Lung Cancer Detection Systems using Deep Learning Models

Thesis submitted by

NANDITA GAUTAM

DOCTOR OF PHILOSOPHY (Engineering)

Department of Computer Science and Engineering
Faculty Council of Engineering & Technology
Jadavpur University,
Kolkata-700032, India
2025

JADAVPUR UNIVERSITY
KOLKATA 700032, INDIA

Index No: 199/22/E

1. Title of the Thesis:

   Lung Cancer Detection Systems using Deep Learning Models

2. Name, Designation & Institute of the supervisor:

   Ram Sarkar
   Professor,
   Department of Computer Science and Engineering,
   Jadavpur University, Kolkata-700032

3. List of Publications:

   a) Journal

   i. Roy, A., Saha, P., Gautam, N. & Sarkar, R. Adaptive genetic algorithm based deep feature selector for cancer detection in lung histopathological images. *Scientific Reports* **15**, 4803 (2025). https://doi.org/10.1038/s41598-025-86362-8

   ii. Gautam, N., Ghosh, S. & Sarkar, R. CNN models aided with a meta-classifier for lung carcinoma classification using histopathological images. *Multimedia Tools and Applications* **84**, 28493–28517 (2025). https://doi.org/10.1007/s11042-024-20289-6

   iii. Gautam, N., Basu, A. & Sarkar, R. Lung cancer detection from thoracic CT scans using an ensemble of deep learning models. *Neural Computing and Applications* **36**, 2459–2477 (2024). https://doi.org/10.1007/s00521-023-09130-7

   iv. Hazra, M., Gautam, N., Dey, J. & Sarkar, R. A stacking ensemble method with multiple color models for lung cancer classification using histopathology images. *Multimedia Tools and Applications* (2025). https://doi.org/10.1007/s11042-025-20942-8

   v. Majumder, S., Gautam, N., Basu, A., Sau, A., Geem, Z. W., & Sarkar, R. MENet: A Mitscherlich function based ensemble of CNN models

to classify lung cancer using CT scans. *PLOS ONE* **19**(3), e0298527 (2024). https://doi.org/10.1371/journal.pone.0298527

b) Conference

i. Gautam, N., Basu, A., Kaplun, D., & Sarkar, R. An Ensemble of UNet Frameworks for Lung Nodule Segmentation. In: *APAMCS 2022, Lecture Notes in Networks and Systems, vol 702*. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-34127-4_44 (Self Presented)

ii. Gautam, N., Ghosh, S., & Sarkar, R. Lung CT Image Segmentation Using Double Adaptive Attention Aided TransUNet. In: *Computational Intelligence in Pattern Recognition (CIPR 2024), Lecture Notes in Networks and Systems, vol 1153*. Springer. https://doi.org/10.1007/978-981-97-8093-8_19 (Self Presented)

c) List of Presentations in Conference:

i. Gautam, N., Basu, A., Kaplun, D., & Sarkar, R. An Ensemble of UNet Frameworks for Lung Nodule Segmentation. In: *APAMCS 2022, Lecture Notes in Networks and Systems, vol 702*. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-34127-4_44 (Self Presented)

ii. Gautam, N., Ghosh, S., & Sarkar, R. Lung CT Image Segmentation Using Double Adaptive Attention Aided TransUNet. In: *Computational Intelligence in Pattern Recognition (CIPR 2024), Lecture Notes in Networks and Systems, vol 1153*. Springer. https://doi.org/10.1007/978-981-97-8093-8_19 (Self Presented)

PROFORMA – 1

"Statement of Originality"

I ..... NANDITA GAUTAM ..... registered on ..... 20/06/2022 ..... do hereby declare that this thesis entitled ... "Lung Cancer Detection Systems using Deep Learning Models" ... contains literature survey and original research work done by the undersigned candidate as part of Doctoral studies.

All information in this thesis have been obtained and presented in accordance with existing academic rules and ethical conduct. I declare that, as required by these rules and conduct, I have fully cited and referred all materials and results that are not original to this work.

I also declare that I have checked this thesis as per the "Policy on Anti Plagiarism, Jadavpur University, 2022", and the level of similarity as checked by iThenticate software is 8%.

