Net	94.97	-
Wang et al. (2017)	Central focused CNN	88.15	71
Messay et al. (2015)	Regression neural Network	-	74
Proposed method	Light-weight RefineNet	98	93.5

The proposed architecture is not only better in terms of performance but also is low-complex in nature i.e. it takes fewer trainable parameters. A deep network named dual branch ResNet (DBResNet) was proposed by Cao et al. (2020), which captures multi-view and multi-scale features from the CT scan images. However, the method's performance can be improved as lung nodule segmentation is a critical task. Also, the proposed method is a data-driven approach, requiring a huge amount of data to get good performance. Another residual network was proposed by Dolejsi et al. (2009), which gave better performance, but the deep network is computationally extensive. U-Net architecture is one of the most common segmentation architectures used for segmentation tasks. Ronneberger et al. (2015) have proposed the U-Net model to perform lung nodule segmentation. The method resulted in a low DSC value which is not a reliable factor to segment nodules. A data-driven approach was proposed by

Wang et al. (2017) to perform nodule segmentation from CT scans. The model proposed captures both 2D and 3D multi-scale features. However, the results can still be improved. A regression neural network was proposed to segment pulmonary nodules by Messay et al. (2015). A fully-automated segmentation approach is introduced in this work and also has good IoU performance.

4.2.7.5 Performance evaluation for proposed Lung Nodule Segmentation system on multiple datasets

The proposed method is compared with multiple datasets as mentioned in Table 4.3. The additional datasets considered for evaluation are International Early Lung Cancer Action Program (I-ELCAP), Lung time, and Lung Nodule Database (LNDb). These datasets are the popular datasets used for lung cancer nodule segmentation. I-ELCAP dataset consists of 50 low-dose documented whole-lung CT scans for detection (Reeves et al. 2017). The lung time dataset consists of 157 scans, divided into two sets (Dolejší et al. 2009). LNDb dataset consists of 294 CT scans (Pedrosa et al. 2021). All the datasets are annotated with lung nodule locations. The datasets are evaluated for the detection of nodules in CT scans. However, the nodules could not be further classified into benign and malignant due to a lack of ground truth. The proposed lightweight RefineNet model performed well on all the datasets.

Table 4.3: Performance evaluation for proposed Lung Nodule Segmentation system

Dataset	DSC (in %)	IoU (in %)	SEN (in %)	PPV (in %)
LUNA16 (Iandola et al. 2016)	98.0	93.5	98.5	98.4
I-ELCAP (Reeves et al. 2017)	97.5	95.9	95.0	97.1
Lung TIME (Dolejší et al. 2009)	98.3	95.1	92.0	96.3
LNDB (Pedrosa et al. 2021)	96.8	94.3	94.5	95.1

4.2.7.6 Performance comparison of various Optimization Algorithms along with parameters used

The tuning of hyperparameters is performed using optimization algorithms such as Ant-Colony Optimization (ACO), BAT algorithm, Artificial Bee Colony (ABC) optimization, Genetic Algorithm (GA), and Particle Swarm Optimization (PSO). The algorithms are bio-inspired optimization techniques used to achieve the optimal

4. Lung Cancer Nodule Segmentation

solution for a given problem. Each algorithm's working is carried out with different parameter initialization. The parameters considered for the different optimization algorithms combined with CondenseNet architecture and its respective performance results achieved are presented in Table 4.4. The initial parameters considered are chosen in an empirical manner. The parameters providing the highest performance results are presented in the table. The best results for lung nodule classification are achieved using the PSO algorithm fused with CondenseNet architecture.

Table 4.4: Performance comparison of various Optimization Algorithms used

Optimization Algorithm	Parameters used	ACC (in %)	SEN (in %)	SPE (in %)
ACO-based CondenseNet	Number of ants=15 Initialisation of pheromone=0.5 Weight of pheromone of decision=0.5 Number of iterations=100 Number of generations=500	90.84	89.03	90.31
BAT-based CondenseNet	Population size=50 Loudness=0, 30 Pulse rate=0, 5 Minimum frequency=0 Maximum frequency=100 Number of iterations=100	75.48	76.71	74.98
ABC-based CondenseNet	Colony size=200 Limit=100 Number of onlookers=100 Number of employed bees=100 Number of scouts=1	79.3	78.83	79.94
GA-based CondenseNet	Initial population size=50 Offspring/parent populations size=50 Mutation rate=15 Fitness function=Maximize Sorting method=Ascending Probability of crossover=85% Number of iterations=100	66.4	65.9	67.4
PSO-based CondenseNet	Number of iterations=100 inertia weight=0.7 cognitive and social parameter=2 Number of iterations=100	98.7	98.8	97.9

4.2.7.7 AUC-ROC graph representing various Optimization Algorithms combined with CondenseNet architecture

The proposed CondenseNet architecture is evaluated with various optimization algorithms such as ACO, ABC, BAT, GA, and PSO, illustrated in Figure 4.7. The algorithms displayed varying performances when fused with CondenseNet architecture for tuning of hyperparameters. The ideal Area Under the Receiver Operating Characteristic Curve (AUC-ROC) curve will be nearer to a value of 1.0. The higher the AUC, the better the classifier. From the experimentation, it can be concluded that the best AUC value is achieved by the PSO-based CondenseNet classifier.

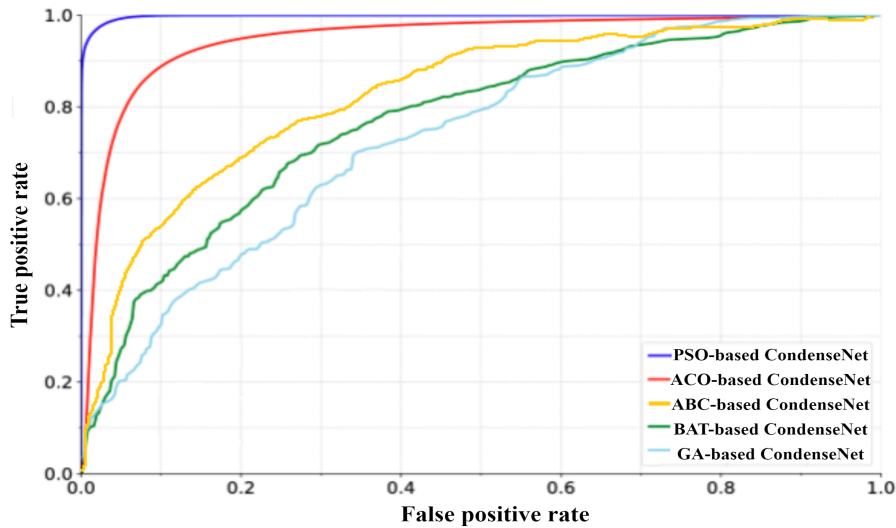


Figure 4.7: AUC-ROC graph representing various Optimization Algorithms combined with CondenseNet architecture

4.2.7.8 Analysis of accuracy and loss for the proposed Classification method

The proposed PSO-based CondenseNet method is trained for 50 epochs. Figure 4.8 illustrates the accuracy and loss values throughout the training and validation phase of the network. In order to attain the optimal performance of the system, the accuracy value should be higher, and the loss value should be lower towards the end of the network's training. The graph shows that the proposed method is trained similarly, resulting in the best accuracy and lowest loss for the classification problem. Because different inputs are supplied in each batch, the curve is not exactly exponential, which may result in poor performance. However, at the end of the training phase, the model

4. Lung Cancer Nodule Segmentation

learns better, and the best weights are chosen from the validation phase and used to evaluate the unknown data.

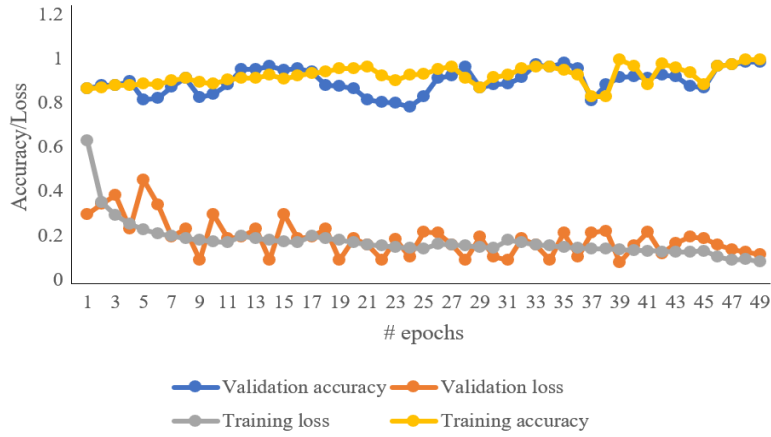


Figure 4.8: Graph representing the accuracy and loss achieved

4.2.7.9 Performance of Lung Nodule Classification system with existing methods

The detected lung nodules are further classified into benign and malignant nodules using a novel classification method named PSO-based CondenseNet. The performance metrics used to evaluate the proposed method are ACC, SEN, SPE, PR, F1-score, FPR, and FNR. The proposed method is evaluated with SOTA lightweight classification systems: SqueezeNet, Squeeze-and-Excitation Net (SENet), ShuffleNet, MobileNetv2, and CondenseNet. The results obtained for all the methods are presented in Table 4.5. The proposed method achieves the best results with 98.7%, 98.8%, 97.9%, 95.3%, 96.1%, 0.0495, and 0.0231 for ACC, SEN, SPE, PR, F1-score, FPR, and FNR values, respectively.

Table 4.5: Performance of the proposed Lung Nodule Classification method

Models	ACC (in %)	SEN (in %)	SPE (in %)	PR (in %)	F1-score (in %)	FPR	FNR
SqueezeNet	88.3	89.2	86.9	88.1	87.5	0.0640	0.0438
SENet	89.9	90.5	89.5	90.4	91.3	0.0545	0.0441
ShuffleNet	90.3	92.1	94.9	93.6	95.1	0.0622	0.0458
MobileNetv2	95.4	93.0	92.7	94.5	95.3	0.0521	0.0372
CondenseNet	95.1	96.3	93.4	92.5	96.0	0.0568	0.0454
PSO-based CondenseNet	98.7	98.8	97.9	95.3	96.1	0.0495	0.0231

4.2.7.10 Comparison of PSO-based CondenseNet with SOTA Nodule Classification CAD systems

The proposed PSO-based CondenseNet method is compared with the SOTA nodule classification systems and is provided in Table 4.6. The existing CAD systems are dominated by CNN models for the classification task. Hussein et al. (2017) have proposed a deep learning CNN model with high-level attributes in which the transfer learning technique has been adapted. A stacked autoencoder method was proposed by Naqi et al. (2020) to classify lung nodules. The method uses multiple CNN models to achieve better performance. However, the number of parameters for multiple CNNs will be higher, causing the model to be complex. A 3D Deep Neural Network was proposed by Liao et al. (2019), which extracts volumetric information from the input scans. However, the system's performance is less and can not be used for real-time deployment. Tran et al. (2019) have introduced an autoencoder model with deep features. The sensitivity and FP/scan value obtained in their method are not satisfactory for a reliable CAD system. Also, one of the major drawbacks of using deep learning models is that they can overfit for less data and also consume a lot of resources to train better. Therefore, in this work a lightweight CNN model is proposed for the classification of lung nodules. The results obtained by the proposed method are better than the existing CAD systems in terms of performance and also training time.

Table 4.6: Comparison of the proposed PSO-based CondenseNet with SOTA systems

Ref.	Methods used	ACC (in %)	SEN (in %)	SPE (in %)
Hussein et al. (2017)	High-level attributes, CNN	91.26	-	-
Naqi et al. (2020)	Stacked autoencoders	96.9	95.6	97
Liao et al. (2019)	3D DNN	81.42	-	-
Tran et al. (2019)	Deep features with autoencoder	75.01	83.35	-
Proposed method	PSO-CondenseNet	98.7	98.8	97.9

4.2.7.11 Comparison of Time and Space Complexity for proposed models

Computational complexity is a significant factor considered while deploying any model for real-time usage. The reason being less availability of high configuration

4. Lung Cancer Nodule Segmentation

systems at every remote location. Therefore, the more the model is light-weight, the better the model for using it on mobile devices. Here, training time is considered per epoch. The time and space complexities are mentioned in Table 4.7 for segmentation and classification tasks in terms of training time in seconds (s) and number of parameters in millions (M). It can be observed that both the proposed models perform better in comparison to other existing models, making it feasible for real-time usage.

Table 4.7: Time and Space Complexity comparison for proposed models

Segmentation			Classification		
Models	Parameters (M)	Training Time (s)	Models	Parameters (M)	Training Time (s)
U-Net	30	232	SqueezeNet	27.5	580
V-Net	27	192	SENet	25.6	394
FCN U-Net	130	780	ShuffleNet	5.3	289
Mask RCNN	63	423	MobileNetv2	3.5	188
RefineNet	26	165	CondenseNet	2	103
Proposed Light-weight RefineNet	24	142	PSO-based CondenseNet	1.7	82

4.3 PROPOSED LUNG CANCER NODULE SEGMENTATION SYSTEM-II

The CT scans consist of candidates that need to be identified as nodules and non-nodules, which are performed using the EFCM method. Once the nodules are categorized, it needs to be further classified into benign and malignant nodules. The nodule can be assigned a malignant label based on the malignancy score provided by expert radiologists. The average score of all the radiologists is calculated, and the nodule is assigned the corresponding label. Once we get the two classes, the classification is performed using two sets of feature representations, BoVW, deep features, and the combination of both features. The classifier used is the SVM for the final decision. Figure 4.9 demonstrates the architecture of the proposed method. The details of each module in the architectures are discussed in subsections 4.3.1, 4.3.2, 4.3.3, 4.3.4, and 4.3.5.