Signature of Candidate: Nandita Gautam

Date: 11/11/25

Certified by Supervisor(s):
(Signature with date, seal)

1. Ram Sarkar 11/11/2025

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata-700032

PROFORMA - 2

CERTIFICATE FROM THE SUPERVISOR/S

This is to certify that the thesis entitled **"Lung Cancer Detection Systems using Deep Learning Models"** submitted by **NANDITA GAUTAM**, who got his/her name registered on **20/06/2022** for the award of Ph.D. (Engg.) degree of Jadavpur University is absolutely based upon his own work under the supervision of **Prof. RAM SARKAR (Dept. of C.S.E., JU)** and that neither his/her thesis nor any part of the thesis has been submitted for any degree/ diploma or any other academic award anywhere before.

1. Ram Sarkar 11/11/25

Signature of the Supervisor
and date with Office Seal

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata-700032

2. _____

Signature of the Supervisor
and date with Office Seal

Dedicated to my Grandmother

Acknowledgements

I am deeply indebted to many people who have contributed in making the completion of this thesis work possible. Firstly, I would like to express my earnest gratitude and respect to my supervisors Prof. Ram Sarkar for his relentless guidance, tremendous support, endless care, invaluable suggestion, and constant monitoring during the course of this work.

I would also like to express my sincere respect to ..Prof. Nirmalya Choudhhury.., Head, Department of Computer Science and Engineering, Jadavpur University (JU); ...Prof. Nandini Mukhopadhyay.., Former Head, Computer Science Engineering department, JU; ..Prof. Mahantapas Kundu.. Coordinator CMATER Laboratory, JU, for providing me constant inspiration, support and useful suggestions during the course of this thesis work. I am also thankful to other faculty and staff members of Department of Computer Science and Engineering, Jadavpur University, for their support during my course of work. I am grateful to ..Dr. Neelotpal Chakraborty.., ..Dr. Samir Malakar.. for their encouragement throughout my research work. I am also thankful to all my co-researchers ..Abhishek Basu.., ...Surya Majumder..., ...Arup Sau...., ...Manish Hazra..., ...Jaydip Dey.., ...Sohini Ghosh... for their sincere support during my research work.

I am indebted to my mentor and guide HG Dauji Kripa Das Prabhuji for his prayers and blessings that kept me motivated and inspired me in every way to complete this thesis. I owe my encompassing debt to my respected Father (Mr. .Ajay Gautam..), Mother (Mrs. .Madhuri Gautam..) Finally, I am thankful to the person who saw it all, who encouraged me to pursue it from the beginning, till the completion of this work and was there by my side to support through it all, my husband (Mr. Devajyoti Das Adhikary) without whom this thesis would have never been possible.

Date: 11.11.2025

Nandita Gautam

Place: Kolkata

(.....)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Lung cancer continues to be one of the most prevalent and fatal cancers worldwide, accounting for a significant proportion of cancer [1]. Despite advances in treatment, the prognosis remains poor, largely due to late-stage diagnosis. Many studies highlight the urgent need to improve early detection methods to enhance patient survival rates. The variability in lung cancer histopathology and the complexity of tumor morphology further complicate diagnosis. This global health challenge necessitates innovative approaches that can effectively support clinicians in detecting and characterizing lung cancer at earlier, more treatable stages. Imaging modalities such as Computed Tomography (CT) scans have become the gold standard for detecting lung nodules, offering high-resolution, three-dimensional visualization of lung tissue [2]. Techniques like convolutional neural networks (CNNs), recurrent networks, and transformer models have demonstrated superior performance in segmenting complex structures and classifying cancerous tissues compared to traditional methods. These models automatically learn hierarchical features directly from data, circumventing the need for manual feature engineering. Many recent studies have successfully applied deep learning to CT and histopathology images, demonstrating improved accuracy, reproducibility, and speed. The ability of deep learning to handle large-scale datasets and capture subtle image patterns presents an opportunity to enhance lung cancer diagnosis significantly. The literature shows that researchers are extensively experimenting to solve this problem. However, given their real-life societal impacts, their performances are still inadequate. Furthermore, the increasing technological advancements in this domain necessitate more such work to find a solution. With significant advancements in ML and DL over the last decade, several methods have been proposed that improve the accuracy of determining whether a lung cancer nodule is benign or malignant. In ML algorithms, need to extract features from the inputs in a handcrafted manner, and pass those features to the classifier for classification. On the other hand, DL based models like different CNN models extract the features from the input automatically, i.e., without any manual intervention [3]. DL algorithms, such as CNNs and recurrent neural networks (RNNs), have demonstrated exceptional capabilities in automatically learning hierarchical representations from raw imaging and genomic data. By analyzing massive datasets comprising radiological images, histopathological slides, CT scans, etc.,

DL models can detect hidden patterns, facilitating the early detection of lung cancer and accurate prediction of patient prognosis. The study by Talib et al. [4] presents two innovative models, TransSegNet for lung cancer segmentation and MinClassNet for classification, demonstrating significant accuracy improvements over existing methods. TransSegNet achieved a detection accuracy of 99.62% on the ACDC-LungHP dataset, while MinClassNet obtained 98.39% accuracy in classifying healthy and abnormal cells. An example of a Computer Aided Diagnosis (CAD) system available from literature, has been depicted in Figure 1.1.

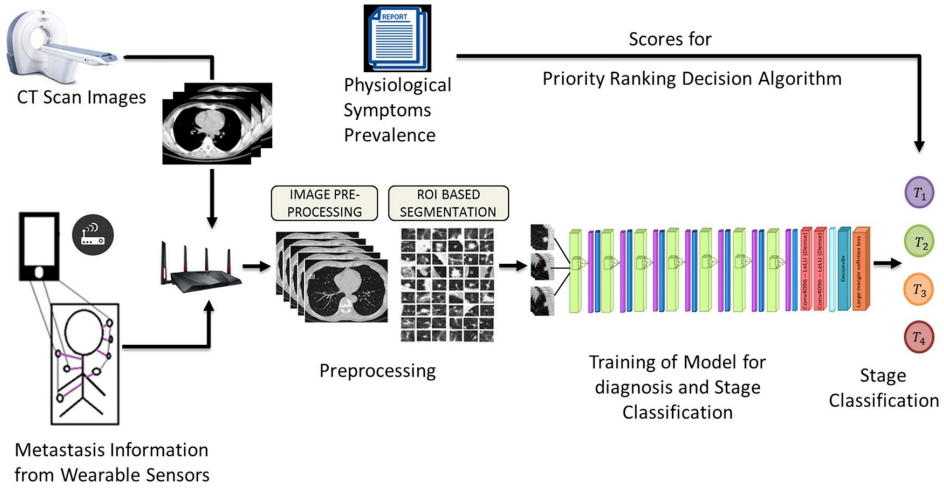


Figure 1.1: Computer-Assisted Decision Support System in Lung Cancer detection and stage classification

[5]

1.1 Origin of the Problem

One critical application of DL in lung cancer research is medical imaging analysis. The utilization of CNNs in automated feature extraction and classification of lung nodules on CT scans [6] is further explored. These models can assist radiologists in rapidly identifying suspicious lesions, distinguishing between malignant and benign nodules, and providing valuable insights into tumor characteristics such as size, shape, and texture. The present thesis work is primarily focused on the development of deep learning-based systems for automated detection, classification, and segmentation of lung cancer using medical imaging data. With the increasing availability of annotated radiological scans such as CT, CNNs and attention-based architectures have emerged as powerful tools in the analysis of thoracic abnormalities. Despite their promise, several fundamental and practical challenges persist that hinder their clinical deployment.

Lung cancer lesions exhibit considerable variability in terms of size, shape, texture,

and spatial location. This intrinsic heterogeneity complicates the design of generalizable deep learning models, particularly when trained on datasets that lack diversity in demographic, pathological, or imaging acquisition parameters. CNNs, while proficient at capturing local spatial features, often struggle with global context—an issue especially critical in tasks like tumor boundary detection and cancer staging. Attention mechanisms have been introduced in recent years to mitigate this limitation by enabling models to focus selectively on clinically relevant regions of interest, yet their integration with CNNs introduces additional architectural and computational complexity.