4.3. Proposed Lung Cancer Nodule Segmentation System-II

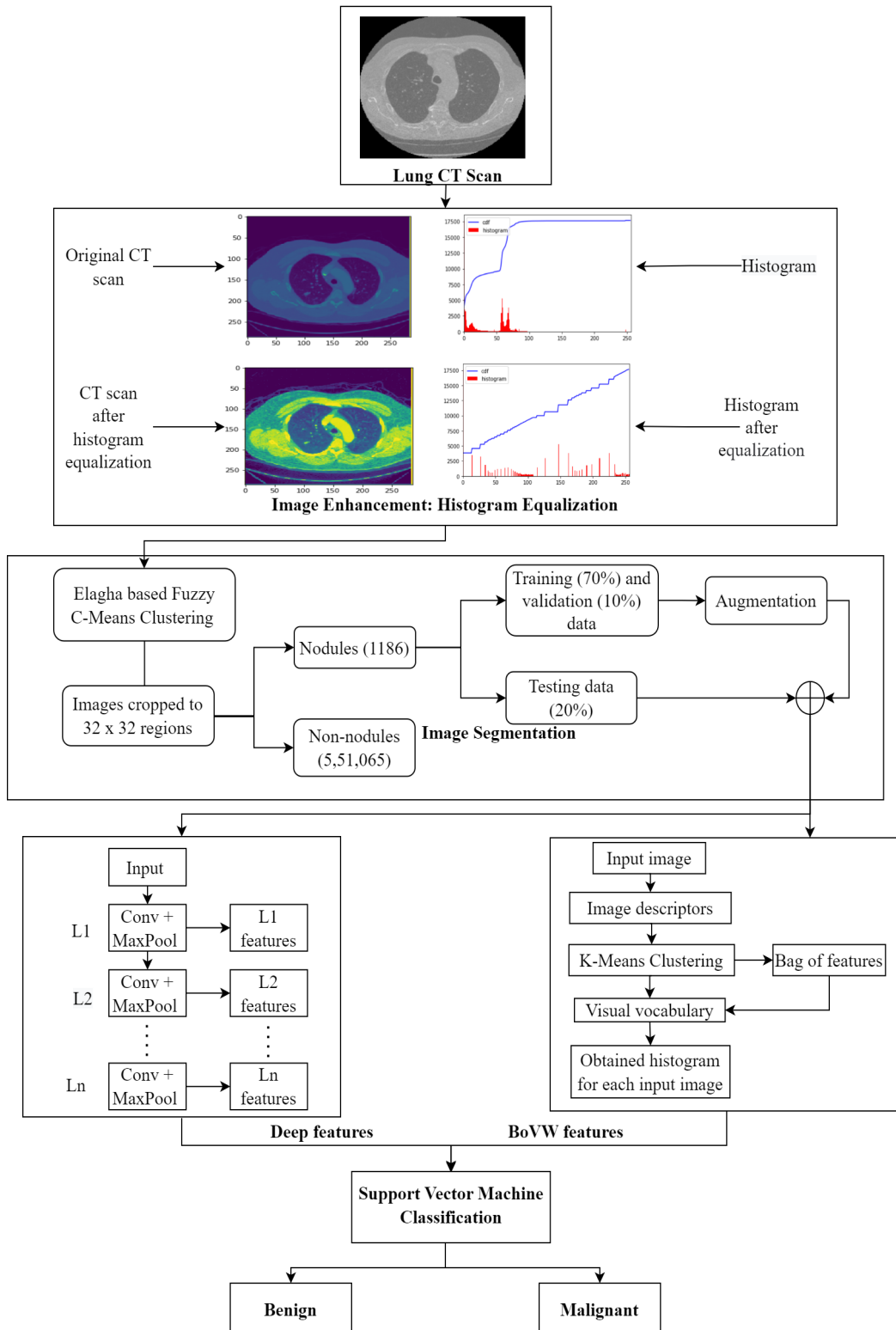


Figure 4.9: Schematic architecture of the proposed method

4.3.1 Image Enhancement

Enhancing the images draws more attention toward certain characteristics of an image, making the images more precise, sharp, and detailed. This, in turn, can be used for better analysis and information extraction from the images. In this work, the Histogram Equalization (HE) technique is used where the contrast is altered by adjusting the intensity of the image, which provides an enhanced CT scan image. The comparative frequency of occurrence of different gray levels in the image is represented in the histogram (Salem et al. 2019).

The histogram $h(r_k)$ of an image consists of a L total intensity values r_k within a range of [0-255] (refer equation 4.14).

$$h(r_k) = n_k \quad (4.14)$$

where, n_k is the number of pixels with an intensity value of r_k in the image.

The histogram can be obtained by plotting the $p(r_k)$ which is shown in below equation 4.15:

$$p(r_k) = \frac{h(r_k)}{\text{number of rows}(M) * \text{number of columns}(N)} = \frac{n_k}{MN}; k = 0, 1, 2, \dots, (L-1) \quad (4.15)$$

The HE of an image is a transformation function i.e, Cumulative Distribution Function (CDF) is given in below equations 4.16 and 4.17:

$$cdf(k) = \sum_{i=0}^k p_r(r_i), k = 0, 1, \dots, L - 1 \quad (4.16)$$

$$s(k) = T(r_k) = \left\lfloor (L - 1) \sum_{i=0}^k p_i \right\rfloor = \left\lfloor \frac{(L - 1)}{MN} \sum_{i=0}^k n_i \right\rfloor; k = 0, 1, \dots, (L - 1) \quad (4.17)$$

where, k is total intensity value, p_r is plot of histogram, and r is intensity value of a pixel in the image.

4.3.2 Image Segmentation

The proposed work uses a novel clustering approach named Elagha initialization-based Fuzzy C-Means clustering to segment the nodule region from the given input CT scan. At first, FCM partitions the image into several clusters, and then the cluster centroids are selected randomly to compute the Euclidean distance. This random selection of initial centroids may lead to the local optimum solution. Thus, to overcome this drawback, Elagha initialization is used to initialize centroids. It generates the initial centroids based on the overall shape of the data. This modification in traditional FCM is termed EFCM.

EFCM method divides the input image (X) into M clusters such that $x_j = x_1, x_2, \dots, x_m$. Then, Elagha initialization calculates the initial cluster centroids by identifying the boundaries of data points and divides them into F rows and F columns to calculate the initial centroids. The width w_j and height h_j of the grid cell is computed as shown in equations [4.18](#) and [4.19](#):

$$w_j = \frac{w_{j,max} - w_{j,min}}{F} \quad (4.18)$$

$$h_j = \frac{h_{j,max} - h_{j,min}}{F} \quad (4.19)$$

Where $w_{j,max}$ and $w_{j,min}$ represents the maximum and minimum widths, $h_{j,max}$ and $h_{j,min}$ signifies the maximum and minimum heights respectively. The N number of initial cluster centroids (c_i) is given by equation [4.20](#).

$$c_i = \frac{w_j}{2} + \frac{h_j}{2}, i = 1, 2, \dots, N \quad (4.20)$$

After initialization of centroids, the membership function calculation of each pixel is done using equation [4.21](#):

$$\mu_{ij} = \frac{1}{\sum_{i=1}^N \sum_{j=1}^M \left(\frac{1}{d_{ij}} \right)^{\frac{2}{q-1}}} \quad (4.21)$$

Where, q indicates the power exponent, d_{ij} is the Euclidean distance between samples x_j and cluster centroid c_i and is given by, equation 4.22:

$$d_{ij} = \sqrt{\sum_{i=1}^N \sum_{j=1}^M (x_j - c_i)^2} \quad (4.22)$$

The objective function ξ used for the initialization of the FCM algorithm is given by equation 4.23:

$$\xi = \sum_{i=1}^N \sum_{j=1}^M \mu_{ij} d_{ij}^2 \quad (4.23)$$

The clusters are formed for nodule and non-nodule regions into separate groups based on the Euclidean distance. The output of the EFCM algorithm is a segmented image consisting of lung nodule regions.

4.3.3 Bag-of-Visual Words

Learning feature representations from images using the BoVW method is a two-tiered process. The information from the segmented images is extracted from a pre-generated codebook or dictionary consisting of low-level local features, also known as visual words. The image descriptors used in this work are Scale-Invariant Feature Transform (SIFT) features. A visual dictionary is represented using a histogram named “Bag of Visual Words”, which is used as a mid-level feature representation (Sundarambal et al. 2021). The words in the image mean information in a patch of an image. The patch size must be larger than a few pixels to retrieve more and better information, as it should consider key parts like corners or edges.

SIFT operation is based on the local edge histogram technique. The SIFT technique is one of the popular methods that work very effectively for the BoVW method. Densely sampled SIFT features are extracted from the images. K-means algorithm is used to get cluster centers to generate a visual codebook or dictionary on these features. A histogram is built to the nearest code in the codebook based on the number of occurrences of a feature in each image. The image is then divided into sub-regions of size 2×2 , and histograms are built for each sub-region. Once all the

histograms are generated, all the sub-region histograms are concatenated to form a single feature vector.

4.3.4 Deep Features

In medical imaging, deep architectures are mostly used for final decision-making. However, in this work, deep architecture is used as a feature representation. The deep learning models are well-known for learning hierarchical information from the input images. The higher the layers, the more information the network learns. This novel set of features is used for classifying cancerous and non-cancerous nodules. Images of both categories are trained separately using a deep CNN architecture, and intermediate features of both classes are extracted. The deep features learn better representations as the network gets deeper.

4.3.5 Nodule Classification

Classification of detected nodules into cancerous and non-cancerous is performed using an SVM classifier. The model is trained using BoVW, deep features, and the combination of these features. The kernel used for SVM is linear. The hyperparameters set for the SVM model are cross-validation parameter set to 5 and the cost parameter set to 0. The model is tested using probability estimates generated from the trained model for the classification.

4.3.6 Results and Discussion

This section discusses about the results obtained by the proposed segmentation and classification models. The performance metrics considered for segmentation task are DSC, IoU, SEN, and PPV, and for classification task are accuracy, error rate, specificity, sensitivity, FPR, and F-score.

4.3.6.1 Nodule Segmentation

Figure 4.10 shows the dominance of the proposed EFCM segmentation method. The proposed model obtains the DSC of 97.10%, whereas existing methods obtain lower values, such as U-Net of 80.36%, V-Net (92.86%), Fully Connected Network (FCN) U-Net (91.20%), and Mask Region-based CNN (Mask RCNN) (71.16%). Also, the

4. Lung Cancer Nodule Segmentation

IoU of the proposed technique is 91.96%, but the existing methods show lesser values. Likewise, the SEN value of 95.35% makes the proposed model preferable to the current techniques. On the other hand, the proposed system attained the PPV of 96.30%, which is higher than the existing methods. However, the second best performing method was V-Net followed by FCN U-Net. It can be inferred that the proposed EFCM segmentation model displayed an improved performance in terms DSC, IoU, SEN, and PPV.

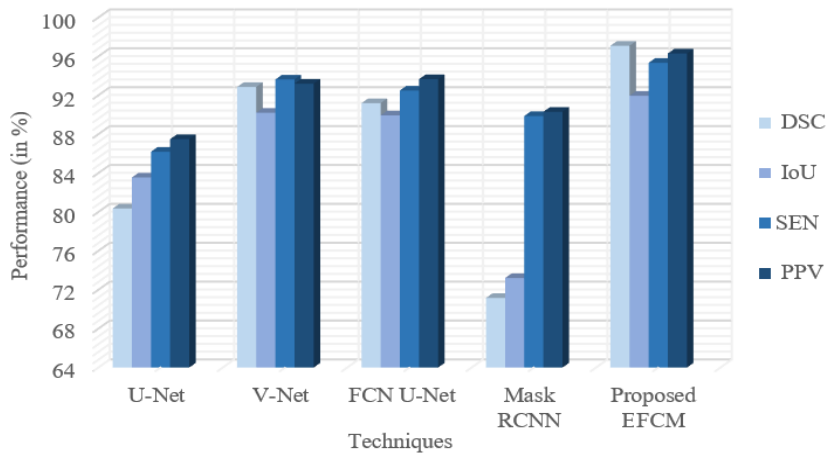


Figure 4.10: Comparison of proposed EFCM model with existing techniques

The evaluation of the nodule segmentation system for the LUNA16 dataset is performed using a primary performance metric named FPs/scan. Figure 4.11 illustrates the FROC curve achieved for the proposed EFCM method. The graph depicts that the proposed method resulted in low FPs/scans, proving it is a better performing system. The FPs/scan result for the proposed EFCM model is 2.7 FPs/Scan with a sensitivity of 95.35%.

The proposed EFCM segmentation method is compared with the existing lung nodule segmentation systems in Table 4.8. The methods considered for comparison are mostly deep learning architectures such as U-Net (Ronneberger et al. 2015), dual branch residual network (Cao et al. 2020), CNN (Wang et al. 2017), deep Fully Convolution Networks (FCN) (Roy et al. 2019), and so on. In recent trends, deep learning architectures have taken over image segmentation techniques. However in the proposed method, a clustering approach for segmentation attained improved results of 97.10% for DSC, and 91.96% of IoU respectively.

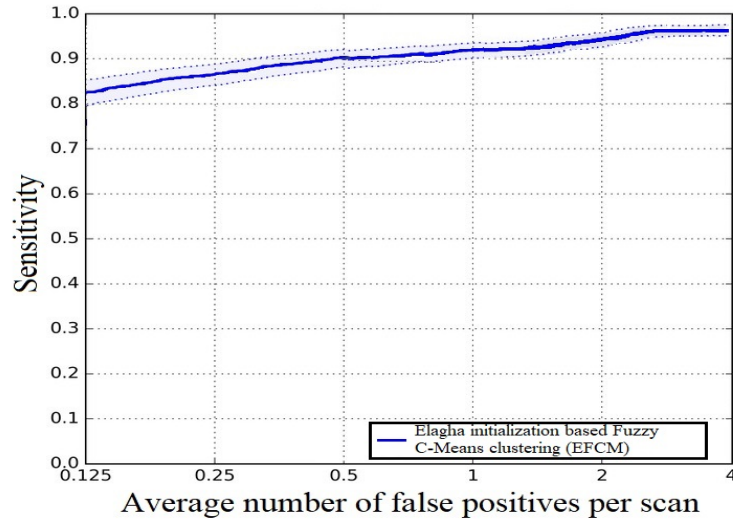


Figure 4.11: FROC curve for the performance of proposed EFCM at 2.7 FPs/scan

Table 4.8: Comparison of the proposed system with the SOTA systems

Ref.	Methods used	DSC (in %)	Ref.	Methods used	IoU (in %)
Ronneberger et al. (2015)	U-Net	94.97	Wu et al. (2018)	Segmentation Attributes and Malignancy Prediction	58.00
Cao et al. (2020)	Dual branch residual network	82.74	Aresta et al. (2019)	iW-Net	55.00
Roy et al. (2019)	Deep FCN	93.00	Messay et al. (2015)	Regression Neural Network	74.00
Wang et al. (2017)	Central focused CNN	82.15	Wang et al. (2017)	Central focused CNN	71.00
Proposed method	EFCM	97.10	Proposed method	EFCM	91.96

4.3.6.2 Nodule Classification

The performance of the lung nodule classification system is evaluated on the publicly available LUNA16 dataset. The accuracy obtained for the lung nodule classification task is 96.87%. The performance metrics considered for the evaluation of the proposed method are accuracy, error rate, specificity, sensitivity, FPR, and F-score. The results are presented in Table 4.9.

4. Lung Cancer Nodule Segmentation

Table 4.9: Performance of the Lung Nodule Classification for the proposed system

Performance metric	BoVW	Deep features	BoVW + Deep features
ACC (in %)	93.48	95.32	96.87
Error rate (in %)	6.52	4.68	3.13
SPE (in %)	93.04	95.01	96.60
SEN (in %)	93.94	95.61	97.15
FPR	0.0696	0.0499	0.0340
F-score	0.9337	0.9522	0.9681

Figure 4.12 illustrates the quantitative analysis of the three feature representations using the Receiver-Operating Characteristics (ROC) curve. The ideal system provides an Area Under Curve (AUC) of 1. The AUC values attained for BoVW, deep, and BoVW + Deep features are 0.83, 0.88, and 0.92, respectively. The classification performance analysis is presented for the SVM classifier. It can be noted from the Figure 4.12 that BoVW + Deep features resulted in the highest AUC of 0.92. The feature combination worked effectively to improve the performance of the system.