Furthermore, medical imaging data frequently suffer from low contrast, noise, and overlapping anatomical structures. This makes segmentation of tumors a highly non-trivial task. Deep learning models must not only localize suspicious regions accurately but also preserve semantic consistency across slices and modalities. In practice, these systems often face a trade-off between model complexity, inference time, and interpretability—each of which is critical in real-world diagnostic workflows.

In order to address these multifaceted challenges, this research aims to systematically investigate and design deep learning-based decision support frameworks that combine the spatial hierarchies learned by CNNs with the contextual reasoning power of attention modules. The goal is to enhance both classification accuracy and segmentation precision, thus contributing toward more reliable and interpretable lung cancer detection systems that can support early diagnosis and clinical decision-making.

1.2 Problem Definition and Research Objectives

1.2.1 Challenges in different Lung Image Modalities

Lung cancer imaging spans multiple modalities—each with its unique set of technical and diagnostic challenges that complicate the development of reliable automated systems. Among these, CT and histopathological images remain widely used for tumor localization and staging, while modalities such as positron emission tomography (PET), magnetic resonance imaging (MRI), and chest X-rays serve complementary roles in clinical decision-making. However, the inherent complexities of each imaging type demand advanced deep learning models that can adapt across spatial, spectral, and contextual variations.

In **CT imaging**, high inter-patient variability, irregular nodule morphology, and heterogeneous intensity profiles pose significant barriers to automated segmentation and classification. Lung nodules may vary in shape, texture, and boundary clarity,

often mimicking benign structures such as blood vessels or bronchi. Moreover, noise, motion artifacts, and slice-wise inconsistencies reduce signal clarity, requiring models to extract robust spatial features while preserving fine-grained anatomical details.

Histopathological images introduce challenges of a different nature. Inconsistencies in staining protocols, variations in slide preparation, and differences in scanner calibration lead to significant color and texture variability across samples. The complex spatial arrangement of cells, overlapping nuclei, and high intra-class similarities make it difficult to distinguish between benign and malignant tissues. Additionally, the sheer resolution and size of whole-slide images (WSIs) necessitate patch-wise analysis, which increases computational demands and introduces additional context-related ambiguity.

Positron Emission Tomography (PET), though valuable for identifying metabolically active regions, suffers from lower spatial resolution and higher noise levels compared to structural imaging modalities. This limits the accuracy of boundary delineation and requires fusion with CT or MRI to improve anatomical localization. Deep learning models must handle this multimodal fusion effectively, often requiring architectures that can learn both complementary and hierarchical features from multiple modalities.

Magnetic Resonance Imaging (MRI), while offering excellent soft-tissue contrast, is less frequently used for lung imaging due to respiratory motion artifacts and susceptibility to low signal in aerated lung regions. Nonetheless, for specific lung cancer subtypes or metastasis evaluation, MRI data must be integrated into diagnostic pipelines. Variability across imaging sequences (T1, T2, diffusion-weighted, etc.) adds further complexity, necessitating modality-aware feature extraction mechanisms.

Chest X-rays, though low-cost and widely available, provide only 2D projections of complex 3D anatomical structures. This can obscure nodules and limit the visibility of early-stage tumors. The presence of overlapping tissues and variability in patient positioning introduces additional ambiguity, requiring models with high sensitivity and contextual reasoning capabilities.

Across all these modalities, the common challenge lies in developing deep learning systems that can generalize across diverse imaging characteristics, institutions, and patient populations. Effective solutions demand robust architectures capable of capturing both local and global contextual information—often achieved through multi-scale CNNs, attention mechanisms, and modality-specific pre-processing strategies. Furthermore, training these models requires access to large, annotated datasets, which is often limited in the medical domain. This makes semi-supervised learning, domain adaptation, and data augmentation essential components of the overall system design.

This thesis utilizes the CT Scan Images and Histopathological images for developing deep learning based application for lung cancer detection.