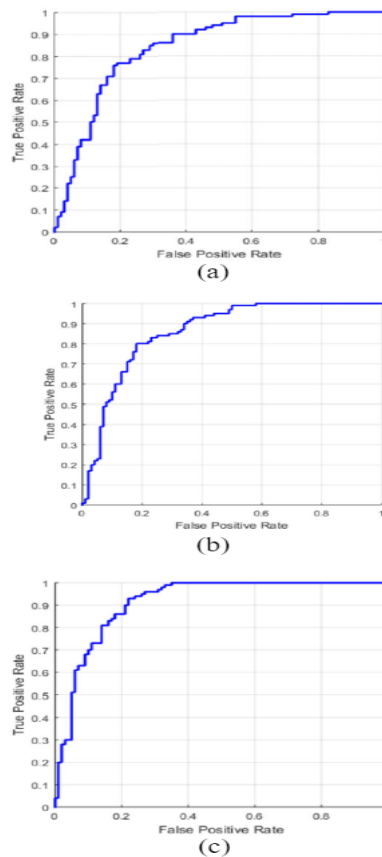
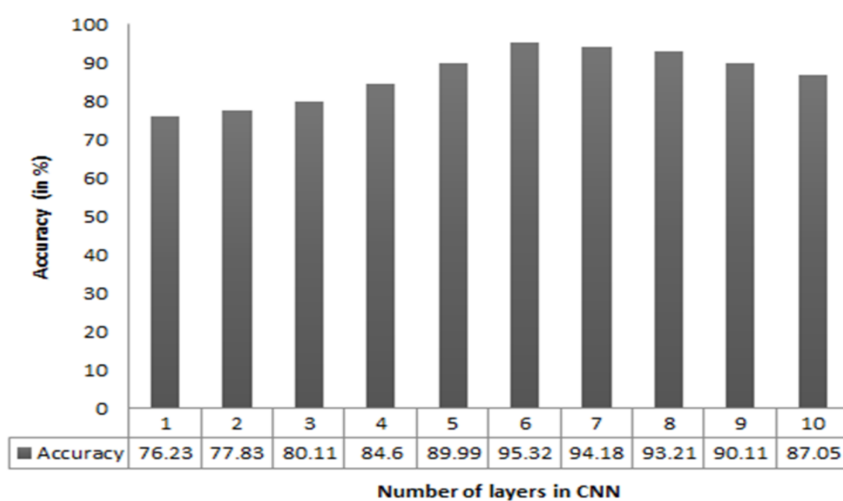
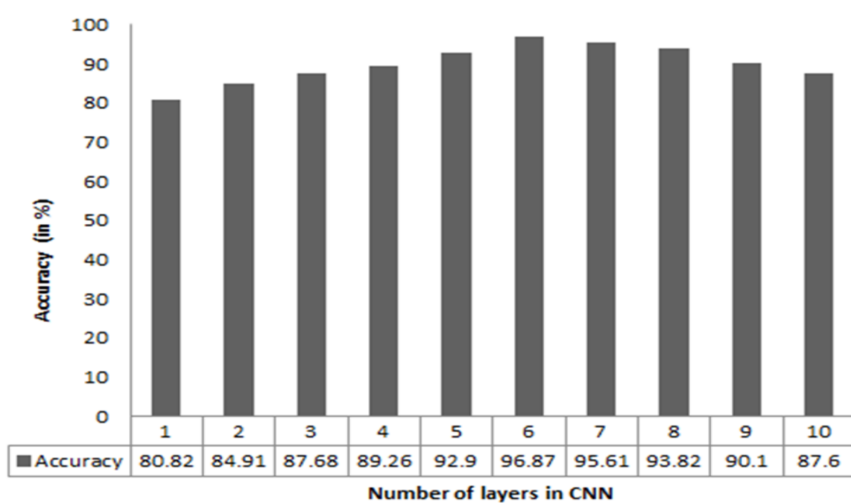


Figure 4.12: ROC curves for (a) BoVW, (b) Deep features, (c) BoVW + Deep features

A layer-wise feature comparison is performed for deep and BoVW + deep features. Figure 4.13 illustrates the accuracy values obtained for different intermediate layers for deep features. The number of layers considered is from 1 to 10. It can be noticed that the performance of the system increases with the increase in layers. After 7th layer, there is a degradation of accuracy observed in the figure. This is due to overfitting the model for more layers as complexity increases, and less data is available to train the deep architecture. Therefore, the number of layers must be monitored to prevent the model from overfitting.



(a)



(b)

Figure 4.13: Layer-wise accuracy values for (a) Deep features, (b) BoVW + Deep features

4. Lung Cancer Nodule Segmentation

The proposed method is compared with SOTA lung nodule classification systems. The systems previously proposed for performing lung nodule classification system utilizes deep learning architectures such as ANN, CNN, multi-scale CNN, Stacked Auto Encoder (SAE), etc. The results are presented in Table 4.10. It can be noted from the table that the proposed method achieved better performance as compared to the SOTA lung nodule classification systems.

Table 4.10: Comparison of the proposed system with the SOTA systems

Ref.	Methods used	ACC (in %)	SEN (in %)	SPE (in %)
Silva et al. (2016)	Taxonomic indexes and phylogenetic trees, SVM	88.44	84.22	90.06
Song et al. (2017)	CNN, DNN, SAE	84.15	N/A*	N/A
Shen et al. (2015)	Multi-scale CNN, Random forest	86.84	N/A	N/A
Gupta et al. (2018)	Super-Resolution CNN, SVM	85.70	N/A	N/A
Shaukat et al. (2019)	Intensity, shape, texture features and ANN	93.70	95.50	94.28
Proposed method	BoVW + Deep features, SVM	96.87	97.15	96.60

N/A*-Not Available

Another major issue in training these networks is it is computationally expensive as it requires a lot of time to train a deep model. However, in our method, we used CNN as a feature extractor rather than a classifier. It does not require much time to extract intermediate features. It is also computationally less expensive as the classifier used does not require much time to learn BoVW and deep features. Because BoVW features are encoded, representations do not take up more time for calculation.

4.4 SUMMARY

Number of CAD systems have been proposed in the past years. However, some limitations of the existing systems are that they are trained on less data, the complexity of the model is too high, the model is computationally intensive, the system's performance is not reliable for deployment in real-time, etc.

This chapter is mainly focused on two proposed novel segmentation algorithms. The first proposed system introduces a novel low-complex deep learning approach to perform lung nodule segmentation and classification. There is a need to develop low-complex models as the conventional CNN models take up a lot of computer resources and are largely dependent on the size of the dataset. A novel lightweight RefineNet is introduced to segment lung nodule candidates from CT scans. Before feeding the identified nodules to the classifier, there is a need for more cancerous and non-cancerous nodules. More images are generated using a DC-GAN model, a SOTA data augmentation technique. The identified nodules are further fed to a lightweight PSO-based CondenseNet to classify them into cancerous and non-cancerous nodules. The primary goal of this work is to create a CAD system that is both performance and computationally-efficient. The proposed system achieves performance at par with the SOTA CAD systems and is lightweight. The key contribution of this work is that the models proposed for both segmentation and classification take significantly fewer parameters to achieve high performance in terms of DSC, IoU, ACC, SEN, and SPE.

In this second proposed system, a novel clustering-based segmentation method named EFCM is designed to extract lung nodules from the given CT scan. A hybrid of two different types of feature representations for lung nodule classification is proposed in the work. The method glorifies that deep learning can be used as a classifier and as a suitable feature extractor. The segmentation method introduced in this work performs better than existing segmentation methods in terms of DSC, IoU, and PPV. The proposed method acknowledges that a combination of certain feature representations can enhance the system's performance in terms of various evaluation metrics such as ACC, SEN, etc. It also reduces the computational cost of the system by reducing the system's learning parameters. The proposed system effectively combines the encoded feature representation method BoVW and deep features extracted from intermediate layers of a CNN. The performances obtained in the medical imaging tasks are rather critical and also require quicker output. The proposed method provided better and faster results than other CAD systems proposed for the lung nodule classification tasks.

CHAPTER 5

LUNG CANCER NODULE CLASSIFICATION

5.1 INTRODUCTION

Cancer is one of the world's deadliest illnesses, with a high death rate. Cancer is formed by abnormal cell growth in any tissue, which leads to the formation of tissue lumps, masses, or nodules. Lung cancer is one of the most life-threatening cancers, accounting for the majority of cancer-related deaths. There has been increasing interest in research in the early identification of lung cancer by investigating lung nodules. Some CAD systems have been previously developed, but there is still scope for improving their performance to identify and classify lung nodules. Nodules in lung cancer may be characterized into two categories, namely, cancerous and non-cancerous. Malignant nodules are cancerous, while benign nodules are not. One of the primary determinants of the nodule type is size. In this chapter, a novel nodule classification method is proposed to categorize the nodules into cancerous and non-cancerous.

The contributions of the chapter are :

- A new preprocessing technique named Boosted Bilateral Histogram Equalization (BBHE) is introduced in this work to improve the quality of CT scans.
- Level-1 classification is performed using the proposed Cauchy Black Widow Optimization-based Convolutional Neural Network (CBWO-CNN), in which Cauchy mutation is used to choose the best weights.

- A novel Squeeze-and-Excitation-Xception (SE-Xception) CNN model with shared network parameters is proposed for performing Level-2 classification of lung cancer in CT scans.

The chapter is organized as follows: Section 5.2 discusses the proposed methodology in detail. Section 5.3 briefs the results and discussion, and Section 5.4 summarizes the proposed method and its significance.

5.2 PROPOSED METHODOLOGY

This work aimed to perform lung nodule classification using a bi-level classification approach. In Level-1, candidates are classified into nodules and non-nodules, and in Level-2, the detected nodules are further classified into benign and malignant. The proposed classification approach is divided into four main stages: data preparation, Level-1 classification, transfer-learned knowledge, and Level-2 classification. The block diagram of the proposed methodology is shown in Figure 5.1.

In the first stage, data preparation involves the preprocessing of input CT scans. Pre-processing helps to denoise the scan, improve its visibility which in turn enhances the quality of the image. The preprocessed scans are fed as an input to the Level-1 classification task. Boosted Bilateral Histogram Equalization filtering technique is used to preprocess input CT scans and is explained in detail in section 5.2.1. The second stage is to differentiate between nodules and non-nodules. The dataset consists of very few positive nodules as compared to non-nodules. To mitigate this issue, positive lung nodule data is augmented only for training and validation dataset splits. This process is described in the data augmentation section 5.2.2. The level-1 classification task is performed using the proposed CBWO-CNN. The third stage is the transferring of learned knowledge from the CBWO-CNN model to the proposed SE-Xception model. In machine learning, transfer learning (Torrey and Shavlik 2010) focuses on using the knowledge acquired while completing one task to complete another task that is linked to it. The reason for adapting the transfer learning approach is that the size of the data used for classifying benign and malignant nodules is significantly less. Deep learning models perform well on large datasets. Therefore, to

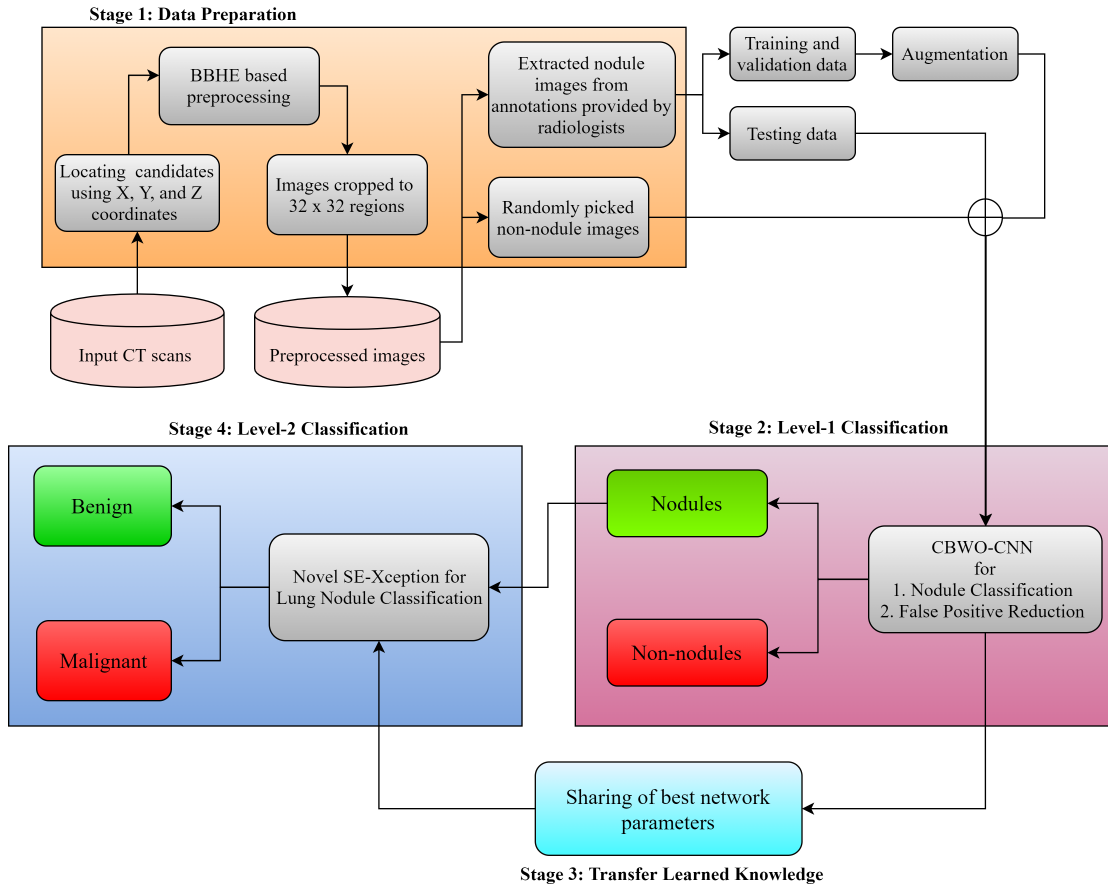


Figure 5.1: Block diagram of the proposed approach

address this issue, a weight-sharing scheme is adapted in this work. The best pre-trained weights obtained from the trained CBWO-CNN model for nodule and non-nodule images are used as initialization weights for training the SE-Xception model for benign and malignant nodules. The fourth stage of the proposed method is the classification of lung nodules into benign and malignant. An architecture named “SE-Xception” is proposed in this work. A detailed description of all the methods is provided in the below sections.