1.2.2 Need for Automated Lung Cancer Detection Systems

Manual interpretation of lung cancer images is labor-intensive and subject to inter- and intra-observer variability, leading to inconsistent diagnostic outcomes. Automated segmentation of lung nodules in CT scans can provide precise tumor boundaries, facilitating accurate volume estimation and treatment planning. Similarly, automated classification of histopathology images can assist pathologists in identifying cancer subtypes and grades more efficiently. Deep learning-based automated systems promise to reduce workload, standardize diagnoses, and enable scalable screening, ultimately benefiting clinical workflows and patient care.

1.2.3 Objectives of the Thesis

This thesis aims to address the limitations of current lung cancer imaging analysis by developing advanced deep learning models for segmentation and classification tasks. The specific objectives include:

- Designing a robust segmentation framework for CT images that effectively captures variable lung nodule characteristics.
- Developing accurate classification models for both CT and histopathology images to differentiate between benign and malignant tissues.
- Exploring ensemble methods that combine multiple models to improve prediction reliability and diagnostic accuracy.
- Develop attention based methods for precise segmentation of lung nodules in order to assist accurate lung cancer classification.

1.3 Research Gaps and Motivation

1.3.1 Segmentation of Lung Regions using Novel UNet Variants

The segmentation of lung regions, particularly the identification of small or irregular nodules, remains one of the most pressing challenges in medical image analysis. Traditional UNet-based segmentation models, although widely adopted due to their en-

coder–decoder architecture and ability to capture spatial context, often fail to provide reliable accuracy when faced with the heterogeneous characteristics of lung nodules. Many nodules in computed tomography (CT) and other imaging modalities exhibit irregular morphologies, highly variable sizes, and heterogeneous intensity patterns that make their boundaries indistinct. Low-contrast nodules, in particular, present significant difficulty as their appearance blends with surrounding tissues, making the segmentation task prone to false negatives or inaccurate delineation. Another major issue is the scarcity of accurately annotated regions of interest (ROIs) in real-world datasets. Obtaining ground-truth annotations requires the expertise of radiologists and pathologists, which is both time-consuming and resource-intensive. Consequently, available annotated datasets are often limited in size, further complicating the training of deep learning models. This scarcity leads to overfitting of conventional UNet models and restricts their ability to generalize to unseen data. Furthermore, conventional UNet approaches are not inherently designed to capture long-range dependencies within an image, which are critical for distinguishing between subtle variations in tissue structures. These limitations necessitate the development of novel UNet variants capable of handling irregular morphologies, heterogeneous intensities, and low-contrast nodules with greater robustness. Enhanced architectures, such as recurrent residual UNets (R2UNets), attention-based UNets, and Transformer-augmented UNets (TransUNets), offer potential solutions by improving feature representation, learning global contextual dependencies, and integrating adaptive attention mechanisms. By addressing the shortcomings of conventional segmentation methods, these advanced models aim to improve the accurate extraction of lung regions and nodules, providing a more reliable foundation for downstream diagnostic tasks.

1.3.2 Classification of Lung Cancer Images Using Novel Deep Learning Algorithms

While segmentation serves as the first step in lung image analysis, accurate classification of lung cancer images is equally critical for effective diagnosis and treatment planning. However, most existing studies in this domain rely heavily on deep convolutional neural networks (CNNs) without incorporating systematic feature selection strategies. This over-reliance on deep CNNs often leads to models that process large volumes of redundant or less informative features, thereby reducing efficiency and sometimes impairing classification accuracy. The domain of histopathology-based lung cancer classification, in particular, has not extensively explored feature selection methods that could isolate the most discriminative features and enhance the learn-

ing process. To overcome this limitation, adaptive genetic algorithm–based feature selection strategies can be employed, enabling the optimization of feature subsets to improve detection accuracy while reducing computational complexity.

Another limitation of current approaches is the tendency to depend on single deep learning models, which may suffer from inherent biases or limited generalization capability, especially when applied to diverse datasets collected across different institutions and imaging devices. Ensemble learning approaches, such as stacking, boosting, and the use of metaclassifiers, provide an opportunity to mitigate these issues by integrating the strengths of multiple models. Moreover, unique ensemble methods, such as Mitscherlich-based ensembles, can further enhance robustness and stability in classification results. A persistent challenge in medical image datasets is class imbalance, where the number of malignant cases is often significantly lower or higher than benign cases, or where nodules are underrepresented compared to non-nodules. Models trained on such imbalanced datasets tend to be biased toward the majority class, leading to poor sensitivity in detecting critical but underrepresented categories. This imbalance not only undermines the reliability of predictions but also poses a serious risk in clinical applications where missing a malignant case could have severe consequences. Therefore, techniques for addressing class imbalance, such as data augmentation, cost-sensitive learning, and synthetic data generation, must be integrated into the classification pipeline.