5.2.1 Image Preprocessing

Preprocessing provides the quality enhanced image to locate small particles in the scanned image. A CT scan is a medical imaging procedure that uses X-rays and computer technology to produce detailed, cross-sectional images of the body. CT scans are stored in the image format of the ‘raw’ and ‘mhd’ extensions. These CT

scans are loaded using the python tool SimpleITK (Lowekamp et al. 2013). The location of candidates is provided in X, Y, and Z coordinates. The candidate location is chosen based on the center of the nodule or non-nodule. A 32 x 32 region is cropped out from each candidate location and saved into two classes, positive and negative, i.e., nodules and non-nodules. These classes are decided based on the annotations provided by the radiologists. The reason for reducing the size of the image to 32x32 is that the size of the nodules is very small in the CT scan and can consist of background noise which can lead to mis-classification. When a scan is cropped to a size of 48 x 48 or larger, more background noise is introduced, which causes mis-classification. When the scan size is reduced to 16 x 16, the nodule is clipped and significant data is lost.

Various preprocessing methodologies have been developed, but still, quality remains to be a challenge. A Boosted Bilateral Histogram Equalization algorithm has been developed to improve the image quality so that small parts can be seen clearly. The detailed illustration of the BBHE is as follows:

Initially, the histogram is generated for the image $\bar{h}(\chi_j) = n_j^p$. Then the histogram is smoothed by bilateral filtering to preserve the edges in an image given by equation 5.1.

$$\bar{h}_\chi(\chi_j) = \bar{h}(\chi_j) * \psi \quad (5.1)$$

Where, χ_j represents the j^{th} intensity level, n_j^p denotes the number of pixels having intensity level χ_j , and ψ denotes the bilateral filtering. After that, boosting is applied on the edge preserved histogram given by equation 5.2. Boosting is a technique that enhances the edges detected by the bilateral filtering process.

$$\bar{h}_B(\chi_j) = \begin{cases} \frac{\bar{h}_\chi(\chi_j) - P_{min}}{(P_{max} - P_{min})} (m(k) - P_{min}) \alpha + P_{min}, & \text{if } \bar{h}_B(\chi_j) > P_{min} \\ \bar{h}_\psi(\chi_j), & \text{otherwise} \end{cases} \quad (5.2)$$

Where, $m(k)$ denotes the peak histogram's smoothed value, P_{min} and P_{max} denotes the local minimum and maximum pixel values, α boosts the minor regions which is given by $\alpha = \log(P_{max} - P_{min}) / \log(m(k) - P_{min})$. Now, mapping is done using

Histogram Equalization (HE) and enhancement in the contrast of the image is obtained (χ'). δ is multiplied by the gain (Γ_g) and noise reduction (Γ_n) functions to improve details of the edges. The detail gain function is given as equation [5.3](#).

$$\Gamma_{detail(i,j)} = [\Gamma_g(i,j) \cdot \Gamma_n(i,j)] * B(i,j) \quad (5.3)$$

Where, $\Gamma_g(i,j) = 1 / \{\mathbb{R} \cdot [\chi_s(i,j) + 1.0]^p\}$ and where $\Gamma_n(i,j) = N_{offset} + \{N_{band} / [1 + e^{-\delta(i,j)-E}]\}$.

Finally, by multiplying the details with the detail function (Γ_{detail}) using equation [5.4](#), the improved detail image δ' is obtained.

$$\delta'(i,j) = \sum \Gamma_{detail} \cdot \delta(i,j) \quad (5.4)$$

Equation [5.5](#) is used to integrate the χ' and δ' in the final step.

$$\chi''(i,j) = w \times \chi_s(i,j) + (1 - w) \times \delta'(i,j) \quad (5.5)$$

Where, the final enhanced image is $\chi''(i,j)$, and w is a weighted function. The value of w can be anywhere between 0.0 and 1.0.

After preprocessing of CT scans, sample images of nodules, non-nodules, benign and malignant nodules are shown in Figure [5.2](#). There is a considerable difference in this dataset's number of nodules and non-nodules. To avoid the above-mentioned skewness in the data, the non-nodule data images are sub-sampled, i.e., the images are randomly selected from all the subsets of the dataset. Nodules must be further classified into benign and malignant nodules after the nodule and non-nodule categories have been established.

5.2.2 Data Augmentation

In the above step, there is a clear data imbalance between the two classes. There are a total of 1186 positive lung nodules in the LUNA16 dataset out of 5,51,065 total candidates. To overcome this problem, the augmentation of positive nodules is performed. The images are augmented by performing some image operations such as

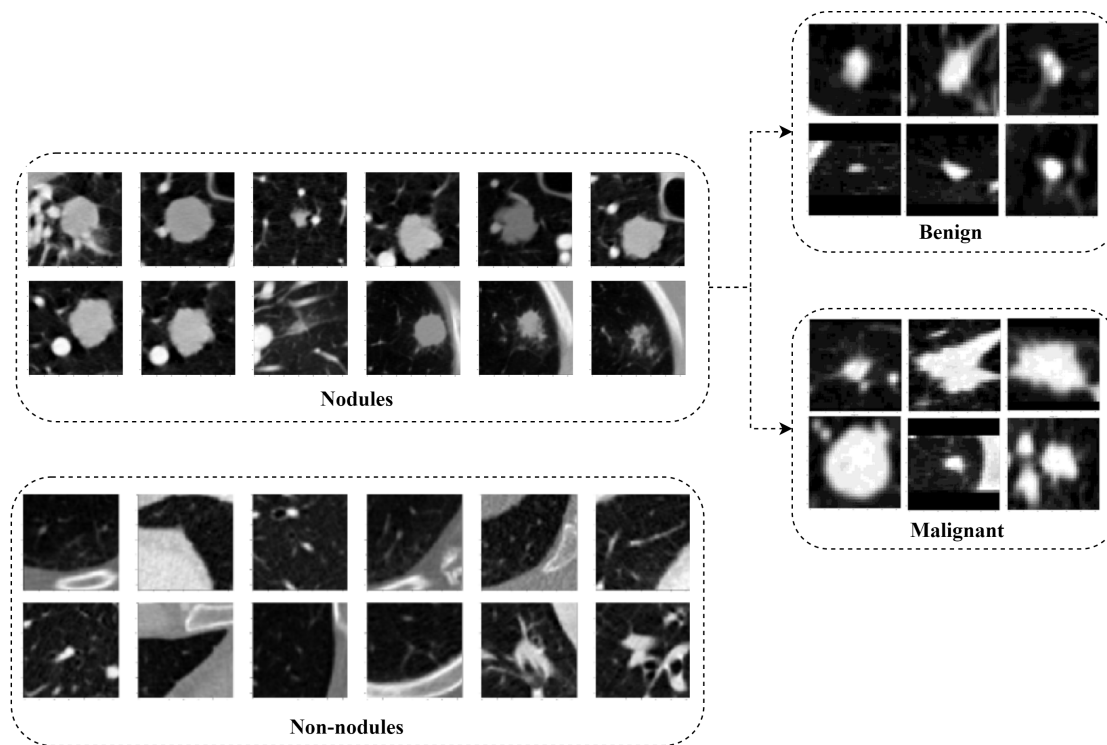


Figure 5.2: Nodule, non-nodule, benign, and malignant images after preprocessing of CT scans

modifying brightness, contrast, random rotation of the image to 90 degrees, transposing, scaling the images, and flipping the images horizontally and vertically. Figure 5.3 shows some of the augmented images of a nodules' CT scan. The augmentation images shown in the figure are taken only from one CT scan.

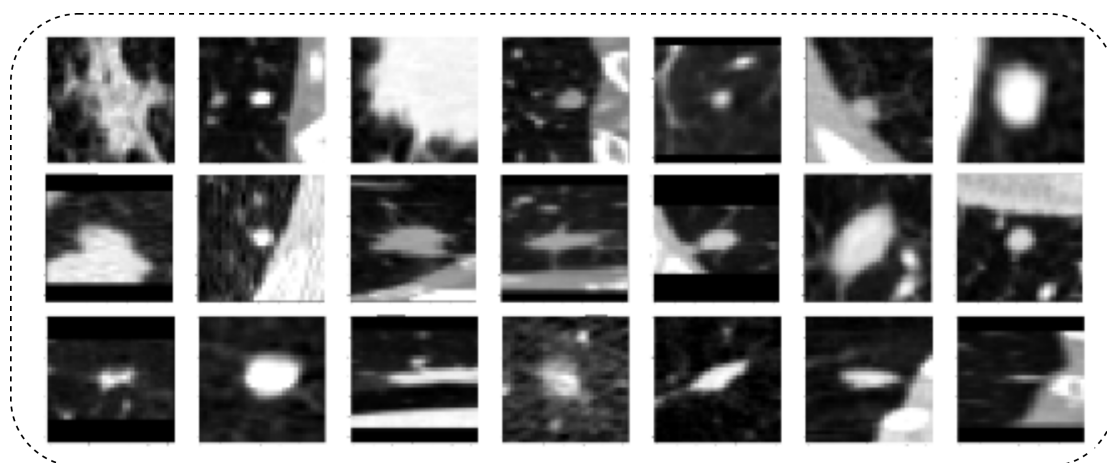


Figure 5.3: Augmented images of a nodule CT scan

5.2.3 Level-1 Classification: CBWO-CNN

Classifying candidates in a lung CT scan into nodules and non-nodules is critical because some nodules resemble tissues or organs in that area, making it difficult to classify the nodule correctly. The training should be done properly to avoid high bias and low variance as well as low bias and high variance.

Algorithm 1 : CBWO-CNN algorithm

- 1: **Input:** Extracted features $F'_{EXT} = [\xi_1^+, \xi_2^+, \xi_3^+, \xi_4^+, \dots, \xi_N^+]$
 - 2: **Training:** Initialize the kernel ($K_{(x,y)}$), bias φ_b , pooling layers (l_{layers}^{PL}), weights
 - 3: **Loop:** for i pixels in image $\xi_{x,y}^+$
 - 4: **Step 1:** Evaluate the convolution operation using
 - 5: $\varphi_2^{conv} = L_{relu}(\varphi_B + \sum_{x=0}^2 + \sum_{y=0}^2 w_{i,j} \xi_{x,y}^+)$
 - 6: **Step 2:** Evaluate the pooling layer using
 - 7: $\varphi_3^{pooling} = l_{layers}^{PL}(|\xi_{x,y}^+|)$
 - 8: **Step 3:** Generate flattened feature vectors
 - 9: $\varphi_3^{pooling}(\zeta_{flatten}) = [\xi_1^+, \xi_2^+, \xi_3^+, \xi_4^+, \dots, \xi_n^+]$.
 - 10: **Step 4:** Evaluate the fully connected layer using
 - 11: $\varphi_4^{FC} = \gamma(\sum_{i=1}^n w_i \zeta_{flatten} + \varphi_b)$
 - 12: **Step 5:** If $\sum(\xi_{x,y}^+ - \bar{\xi}_{x,y}^+) = 0$
 - 13: Lung cancer nodule is detected
 - 14: **Else** Update weights using
 - 15: $\partial_{fitness} = f(W_k^+) = f\{w_1^+, w_2^+, w_3^+, w_4^+, \dots, w_n^+\}$
 - 16: $w_1 = \alpha \times w_1^+ + (1 - \alpha) \times w_2^+$
 - 17: $w_2 = \alpha \times w_2^+ + (1 - \alpha) \times w_1^+$
 - 18: **Step 6:** Based on cannibalism the strong solution is preserved and then
 - 19: Cauchy mutation is performed for better accuracy rate
 - 20: $v_w^{k+1} = (1 - \alpha) w_k^+ v_w^k + \alpha (\eta_i \cdot N(0, \sigma)) + \mu_1 \mu_2 (p_k - w_k^+)$
 - 21: **EndIf**
 - 22: **EndFor**
 - 23: **Step 7:** For $\bar{\xi}_{x,y}^+ - >$ Candidate nodule
 - 24: **If** malignancy rate > 3 and malignancy rate ≤ 5
 - 25: Nodule is positive
 - 26: **Else** Nodule is negative
 - 27: **EndIf**
 - 28: **EndFor**
 - 29: **Output:** Lung cancer nodule detection
-

The detailed discussion of algorithm [1](#) is given as follows:

1. Pre-processed images in terms of features are given as input to the CBWO-CNN algorithm. These features are denoted as F'_{EXT} which are fed as input to a

convolutional layer that offers related feature mapping with not losing the related features that are best significant aimed at prediction.

2. The parameters required for training the model are kernel ($K_{(x,y)}$), bias φ_b , pooling layers (l_{layers}^{PL}), weights. In this step, these parameters are initialized empirically.
3. The first step in training the CBWO-CNN is convolutional layers \wp_2^{conv} . Convolutional neural networks takes an input image matrix $\xi_{x,y}^+$ and then convolves it with filters or else kernels $K_{(x,y)}$ to take out the features. The input image is convolved utilizing a filter, and also this convolution procedure learns the identical feature prevalent on the total image. The output matrix's size with no padding (i.e.) $|\xi_{x,y}^+| \times K_{(x,y)}$, the window slides after every procedure, and the features are learned via the feature maps. The feature maps capture the image's and work's local receptive field utilizing shared weights and also biases.
4. In the second step, the pooling layer is responsible for reducing the Convolved Feature's spatial size $\wp_3^{pooling}$. This is to reduce the computing power needed by dimensionality reduction to process the data. Herein, l_{layers}^{PL} signifies the diverse pooling layer i.e., it may be maximum pooling layer or else minimum pooling layer or else average pooling layer, etc. In this work, maxpooling is used.
5. In the third step, the output from the convolutional layer is flattened $\wp_3^{pooling}(\zeta_{flatten})$. The convolutional, as well as pooling layers, form CNN's layer together. The number of layers like these are incremented aimed at capturing the low-level information much more relying on the images' complexity. Hence, the convolutional layer's output is flattened.
6. In the fourth step, the fully connected layer's (FCL's) input is the convolutional layer's output that is flattened and then fed as input to the FCL. The flattened vector trains the FCL that is identical to that of Artificial Neural Network. The vector's training is executed utilizing, $\wp_4^{FC} = \gamma(\sum_{i=1}^n w_i \zeta_{flatten} + \varphi_b)$. Herein, φ_b signifies the bias that is randomly initialized; w_i implies the respective input

node's weight; γ signifies the activation function and the FCL utilizes the relu activation function aimed at acquiring the input image's vector values

7. Step 5 is performed to minimize the loss using backpropagation, which, minimizes the loss between the observed output and the actual output. Backpropagation calculates the gradient of the loss function with respect to each parameter in the network, which tells us how much changing that parameter would affect the loss. The algorithm then adjusts the parameters in a way that reduces the loss. This is done by iteratively updating the weights and biases in the opposite direction of the gradient (hence the name "backpropagation"), meaning that the network moves down the gradient towards the minimum of the loss function. The magnitude of the updates is controlled by a learning rate hyperparameter, which determines how much the parameters are adjusted in each iteration. CBWO provides with selecting the best weights that is $W_K^+ = \{w_1^+, w_2^+, w_3^+, w_4^+, \dots, w_n^+\}$ and improving the accuracy.
8. In step 6, to update novel weights, a population is created and inside the population, the parents are mated to create offspring's procreation. During mating, diverse eggs are obtained. However, only strong eggs are sustained and stored in an array for reproducing then offspring and from the parents w_1 and w_2 . Now, centred on cannibalism the strong solution is conserved, and next, the Cauchy mutation is executed aimed at an excellent accuracy rate v_w^{k+1}
9. In step 7, based on the malignancy rate, nodule classification is performed. For every candidate nodule, a malignancy rate is calculated by averaging the malignancy score provided by 4 radiologists in LUNA16 dataset. The malignancy score values lie between the range 1 to 5, where 1 being least malignant and 5 being most malignant. If the average score is below 3, the nodule is considered as negative and if the average score is above 3 and below 5, the nodule is considered as positive.
10. The output returned by the algorithm is whether the given candidate input is nodule (positive) or non-nodule (negative).

5.2.4 Level-2 Classification: SE-Xception

Lung cancer nodule classification is a challenging task. SE-Xception, a combination of two best-performing and popular deep learning architectures, SE-Net, and Xception, is proposed in this work. SE-Net architecture consists of a block named as Squeeze-and-Excitation block. In this block, the channel-wise feature responses are adaptively recalibrated using modeling channel inter-dependency explicitly. The SE block is shown in Figure 5.4(a).

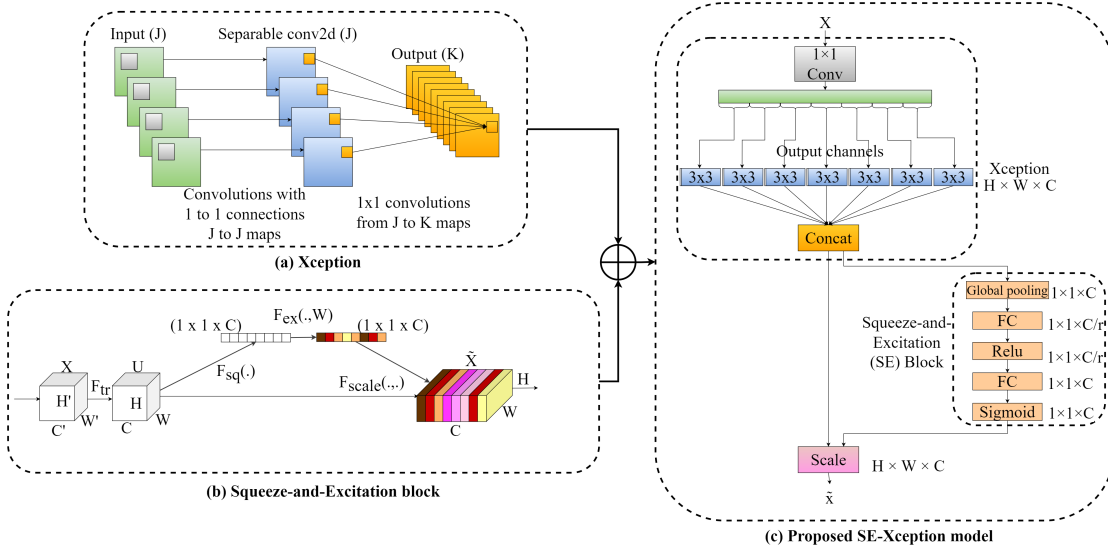


Figure 5.4: Diagrammatic representation of (a) Xception network, (b) Squeeze-and-Excitation block, and (c) proposed SE-Xception model

In figure 5.4(b), the SE block diagram, F_{tr} represents a convolution operation where the input X is transformed to U . The F_{tr} in the proposed work is an Xception block. In previous works, inception and residual blocks are used as convolution operations. F_{tr} is represented in equation 5.6.

$$F_{tr} : X \rightarrow U, X \in \mathbb{R}^{H' \times W' \times C'}, U \in \mathbb{R}^{H \times W \times C} \quad (5.6)$$

The notation $V = [v_1, v_2, \dots, v_C]$ is used to illustrate a set of learned filter kernels, where v_C denotes the parameters of filter kernel. The outputs can be written as: $U = [u_1, u_2, \dots, u_C]$, where u_C is given in equation 5.7.

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s X x^s \quad (5.7)$$

The features that are obtained after performing F_{tr} are U . The first operation carried out in the network is passing features through the squeeze operation (F_{sq}). A channel descriptor is produced from the feature maps when passed through the squeeze operation. The feature maps are aggregated across the spatial dimensions (H x W). An embedding of the global distribution of feature responses is generated channel-wise from this descriptor. It makes the information from the network's global receptive area that all its layers are to use. The squeeze operation is shown in equation [5.8](#). By shrinking U by its spatial dimensions H x W, a statistic $z \in R$ is generated such that c^{th} is calculated.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (5.8)$$

An excitation operation is performed after the squeeze operation to capture the aggregated information of the channel descriptors. This operation fully captures the channel-wise dependencies, where it learns the non-linear and non-mutually-exclusive relationship of the channels. This operation is represented in equation [5.9](#), where Rectified linear unit (ReLU) activation function is denoted using δ notation, $W_1 = \mathbb{R}^{\frac{C}{r} \times C}$, and $W_2 = \mathbb{R}^{\frac{C}{r} \times C}$. ReLU activation function returns the output value from $\max(0, x)$ where x is the input. Consequently, it eliminates the values that are negative and offers a much more simple computation.

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (5.9)$$

The network's SE blocks' capacity and computational cost can be changed by adjusting a hyperparameter named the reduction ratio r . This also helps to investigate the trade-off between performance and computational cost. Using r , a bottleneck with two completely connected (FC) layers is created with dimensionality reduction. This block's final output is given by rescaling U with s activations, which can be expressed in equation [5.10](#).

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (5.10)$$

where $X_c = x_1, x_2, \dots, x_c$, $F_{scale}(u_c, s_c)$ represents channel-wise multiplication of scalar s_c and the feature map $u_c \in \mathbb{R}^{(H \times W)}$.

Xception network architecture is completely based on depth-wise separable convolution blocks. The working of the Xception model is illustrated in Figure 5.4(b). The hypothesis includes calculating spatial correlations, so it is possible to decouple cross-channel correlations in the CNN feature maps fully. The name Xception means “Extreme Inception”. It is named Xception because it has a more robust hypothesis than the underlying architecture of InceptionV3 (Xia et al. 2017). Xception architecture consists of 36 convolution layers. These convolution layers are organized into 14 different modules. Every module consists of linear residual connections except for the first and last layers. In brief, the Xception network can be said to be the stacking of depth-wise separable convolution layers in a linear fashion, consisting of residual connections. The input data is mapped to spatial correlations for each output channel separately, and then a 1 x 1 depthwise convolution operation is performed. This operation captures the cross-channel correlation. These correlations can be pictured as a 2D+1D mapping instead of a 3D mapping. Here, the 2D space correlations are performed first, and then the 1D space correlation is performed. Xception proved to provide slightly better results as compared to InceptionV3 on the LUNA16 dataset.

The proposed SE-Xception model is a combination of SE-Net and Xception. SE-Net consists of a squeeze-and-excitation block, which performs the operations mentioned in the above sections. The addition of these modules in the Xception reduces the parameters of the model. Figure 5.4(c) shows the graphical representation of the proposed methodology. The figure represents the operations performed in the SE-Xception model. Only one module operation is illustrated. Input X is first passed to an Xception module. The input is given to a convolution filter with a filter size of 1, and then it is passed to a convolution filter with a filter size of 3. These two operations are concatenated, which is known as depth-wise separable convolution. This operation

is followed by a SE-block where the squeeze and excitation operation is performed. Initially, to generate a channel descriptor, a global pooling operation is performed on the Xception module's output. The pooling operation is followed by a Fully Connected (FC) layer with the ReLU as the activation function. The sigmoid activation function acts as a simple gating mechanism. This operation is known as the excitation operation. Once this step is completed, the output is scaled, represented as \tilde{X} .

5.3 RESULTS AND DISCUSSION

5.3.1 Level-1 Classification

This section describes the performance of the proposed models. As the data is skewed, the non-nodules are selected by performing sub-sampling, where a set of non-nodule images are chosen from each subset to balance the data. The data imbalance (Setio et al. 2017b) may be the reason for overfitting (Banik and Bhattacharjee 2021) the model, which affects the model's performance. The model is validated on 10% of the total data and tested on 20% of the data. Adam optimizer is used in both networks. The loss function used is binary cross-entropy (Ruby and Yendapalli 2020). The total number of epochs is set to 200. An early-stopping criterion is used while training, in which if no improvement is found in the validation loss after 20 epochs, the training is terminated. CBWO-CNN model is initially evaluated on four different activation functions, namely, Exponential Linear Unit (ELU) (Clevert et al. 2015), Tanh (Nwankpa et al. 2018), LeakyReLU (Glorot et al. 2011), and ReLU (He et al. 2015). The model performance of various activation functions on the CBWO-CNN model is demonstrated in Figure 5.5.

Considering ELU as an activation function in the networks, its convergence is faster than other activation functions (Clevert et al. 2015). The difference between tanh and sigmoid lies in the range of the output value it returns. Tanh output value ranges from -1 to 1, whereas the sigmoid output value ranges from 0 to 1. However, the tanh function's performance in the models' hidden layers is relatively poor compared to other activation functions. LeakyReLU activation function is an extension of ReLU. The alpha value is added to the function to solve the issue of "dying ReLU" in this activation function.

5. Lung Cancer Nodule Classification

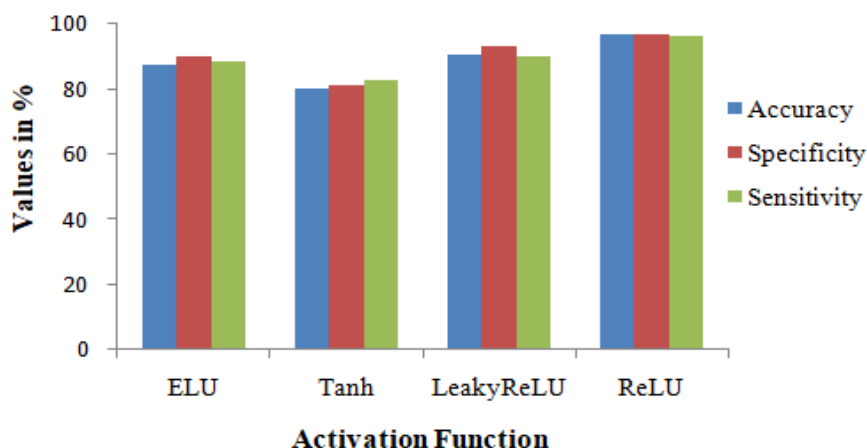


Figure 5.5: CBWO-CNN model evaluated on various activation functions on testing data

The results obtained from the LeakyReLU activation function in the hidden layer are second-best among the other activation functions used to evaluate the model. ReLU provides faster and more accurate results as it removes the negative values to pass to the next layer. The model performed best with the ReLU activation function in the hidden layer.

The performance metrics used to evaluate the CBWO-CNN model are accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PR), F1-score, and false positive rate (FPR). The corresponding results achieved from these confusion matrices are illustrated in Table 5.1. The average results obtained from the 5-fold validation of the model are depicted in the table. The table represents the performance measures stated above with their corresponding results.

Table 5.1: Performance of CBWO-CNN model

Performance measures	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
ACC (in %)	95.21	94.04	96.18	95.77	97.63	96.37
SEN (in %)	96.16	96.81	95.23	94.57	97.73	96.10
SPE (in %)	94.29	97.26	97.11	96.94	97.54	96.63
PRE (in %)	94.26	97.18	96.98	96.79	97.48	96.53
F1-score (in %)	95.21	97.00	96.10	95.67	97.61	96.32
FPR	0.057	0.027	0.028	0.030	0.024	0.033

Several popular deep-learning architectures are used to evaluate the lung cancer

classification system. The architectures chosen for the evaluation of the Level-1 classification system are CNN, VGG-19, Inception-V3, Deep Belief Network, ResNet, Recurrent Neural Network, and proposed CBWO-CNN. ACC, SEN, and SPE are the performance metrics that are evaluated. The results obtained are presented in Table 5.2. Hence, it can be noted that the proposed CBWO-CNN model performed better for Level-1 classification.

Table 5.2: Evaluation of CBWO-CNN

Models	ACC (in %)	SEN (in %)	SPE (in %)
CNN	84.65	82.88	83.47
VGG-19	87.89	86.77	86.56
Inception-V3	86.51	85.98	87.40
Deep Belief Network	89.95	87.21	88.68
ResNet	94.65	92.80	93.91
Recurrent Neural Network	95.83	94.41	95.03
Proposed CBWO-CNN	96.37	96.10	96.63

5.3.2 Level-2 Classification

The task of classifying positive lung nodules into benign and malignant nodules is known as lung nodule classification. In this work, the lung cancer nodule classification is performed using the proposed SE-Xception model with shared parameters from the CBWO-CNN model trained in Level-1 classification. To get the best performing model, the proposed model is evaluated on four activation functions in the model's hidden layer. The activation functions used are ELU, Tanh, LeakyReLU, and ReLU. The network proved to provide better performance for the ReLU activation function. The four activation functions' performance has been assessed for three models, SE-Net, Xception, and proposed SE-Xception is demonstrated in Figure 5.6.

The proposed SE-Xception is evaluated with performance metrics such as accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PR), F1-score, and false positive rate (FPR). The results achieved are presented in Table 5.3. The results are presented without shared network parameters and with shared network parameters for the SE-Xception model. The model trained without shared parameters resulted in overfitting of the model due to fewer training images. The accuracy achieved for the proposed

5. Lung Cancer Nodule Classification

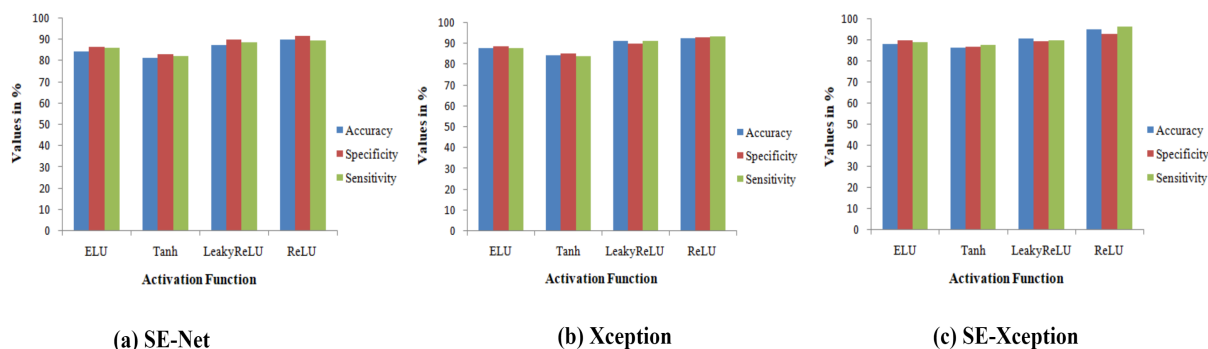


Figure 5.6: (a) SE-Net, (b) Xception, and (c) proposed SE-Xception evaluated on various activation functions on testing data for Level-2 Classification

SE-Xception is 82.29%. The improvement in the result is performed using pre-trained weights from Level-1 classification. This improved accuracy by almost 12%. The accuracy achieved for the proposed SE-Xception was 94.76%. Therefore, the practice of transfer learning using pre-trained models for a new task rather than training a new model from scratch provides some advantages such as improved performance, reduced training time and data requirements.

Table 5.3: Performance assessed without and with shared network parameters

Performance measures	Without shared parameters	With shared parameters
ACC (in %)	82.29	94.75
SEN (in %)	85.85	96.14
SPE (in %)	78.49	92.83
PRE (in %)	80.95	94.89
F1-score (in %)	83.33	95.52
FPR	0.21	0.07

The proposed SE-Xception model is compared with the previous works performing lung cancer nodule classification. The results achieved from the proposed model outperformed the previous works. Performance comparison of the previous works is illustrated in Table [5.4](#).

The effect of the proposed SE-Xception method is visually demonstrated in Figure [5.7](#). The images given in the green box are malignant nodules correctly classified as malignant nodules, and the images presented in the red box are benign nodules

Table 5.4: Comparison of previous works with proposed model

Reference	Methods used	ACC (in %)	SEN (in %)	SPE (in %)
Gupta et al. (2018)	Super-Resolution CNN, SVM	85.70	-	-
Shen et al. (2015)	Multi-scale CNN, Random forest	86.84	-	-
Silva et al. (2016)	Taxonomic indexes and phylogenetic trees, SVM	88	82	94
Song et al. (2017)	CNN, DNN, SAE	84.15	-	-
Shaukat et al. (2019)	Intensity, shape, texture features and Artificial Neural Network	93.70	95.50	94.28
Cao et al. (2020)	3D tensor filtering, 3D level set segmentation, correlation feature selection, and random forest	-	84.62	-
Proposed work	SE-Xception	94.75	96.14	92.83

misclassified as malignant nodules. The reason for some of this misclassification is the indistinguishable similarity in the anatomical structure of the nodules. Even though the FPR of the proposed method is significantly less, the method still has scope for improvement, as minimal misclassification is also not acceptable in medical applications.

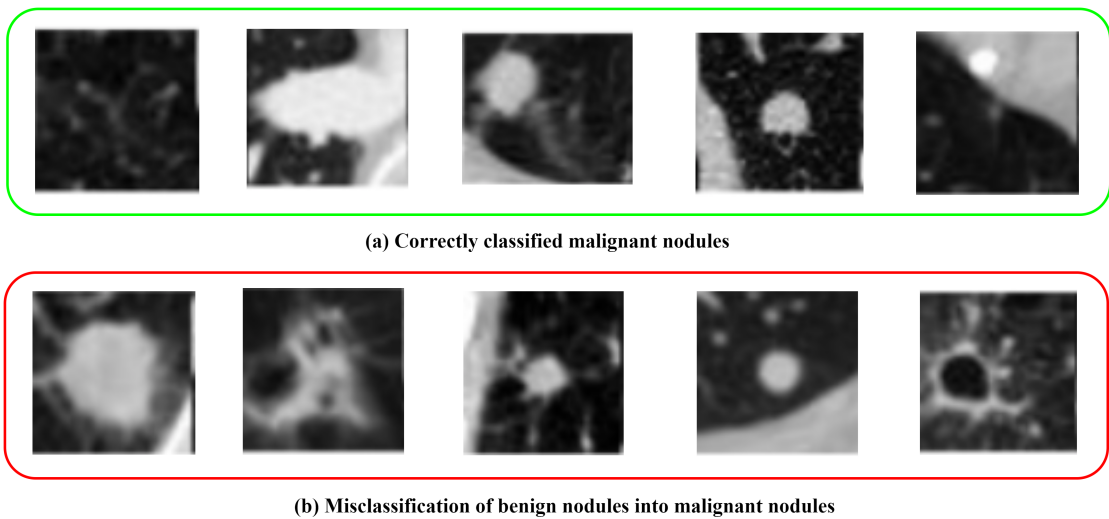


Figure 5.7: Visual depiction of correctly and mis-classified nodules for Level-2 Classification

5.4 SUMMARY

This chapter introduced a bi-level classification model for the classification of lung cancer nodules. To enhance the CT scan images; a new BBHE technique is used. Level-1 classification performs bifurcation among nodules and non-nodules found in the candidates based on the locations provided by the radiologists. Level-2 classification performs the classification of benign and malignant lung cancer nodules from the nodules identified in the Level-1 classification. A deep learning architecture, CBWO-CNN, is introduced to perform Level-1 classification, and a SE-Xception network is proposed to perform Level-2 classification. The proposed SE-Xception model uses the shared network parameters from the CBWO-CNN model used for training nodules and non-nodules. This hybrid network is designed and developed using recently proposed and best-performing models, namely, SE-Net (Hu et al. 2018) and Xception (Chollet 2017). SE-Net is made use to obtain a lesser number of parameters. The Xception network is an extreme version of the InceptionV3 network (Xia et al. 2017). CBWO-CNN and SE-Xception models are the core contributions of this chapter.

CHAPTER 6

CONCLUSIONS AND FUTURE SCOPE

Lung nodule detection using computer-aided algorithmic techniques is challenging, and the research has been carried out for three decades. Deep learning algorithms are a recent trend that has displayed drastic performance improvement in medical imaging tasks. The analysis in the report includes an in-depth study of the specific traditional/conventional and deep learning algorithms for the identification and classification of lung cancer nodules. There is a need to develop a better automated and time-tested pulmonary detection system to increment the life span of lung cancer patients and provide early diagnosis of lung cancer. So, our aim was to focus on the challenges in the detection, segmentation, and classification of lung cancer nodules and target to build a reliable CAD system that will help the radiologist for a second opinion.

The first work introduced a novel deep-learning architecture named RFR V-Net for detecting lung cancer nodules with reduced FP. Further, nodule classification is performed using a new combination of SqueezeNet and ResNet, named NCNet. In our second work, a new lightweight RefineNet is introduced to segment lung nodule candidates from the CT scans. More images are generated using a DC-GAN model, a SOTA data augmentation technique. The identified nodules are further fed to a lightweight PSO-based CondenseNet to classify them into cancerous and non-cancerous nodules.

In our third work, a new clustering-based segmentation method named EFCM is

proposed to extract lung nodules from the given CT scan. The work also encourages that deep learning can be used as a classifier and as a suitable feature extractor. The proposed system effectively combines the encoded feature representation method BoVW and deep features extracted from intermediate layers of a CNN. In our fourth work, a bi-level classification model is proposed for the classification of lung cancer nodules. To enhance the CT scan images, a novel BBHE technique is used. A new deep learning architecture, CBWO-CNN, is introduced to perform Level-1 classification, and a novel SE-Xception network is introduced to perform Level-2 classification. The proposed SE-Xception model uses the shared network parameters from the CBWO-CNN model used for training purposes.

The proposed models' top results for the detection, segmentation and classification tasks were 98.21% ACC, 98.0% DSC, and 98.7% ACC, respectively. Hence, the proposed approaches provided improved results than existing lung cancer CAD systems.

6.1 FUTURE SCOPE

The techniques proposed in this thesis for lung cancer detection, segmentation, and classification outperform existing models. However, there is significant scope for future work. The work and ideas presented in this thesis may be further extended and improved. The following are potential future research directions in this area:

- From a clinical perspective, a set of obstacles can be observed in developing a reliable CAD system. Due to the increased demand for large amounts of data for data-hungry methods like deep learning, the data made available from the providers lack quality assurance, proper annotations, fitness, and correct segmentation. Therefore, one of the important obstacles is that if automated methods are developed for these datasets, the ultimate results may not be accurate.
- The multi-dimensionality of the images retrieved is another reason for not having proper annotations of the data. The images generated from a CT scanner can be

of 100-400 slices. The more slices, the higher the data. For a scan with more slices, the annotations will be tedious as each slice has to be carefully examined manually by the radiologists. Hence, there is a need for standardized data that has been curated from trained professionals and clinicians.

- When viewed from an ethical perspective, there is no standard system to check the ethics in which the algorithm is developed. The components in the algorithm can be illegal. The deployment of these algorithms in the hospital environment can cause more harm than good. A law or regulation must be passed to verify the AI algorithms from an ethical perspective.
- The current diagnostics used for lung cancer identification mostly consider only visually recognizable findings from imaging modalities. The presence or absence of lung cancer nodules is solely determined from images such as chest radiographs, MRI, or CT scans. However, the important factor that needs to be considered is pathological differences observed in the patient along with visually recognizable findings. This can provide an accurate diagnosis of lung cancer.
- One issue that needs to be addressed is the early diagnosis of malignant nodules in the patient. The patient's survival rate can be increased by identifying the malignant nodules at the initial stage of development.
- Another issue that most developers and researchers often oversee is the type of CT scanners used across the globe. The type of CT scanners varies across the world. In well-developed countries, the CT scanners' configuration is advanced, which can generate CT slices in the image of up to 400 or more. However, the CT scanners used in developing or under-developed countries may not be that advanced. The number of slices taken from the patient may be up to 8 or 16. This huge difference in the number of slices in an image is an important drawback that needs to be taken care of when developing a reliable and deployable automated system worldwide.
- Multiple challenges occur in the visibility of the images. The structure of a malignant nodule and a normal pulmonary structure appears to be similar.

Because of the difference in the nodule's type and size, distinguishing between cancerous and non-cancerous nodules is complicated. The images that are made available for research purposes are of poor quality and noisy. Therefore, image pre-processing needs to be performed to enhance the visibility of the images.

- The variability of storage formats of the image makes it tedious to build a generalized model.
- It is also challenging to identify reliable feature representations and integrate them into CAD systems to perform the classification of benign and malignant nodules.
- The recent advent of deep learning has made things easier as they do not require hand-crafted features for training the classifier. So, there is a lot of scope in deep learning and neural network architectures to build CAD systems for lung cancer identification.
- In conventional CAD systems, annotations are quite expensive as it has to be done manually by experienced radiologists. With the help of deep learning, annotation costs can be cut down by automating the annotation process. It can be achieved by training a deep learning model with an extensive data set of annotations. This model can now be used for unseen CT scans' annotation.
- There is a need for providing clinically relevant explanations for the features discovered by the different learning algorithms as it makes the CAD system more reliable for real-time usage.
- In the previously developed CAD systems, micronodules (i.e., nodules < 3 mm) are usually truncated at the system's training as it is difficult to detect these sets of nodules. Considering these sets of nodules can provide better insight for earlier detection of nodules of lung cancer.
- The CAD system can be designed in such a way that it learns from both clinical records and medical images for multi-modal analysis.

- Many authors have considered only a subset of data in the literature, which is not a good practice as it may not be a robust model. In turn, it also affects the system's performance in realistic scenarios. Hence, there is a need to evaluate CAD systems on larger datasets. More experiments on larger datasets will improve the systems' performance and generalizability.
- One of the major concerns of developing fully functional, real-time deployable lung cancer CAD systems is the data imbalance found in the cancerous and non-cancerous nodules in the CT scans and the stage classification of the cancer nodules. Therefore, there is a need to develop a deep learning model that handles data imbalance issues and focuses more on the cancer nodules in the CT scans. This can be achieved using attention-based deep learning methods. Also, the low data issue can be temporarily solved using simulated CT scans using augmentation techniques such as Generative Adversarial Networks. In this current trend, the GAN model is gaining more attention to simulate data with fewer samples.
- Deep models are, in general, considered to be computationally intensive. Therefore to overcome these issues, in the recent trends, low-complexity deep learning models are used. These models minimize the additions and multiplications in the neural network training and, in turn, reduce the overall complexity of the model in both the time and space domains. The reason for doing this is to deploy the CAD systems in low-memory devices such as hand-held smart devices, smartphones, etc. This will help the common individual to be aware of the disease condition with the help of professional doctors.

REFERENCES

- Aggarwal, T., Furqan, A. and Kalra, K. (2015). “Feature extraction and LDA based classification of lung nodules in chest CT scan images.” In *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 1189–1193.
- Akram, S., Javed, M. Y., Akram, M. U., Qamar, U. and Hassan, A. (2016). “Pulmonary nodules detection and classification using hybrid features from computerized tomographic images.” *Journal of Medical Imaging and Health Informatics*, 6(1), 252–259.
- Alilou, M., Kovalev, V., Snezhko, E. and Taimouri, V. (2014). “A comprehensive framework for automatic detection of pulmonary nodules in lung CT images.” *Image Analysis & Stereology*, 33(1), 13–27.
- Aresta, G., Cunha, A. and Campilho, A. (2017). “Detection of juxta-pleural lung nodules in computed tomography images.” In *Medical Imaging 2017: Computer-Aided Diagnosis*, volume 10134, International Society for Optics and Photonics, 952–958.
- Aresta, G., Jacobs, C., Araújo, T., Cunha, A., Ramos, I., van Ginneken, B. and Campilho, A. (2019). “iw-net: an automatic and minimalistic interactive lung nodule segmentation deep network.” *Scientific reports*, 9(1), 1–9.
- Baldwin, D. R. (2015). “Prediction of risk of lung cancer in populations and in pulmonary nodules: Significant progress to drive changes in paradigms.” *Lung cancer (Amsterdam, Netherlands)*, 89, 1–3.
- Banik, D. and Bhattacharjee, D. (2021). “Mitigating data imbalance issues in medical image analysis.” In *Data preprocessing, active learning, and cost perceptive*

REFERENCES

- approaches for resolving data imbalance*, IGI Global, 66–89.
- Bhalerao, R. Y., Jani, H. P., Gaitonde, R. K. and Raut, V. (2019). “A novel approach for detection of lung cancer using digital image processing and convolution neural networks.” In *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, IEEE, 577–583.
- Bhavanishankar, K. and Sudhamani, M. (2015). “Techniques for detection of solitary pulmonary nodules in human lung and their classifications-a survey.” *International Journal on Cybernetics & Informatics (IJCI)*, 4(1), 27–40.
- Bonavita, I., Rafael-Palou, X., Ceresa, M., Piella, G., Ribas, V. and González Ballester, M. A. (2020). “Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline.” *Computer Methods and Programs in Biomedicine*, 185, 105172.
- Cao, H., Liu, H., Song, E., Hung, C.-C., Ma, G., Xu, X., Jin, R. and Lu, J. (2020). “Dual-branch residual network for lung nodule segmentation.” *Applied Soft Computing*, 86, 105934.
- Chiang A, F. D. (2019). *Non-small cell lung cancer*, volume 11, Lippincott Williams and Wilkins.
- Choi, W., Oh, J. H., Riyahi, S., Liu, C.-J., Jiang, F., Chen, W., White, C., Rimner, A., Mechalakos, J. G., Deasy, J. O. and Lu, W. (2018). “Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer.” *Medical Physics*, 45(4), 1537–1549.
- Chollet, F. (2017). “Xception: Deep learning with depthwise separable convolutions.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1251–1258.
- Choromańska, A. and Macura, K. J. (2012). “Evaluation of solitary pulmonary nodule detected during computed tomography examination.” *Polish journal of radiology*, 77(2), 22.
- Cicek, O., Abdulkadir, A., Lienkamp, S. S., Brox, T. and Ronneberger, O. (2016). “3d u-net: learning dense volumetric segmentation from sparse annotation.”

- In *International conference on medical image computing and computer-assisted intervention*, Springer, 424–432.
- Cieszanowski, A., Lisowska, A., Dabrowska, M., Korczynski, P., Zukowska, M., Grudzinski, I. P., Pacho, R., Rowinski, O. and Krenke, R. (2016). “MR imaging of pulmonary nodules: detection rate and accuracy of size estimation in comparison to computed tomography.” *PloS one*, 11(6), e0156272.
- Clevert, D. A., Unterthiner, T. and Hochreiter, S. (2015). “Fast and accurate deep network learning by exponential linear units (elus).” *arXiv preprint arXiv:1511.07289*.
- da Nóbrega, R. V. M., Peixoto, S. A., da Silva, S. P. P. and Rebouças Filho, P. P. (2018). “Lung nodule classification via deep transfer learning in CT lung images.” In *31st International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, 244–249.
- da Silva, G. L., da Silva Neto, O. P., Silva, A. C., de Paiva, A. C. and Gattass, M. (2017). “Lung nodules diagnosis based on evolutionary convolutional neural network.” *Multimedia Tools and Applications*, 76(18), 19039–19055.
- Dandil, E. (2018). “A computer-aided pipeline for automatic lung cancer classification on computed tomography scans.” *Journal of Healthcare Engineering*, 2018, 1–12.
- de Carvalho Filho, A. O., de Sampaio, W. B., Silva, A. C., de Paiva, A. C., Nunes, R. A. and Gattass, M. (2014). “Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index.” *Artificial intelligence in medicine*, 60(3), 165–177.
- de Carvalho Filho, A. O., Silva, A. C., de Paiva, A. C., Nunes, R. A. and Gattass, M. (2017). “Computer-aided diagnosis of lung nodules in computed tomography by using phylogenetic diversity, genetic algorithm, and SVM.” *Journal of digital imaging*, 30(6), 812–822.
- de Carvalho Filho, A. O., Silva, A. C., de Paiva, A. C., Nunes, R. A. and Gattass, M. (2018). “Classification of patterns of benignity and malignancy based on ct using topology-based phylogenetic diversity index and convolutional neural network.” *Pattern Recognition*, 81, 200–212.

REFERENCES

- de Sousa Costa, R. W., da Silva, G. L. F., de Carvalho Filho, A. O., Silva, A. C., de Paiva, A. C. and Gattass, M. (2018). “Classification of malignant and benign lung nodules using taxonomic diversity index and phylogenetic distance.” *Medical & biological engineering & computing*, 56(11), 2125–2136.
- Dolejsi, M., Kybic, J., Polovincak, M. and Tuma, S. (2009). “The lung time: Annotated lung nodule dataset and nodule detection framework.” In *Medical Imaging: Computer-Aided Diagnosis*, volume 7260, International Society for Optics and Photonics, 72601U.
- Dolejší, M., Kybic, J., Polovinčák, M. and Tůma, S. (2009). “The lung time—annotated lung nodule dataset and nodule detection framework.” *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 7260.
- Farag, A., Elhabian, S., Graham, J., Farag, A. and Falk, R. (2010). “Toward precise pulmonary nodule descriptors for nodule type classification.” volume 13, 626–33.
- Farag, A., Elshazly, S. and Farag, A. A. (2017). “Feature fusion for lung nodule classification.” *International journal of computer assisted radiology and surgery*, 12(10), 1809–1818.
- Faruqui, N., Yousuf, M. A., Whaiduzzaman, M., Azad, A., Barros, A. and Moni, M. A. (2021). “Lungnet: A hybrid deep-cnn model for lung cancer diagnosis using ct and wearable sensor-based medical iot data.” *Computers in Biology and Medicine*, 139, 104961.
- Feng, J. and Jiang, J. (2022). “Deep learning-based chest ct image features in diagnosis of lung cancer.” *Computational and Mathematical Methods in Medicine*, 2022.
- Fu, L., Ma, J., Ren, Y., Han, Y. S. and Zhao, J. (2017). “Automatic detection of lung nodules: false positive reduction using convolution neural networks and handcrafted features.” In *Medical Imaging 2017: Computer-Aided Diagnosis*, volume 10134, International Society for Optics and Photonics, 101340A.
- Gibaldi, A., Barone, D., Gavelli, G., Malavasi, S. and Bevilacqua, A. (2015). “Effects of guided random sampling of tccs on blood flow values in CT perfusion studies of lung tumors.” *Academic radiology*, 22(1), 58–69.
- Glorot, X., Bordes, A. and Bengio, Y. (2011). “Deep sparse rectifier neural networks.”

- In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 315–323.
- Gong, J., Liu, J.-Y., Sun, X.-W., Zheng, B. and Nie, S.-D. (2018a). “Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules.” *Physics in Medicine & Biology*, 63(3), 035036.
- Gong, J., Liu, J.-y., Wang, L.-j., Sun, X.-w., Zheng, B. and Nie, S.-d. (2018b). “Automatic detection of pulmonary nodules in CT images by incorporating 3d tensor filtering with local image feature analysis.” *Physica Medica*, 46, 124–133.
- Gupta, A., Das, S., Khurana, T. and Suri, K. (2018). “Prediction of Lung Cancer from Low-Resolution Nodules in CT-Scan Images by using Deep Features.” In *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 531–537.
- Hamidian, S., Sahiner, B., Petrick, N. and Pezeshk, A. (2017). “3D convolutional neural network for automatic detection of lung nodules in chest CT.” In *Medical Imaging 2017: Computer-Aided Diagnosis*, volume 10134, International Society for Optics and Photonics, 1013409.
- Han, F., Wang, H., Zhang, G., Han, H., Song, B., Li, L., Moore, W., Lu, H., Zhao, H. and Liang, Z. (2015). “Texture feature analysis for computer-aided diagnosis on pulmonary nodules.” *Journal of digital imaging*, 28(1), 99–115.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015). “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification.” In *Proceedings of the IEEE international conference on computer vision*, 1026–1034.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016). “Deep residual learning for image recognition.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- He, X., Sahiner, B., Gallas, B. D., Chen, W. and Petrick, N. (2014). “Computerized characterization of lung nodule subtlety using thoracic CT images.” *Physics in Medicine & Biology*, 59(4), 897.
- Heuvelmans, M. A., van Ooijen, P. M., Ather, S., Silva, C. F., Han, D., Heussel, C. P.,

REFERENCES

- Hickes, W., Kauczor, H.-U., Novotny, P., Peschl, H. et al. (2021). “Lung cancer prediction by deep learning to identify benign lung nodules.” *Lung Cancer*, 154, 1–4.
- Hollings, N. and Shaw, P. (2002). “Diagnostic imaging of lung cancer.” *European Respiratory Journal*, 19(4), 722–742.
- Hu, J., Shen, L. and Sun, G. (2018). “Squeeze-and-excitation networks.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7132–7141.
- Hua, K.-L., Hsu, C.-H., Hidayati, S. C., Cheng, W.-H. and Chen, Y.-J. (2015). “Computer-aided classification of lung nodules on computed tomography images via deep learning technique.” *OncoTargets and therapy*, 8.
- Hussein, S., Cao, K., Song, Q. and Bagci, U. (2017). “Risk stratification of lung nodules using 3d cnn-based multi-task learning.” In *International conference on information processing in medical imaging*, Springer, 249–260.
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J. and Keutzer, K. (2016). “Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size.” *arXiv preprint arXiv:1602.07360*.
- Jakimovski, G. and Davcev, D. (2019). “Using double convolution neural network for lung cancer stage detection.” *Applied Sciences*, 9(3), 427.
- Javid, M., Javid, M., Rehman, M. Z. U. and Shah, S. I. A. (2016). “A novel approach to CAD system for the detection of lung nodules in CT images.” *Computer methods and programs in biomedicine*, 135, 125–139.
- Kasinathan, G. and Jayakumar, S. (2022). “Cloud-based lung tumor detection and stage classification using deep learning techniques.” *BioMed Research International*, 2022.
- 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.
- Kim, B.-C., Sung, Y. S. and Suk, H.-I. (2016). “Deep feature learning for pulmonary nodule classification in a lung CT.” In *4th International Winter Conference on Brain-Computer Interface (BCI)*, IEEE, 1–3.
- Koutini, K., Eghbal-Zadeh, H., Dorfer, M. and Widmer, G. (2019). “The receptive

- field as a regularizer in deep convolutional neural networks for acoustic scene classification.” In *2019 27th European signal processing conference (EUSIPCO)*, IEEE, 1–5.
- Koutini, K., Eghbal-Zadeh, H., Haunschmid, V., Primus, P., Chowdhury, S. and Widmer, G. (2020). “Receptive-field regularized cnns for music classification and tagging.” *arXiv preprint arXiv:2007.13503*.
- Kulkarni, A. and Panditrao, A. (2014). “Classification of lung cancer stages on CT scan images using image processing.” In *IEEE International Conference on Advanced Communications, Control and Computing Technologies*, 1384–1388.
- Kumar, D., Wong, A. and Clausi, D. A. (2015). “Lung nodule classification using deep features in CT images.” In *12th Conference on Computer and Robot Vision*, IEEE, 133–138.
- Kuruvilla, J. and Gunavathi, K. (2014). “Lung cancer classification using neural networks for CT images.” *Computer methods and programs in biomedicine*, 113(1), 202–209.
- Kvale, P. A., Johnson, C. C., Tammemägi, M., Marcus, P. M., Zylak, C. J., Spizarny, D. L., Hocking, W., Oken, M., Commins, J., Ragard, L., Ping Hu, d Christine Berg, d. and Prorokd, P. (2014). “Interval lung cancers not detected on screening chest x-rays: How are they different?.” *Lung cancer*, 86(1), 41–46.
- Lakshmanaprabu, S., Mohanty, S. N., Shankar, K., Arunkumar, N. and Ramirez, G. (2019). “Optimal deep learning model for classification of lung cancer on CT images.” *Future Generation Computer Systems*, 92, 374–382.
- Lavanya, M. and Kannan, P. (2018). “Lung cancer segmentation and diagnosis of lung cancer staging using MEM (modified expectation maximization) algorithm and artificial neural network fuzzy inference system (ANFIS).” *Biomedical Research*, 29, 2919–2924.
- Li, X.-X., Li, B., Tian, L.-F. and Zhang, L. (2018). “Automatic benign and malignant classification of pulmonary nodules in thoracic computed tomography based on RF algorithm.” *IET Image Processing*, 12(7), 1253–1264.
- Liao, F., Liang, M., Li, Z., Hu, X. and Song, S. (2019). “Evaluate the malignancy of

REFERENCES

- pulmonary nodules using the 3-d deep leaky noisy-or network.” *IEEE transactions on neural networks and learning systems*, 30(11), 3484–3495.
- Liu, J.-k., Jiang, H., Gao, M.-d., He, C.-g., Wang, Y., Wang, P., Ma, H. and Li, Y. (2017). “An assisted diagnosis system for detection of early pulmonary nodule in computed tomography images.” *Journal of Medical Systems*, 41.
- Liu, M., Dong, J., Dong, X., Yu, H. and Qi, L. (2018a). “Segmentation of lung nodule in CT images based on mask R-CNN.” In *9th International Conference on Awareness Science and Technology (iCAST)*, IEEE, 1–6.
- Liu, Y., Hao, P., Zhang, P., Xu, X., Wu, J. and Chen, W. (2018b). “Dense convolutional binary-tree networks for lung nodule classification.” *IEEE Access*, 6, 49080–49088.
- Liu, Y., Wang, H., Gu, Y. and Lv, X. (2019). “Image classification toward lung cancer recognition by learning deep quality model.” *Journal of Visual Communication and Image Representation*, 63, 102570.
- Liu, Y., Wang, Z., Guo, M. and Li, P. (2014). “Hidden conditional random field for lung nodule detection.” In *IEEE International Conference on Image Processing (ICIP)*, 3518–3521.
- Lowekamp, B. C., Chen, D. T., Ibáñez, L. and Blezek, D. (2013). “The design of simpleitk.” *Frontiers in neuroinformatics*, 7, 45.
- Manikandan, T. and Bharathi, N. (2011). “Lung cancer diagnosis from CT images using fuzzy inference system.” In *International Conference on Computational Intelligence and Information Technology*, Springer, 642–647.
- Marentakis, P., Karaiskos, P., Kouloulias, V., Kelekis, N., Argentos, S., Oikonomopoulos, N. and Loukas, C. (2021). “Lung cancer histology classification from ct images based on radiomics and deep learning models.” *Medical & Biological Engineering & Computing*, 59(1), 215–226.
- McNitt-Gray, M. F., Armato III, S. G., Meyer, C. R., Reeves, A. P., McLennan, G., Pais, R. C., Freymann, J., Brown, M. S., Engelmann, R. M., Bland, P. H. et al. (2007). “The lung image database consortium (lidc) data collection process for nodule detection and annotation.” *Academic radiology*, 14(12), 1464–1474.

- Mehre, S. A., Mukhopadhyay, S., Dutta, A., Harsha, N. C., Dhara, A. K. and Khandelwal, N. (2016). “An automated lung nodule detection system for CT images using synthetic minority oversampling.” In *Medical Imaging: Computer-Aided Diagnosis*, volume 9785, International Society for Optics and Photonics, 97850H.
- Meier, L., Van De Geer, S. and Bühlmann, P. (2008). “The group lasso for logistic regression.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1), 53–71.
- Messay, T., Hardie, R. C. and Tuinstra, T. R. (2015). “Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the lung image database consortium and image database resource initiative dataset.” *Medical image analysis*, 22(1), 48–62.
- Milletari, F., Navab, N. and Ahmadi, S.-A. (2016). “V-net: Fully convolutional neural networks for volumetric medical image segmentation.” In *2016 fourth international conference on 3D vision (3DV)*, IEEE, 565–571.
- Monkam, P., Qi, S., Ma, H., Gao, W., Yao, Y. and Qian, W. (2019). “Detection and classification of pulmonary nodules using convolutional neural networks: A survey.” *IEEE Access*, 7, 78075–78091.
- Nalepa, J., Marcinkiewicz, M. and Kawulok, M. (2019). “Data augmentation for brain-tumor segmentation: a review.” *Frontiers in computational neuroscience*, 13, 83.
- Naqi, S. M., Sharif, M. and Jaffar, A. (2020). “Lung nodule detection and classification based on geometric fit in parametric form and deep learning.” *Neural Computing and Applications*, 32(9), 4629–4647.
- Neal Joshua, E. S., Bhattacharyya, D., Chakkravarthy, M. and Byun, Y.-C. (2021). “3d cnn with visual insights for early detection of lung cancer using gradient-weighted class activation.” *Journal of Healthcare Engineering*, 2021.
- Ng, Q. S. and Goh, V. (2010). “Angiogenesis in non-small cell lung cancer: imaging with perfusion computed tomography.” *Journal of thoracic imaging*, 25(2), 142–150.
- Nibali, A., He, Z. and Wollersheim, D. (2017). “Pulmonary nodule classification with

REFERENCES

- deep residual networks.” *International journal of computer assisted radiology and surgery*, 12(10), 1799–1808.
- Niranjana, G. and Ponnaivaikko, M. (2017). “A review on image processing methods in detecting lung cancer using CT images.” In *International Conference on Technical Advancements in Computers and Communications (ICTACC)*, IEEE, 18–25.
- Nwankpa, C., Ijomah, Gachagan, A. and Marshall, S. (2018). “Activation functions: Comparison of trends in practice and research for deep learning.” *arXiv preprint arXiv:1811.03378*.
- Oğul, B. B., Koşucu, P., İzşam, A. and Kanik, S. D. (2015). “Lung nodule detection in x-ray images: a new feature set.” In *6th European Conference of the International Federation for Medical and Biological Engineering*, Springer, 150–155.
- Pedrosa, J., Aresta, G., Ferreira, C., Atwal, G., Phoulady, H. A., Chen, X., Chen, R., Li, J., Wang, L., Galdran, A. et al. (2021). “Lndb challenge on automatic lung cancer patient management.” *Medical image analysis*, 70, 102027.
- Potghan, S., Rajamenakshi, R. and Bhise, A. (2018). “Multi-layer perceptron based lung tumor classification.” In *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, 499–502.
- Rahman, M. S., Shill, P. C. and Homayra, Z. (2019). “A new method for lung nodule detection using deep neural networks for CT images.” In *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, IEEE, 1–6.
- Rajan, J. R., Chelvan, A. C. and Duela, J. S. (2019). “Multi-class neural networks to predict lung cancer.” *Journal of medical systems*, 43(7), 211.
- Reeves, A. P., Xie, Y. and Liu, S. (2017). “Large-scale image region documentation for fully automated image biomarker algorithm development and evaluation.” *Journal of Medical Imaging*, 4(2), 024505.
- Rehman, M. Z., Nawi, N. M., Tanveer, A., Zafar, H., Munir, H. and Hassan, S. (2020). “Lungs cancer nodules detection from ct scan images with convolutional neural networks.” In *International Conference on Soft Computing and Data Mining*, Springer, 382–391.

- Robles, A. and Harris, C. (2017). “Lung cancer field cancerization: Implications for screening by low-dose computed tomography.” *JNCI: Journal of the National Cancer Institute*, 109.
- Ronneberger, O., Fischer, P. and Brox, T. (2015). “U-net: Convolutional networks for biomedical image segmentation.” In *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.
- Roy, R., Chakraborti, T. and Chowdhury, A. S. (2019). “A deep learning-shape driven level set synergism for pulmonary nodule segmentation.” *Pattern Recognition Letters*, 123, 31–38.
- Ruby, U. and Yendapalli, V. (2020). “Binary cross entropy with deep learning technique for image classification.” *Int. J. Adv. Trends Comput. Sci. Eng*, 9(10).
- Sahu, P., Yu, D., Dasari, M., Hou, F. and Qin, H. (2018). “A lightweight multi-section CNN for lung nodule classification and malignancy estimation.” *journal of biomedical and health informatics*, 23(3), 960–968.
- Salem, N., Malik, H. and Shams, A. (2019). “Medical image enhancement based on histogram algorithms.” *Procedia Computer Science*, 163, 300–311.
- Santos, A. M., de Carvalho Filho, A. O., Silva, A. C., de Paiva, A. C., Nunes, R. A. and Gattass, M. (2014). “Automatic detection of small lung nodules in 3d CT data using gaussian mixture models, tsallis entropy and svm.” *Engineering applications of artificial intelligence*, 36, 27–39.
- Setio, Traverso, A., De Bel, T., Berens, M. S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M. E., Geurts, B. et al. (2017a). “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge.” *Medical image analysis*, 42, 1–13.
- Setio, Traverso, A., De Bel, T., Berens, M. S., Van Den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M. E., Geurts, B. et al. (2017b). “Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge.” *Medical image analysis*, 42, 1–13.

REFERENCES

- 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.
- Shakeel, P. M., Burhanuddin, M. and Desa, M. I. (2020). “Automatic lung cancer detection from ct image using improved deep neural network and ensemble classifier.” *Neural Computing and Applications*, 1–14.
- Shaukat, F., Raja, G., Ashraf, R., Khalid, S., Ahmad, M. and Ali, A. (2019). “Artificial neural network based classification of lung nodules in ct images using intensity, shape and texture features.” *Journal of Ambient Intelligence and Humanized Computing*, 10(10), 4135–4149.
- Shaziya, H., Shyamala, K. and Zaheer, R. (2018). “Automatic lung segmentation on thoracic CT scans using u-net convolutional network.” In *International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 643–647.
- Shen, W., Zhou, M., Yang, F., Yang, C. and Tian, J. (2015). “Multi-scale convolutional neural networks for lung nodule classification.” In *Information Processing in Medical Imaging*, Springer, 588–599.
- Shi, H., Lu, J. and Zhou, Q. (2020). “A novel data augmentation method using style-based gan for robust pulmonary nodule segmentation.” In *2020 Chinese Control and Decision Conference (CCDC)*, IEEE, 2486–2491.
- Silva, G. L. F. d., Carvalho Filho, A. O. d., Silva, A. C., Paiva, A. C. d. and Gattass, M. (2016). “Taxonomic indexes for differentiating malignancy of lung nodules on ct images.” *Research on Biomedical Engineering*, 32(3), 263–272.
- Sinha, T. (2018). “Tumors: Benign and malignant.” *Cancer Therapy and Oncology*, 10(3), 555790.
- Song, Q., Zhao et al. (2017). “Using deep learning for classification of lung nodules on computed tomography images.” *Journal of healthcare engineering*, 1–7.
- Sori, W. J., Feng, J., Godana, A. W., Liu, S. and Gelmecha, D. J. (2021). “Dfd-net: lung cancer detection from denoised ct scan image using deep learning.” *Frontiers of Computer Science*, 15(2), 1–13.

- Sun, W., Zheng, B. and Qian, W. (2017). “Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis.” *Computers in biology and medicine*, 89, 530–539.
- Sundarambal, B., Subramanian, S. and Muthukumar, B. (2021). “A hybrid encoding strategy for classification of medical imaging modalities.” *Journal of Ambient Intelligence and Humanized Computing*, 12(6), 5853–5863.
- Surendar, P. B. (2021). “Diagnosis of lung cancer using hybrid deep neural network with adaptive sine cosine crow search algorithm.” *Journal of Computational Science*, 53, 101374.
- Sverzellati, N., Silva, M., Calareso, G., Galeone, C., Marchianò, A., Sestini, S., Sozzi, G. and Pastorino, U. (2016). “Low-dose computed tomography for lung cancer screening: comparison of performance between annual and biennial screen.” *European Radiology*, 26, 3821–3829.
- Sweetlin, J. D., Nehemiah, H. K. and Kannan, A. (2018). “Computer aided diagnosis of pulmonary hamartoma from CT scan images using ant colony optimization based feature selection.” *Alexandria engineering journal*, 57(3), 1557–1567.
- Tajbakhsh, N. and Suzuki, K. (2017). “Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs.” *Pattern recognition*, 63, 476–486.
- Tian, Q., Wu, Y., Ren, X. and Razmjoooy, N. (2021). “A new optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm.” *Biomedical Signal Processing and Control*, 68, 102761.
- Torrey, L. and Shavlik, J. (2010). “Transfer learning.” In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, IGI global, 242–264.
- Tran, G. S., Nghiem, T. P., Nguyen, V. T., Luong, C. M. and Burie, J.-C. (2019). “Improving accuracy of lung nodule classification using deep learning with focal loss.” *Journal of healthcare engineering*, 2019, 1–9.
- Wang, H., Zhao, T., Li, L. C., Pan, H., Liu, W., Gao, H., Han, F., Wang, Y., Qi, Y. and

REFERENCES

- Liang, Z. (2018). “A hybrid CNN feature model for pulmonary nodule malignancy risk differentiation.” *Journal of X-ray Science and Technology*, 26(2), 171–187.
- Wang, J., Wang, J., Wen, Y., Lu, H., Niu, T., Pan, J. and Qian, D. (2019). “Pulmonary nodule detection in volumetric chest ct scans using cnns-based nodule-size-adaptive detection and classification.” *IEEE Access*, 7, 46033–46044.
- Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., Dong, D., Gevaert, O. and Tian, J. (2017). “Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation.” *Medical image analysis*, 40, 172–183.
- Wang, Y., Liu, H. and Wang, X. (2022). “Pseudo color fusion of infrared and visible images based on the rattlesnake vision imaging system.” *Journal of Bionic Engineering*, 19(1), 209–223.
- Wu, B., Zhou, Z., Wang, J. and Wang, Y. (2018). “Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction.” In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, IEEE, 1109–1113.
- Wu, P., Xia, K. and Yu, H. (2016). “Relevance vector machine based pulmonary nodule classification.” *Journal of Medical Imaging and Health Informatics*, 6(1), 163–169.
- Xia, X., Xu, C. and Nan, B. (2017). “Inception-v3 for flower classification.” In *2017 2nd international conference on image, vision and computing (ICIVC)*, IEEE, 783–787.
- Xie, Y., Xia, Y., Zhang, J., Song, Y., Feng, D., Fulham, M. and Cai, W. (2018). “Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT.” *IEEE transactions on medical imaging*, 38(4), 991–1004.
- Xu, X., Wang, C., Guo, J., Yang, L., Bai, H., Li, W. and Yi, Z. (2020). “Deepln: a framework for automatic lung nodule detection using multi-resolution ct screening images.” *Knowledge-Based Systems*, 189, 105128.
- Yokota, K., Maeda, S., Kim, H., Tan, J. K., Ishikawa, S., Tachibana, R., Hirano, Y. and Kido, S. (2014). “Automatic detection of GGO regions on ct images in LIDC dataset based on statistical features.” In *Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*, IEEE, 1374–1377.

- Ypsilantis, P.-P. and Montana, G. (2016). “Recurrent convolutional networks for pulmonary nodule detection in CT imaging.” *arXiv preprint arXiv:1609.09143*.
- Zhang, G., Yang, Z., Gong, L., Jiang, S. and Wang, L. (2019). “Classification of benign and malignant lung nodules from CT images based on hybrid features.” *Physics in Medicine & Biology*, 125011.
- Zhao, X., Liu, L., Qi, S., Teng, Y., Li, J. and Qian, W. (2018). “Agile convolutional neural network for pulmonary nodule classification using CT images.” *International journal of computer assisted radiology and surgery*, 13(4), 585–595.

PUBLICATIONS

JOURNAL PAPERS

1. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2021). A Novel Receptive Field Regularized V-Net (RFR V-Net) and Nodule Classification Network (NCNet) for Lung Nodule Detection. Imaging Systems and Technology, Wiley. (DOI: <https://doi.org/10.1002/ima.22636>) (Impact Factor: 2.177)
2. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2022). Recent Advancements in Deep Learning based Lung Cancer Detection: A Systematic Review. Engineering Applications of Artificial Intelligence, Elsevier. (DOI: <https://doi.org/10.1016/j.engappai.2022.105490>) (Impact Factor: 8)
3. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2022). KAC SegNet: A Novel Kernel Based Active Contour Method for Lung Nodule Segmentation and Classification using Dense AlexNet Framework. International Journal of Information Technology & Decision Making, World Scientific. (DOI: <https://doi.org/10.1142/S0219622023500700>) (Impact Factor: 3.508)
4. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2022). A Light-weight Lung Nodule Detection and Classification System on Computed Tomography Scans. Soft Computing, Springer. **(Under Review)**
5. Shubham Dodia, Annappa B, Jeny Rajan and Mahesh Padukudru. (2022). Early Stage Detection of Lung Cancer using Artificial Intelligence: A Systematic

REFERENCES

Review, Challenges and Future Scope. Artificial Intelligence Review, Springer
(Under Review)

CONFERENCE PAPERS

1. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2022). A Novel Bi-level Lung Cancer Classification System on CT Scans. MIUA 2022: 26th UK Conference on Medical Image Understanding and Analysis, Springer. (https://dl.acm.org/doi/abs/10.1007/978-3-031-12053-4_43)
2. Shubham Dodia, Annappa B, and Mahesh Padukudru. (2021). A Novel Artificial Intelligence-Based Lung Nodule Segmentation and Classification System on CT Scans. CVIP 2021: 6th International Conference on Computer Vision & Image Processing, Springer. (https://link.springer.com/chapter/10.1007/978-3-031-11349-9_48)

BOOK CHAPTER

1. Shubham Dodia and Annappa B. (2020). Machine Learning based Detection and Classification of Lung Cancer. Computational Intelligence in Cancer Diagnosis: Progress and Challenges, Elsevier. (<https://www.elsevier.com/books/computational-intelligence-in-cancer-diagnosis/nayak/978-0-323-85240-1>)

BIODATA

Name: DODIA SHUBHAM JAGDISH

Date of Birth: 5th March, 1994

Gender: Male

Marital Status: Married

Father's Name: Jagdish A. Dodia

Mother's Name: Sushma J. Dodia

Address: Sawarkar Chowk,
Chaitanywadi,
Buldhana-443001
Maharashtra

E-mail: Shubham.dodia8@gmail.com

Mobile: +91-7020535093

Qualification: B.Tech in Computer Engineering, MIT AOE Pune, Pune University, 2016
M.Tech in Computer Science & Engineering, NIT Goa, Brain Computer Interface, 2018

Areas of Interest: Medical Image Processing, Deep Learning