In summary, the proposed research addresses critical gaps in both lung region segmentation and lung cancer image classification by advancing UNet architectures for challenging segmentation tasks, introducing adaptive feature selection techniques, leveraging ensemble learning, addressing class imbalance, and embedding Explainable AI into the framework. Together, these contributions aim to push the boundaries of accuracy, robustness, and interpretability in lung cancer image analysis.

1.4 Limitations in Existing Techniques

Although numerous methods have been proposed for lung nodule segmentation and classification, many rely on handcrafted features or conventional machine learning techniques that do not fully capture complex image patterns. Some deep learning models lack robustness when faced with diverse imaging conditions or fail to leverage contextual and multi-scale information effectively. Furthermore, existing models often struggle with generalizability across datasets from different scanners or patient populations. The current literature also shows limited exploration of combining imaging modalities or integrating domain knowledge into learning frameworks, leaving room

for improvement in diagnostic accuracy and clinical applicability.

1.5 Motivation

The challenges identified motivate leveraging sophisticated deep learning architectures that incorporate residual connections, attention mechanisms, and recurrent units to better model spatial and contextual information within images. Ensemble learning, by aggregating outputs from multiple models, can mitigate individual model weaknesses and enhance overall prediction performance. This approach has demonstrated success in related medical imaging tasks but remains underexplored in lung cancer segmentation and classification. Therefore, this thesis focuses on designing novel architectures and ensemble frameworks that can robustly handle heterogeneous data and improve the reliability of automated lung cancer diagnosis.

1.6 Scope of the Thesis

The thesis focuses on two major aspects of lung image analysis: segmentation and classification, each employing innovative deep learning techniques to address existing limitations in accuracy, robustness, and interpretability.

1.6.1 Segmentation of Lung Regions using Novel UNet Variants

For the segmentation of lung regions, novel UNet variants are developed and evaluated, with the primary goal of enhancing the detection and delineation of small, irregular, and low-contrast nodules that are often missed by traditional methods. The study explores advanced models such as R2UNet, which introduces recurrent residual connections to capture complex contextual features, and TransUNet, which leverages adaptive attention mechanisms to integrate both local and global dependencies within medical images. Furthermore, ensemble UNet strategies are proposed to combine the strengths of multiple architectures, leading to more stable and precise segmentation outcomes across diverse imaging conditions.

1.6.2 Classification of Lung Cancer Images Using Novel Deep Learning Algorithms

On the classification front, the research introduces adaptive genetic algorithm–based feature selection methods designed to optimize feature subsets and improve the accuracy of histopathology-based lung cancer detection. In addition, ensemble learning strategies, including stacking approaches, Mitscherlich-based ensembles, and meta-classifiers, are incorporated to boost model robustness and generalization capability. Special emphasis is placed on handling class imbalance, a common challenge in medical datasets, through tailored techniques that ensure balanced performance across all cancer subtypes. The study also explores the use of multi-color models to better capture the intricate patterns present in histopathology images, thereby improving classification accuracy. Finally, to ensure the reliability and transparency of predictions, Explainable AI techniques such as Grad-CAM are integrated, providing clear visualizations of decision-making processes and enhancing clinical trust in the developed models. Together, these contributions aim to advance the state of the art in both lung region segmentation and cancer classification, paving the way for more accurate, reliable, and interpretable computer-aided diagnosis systems.

1.7 Organization of the Thesis

The scope of this thesis encompasses the development and evaluation of deep learning–based methodologies for the segmentation and classification of lung cancer images. The work primarily focuses on CT scan data, complemented by histopathology images, and explores ensemble techniques to enhance diagnostic robustness. The thesis is organized into five chapters, each addressing a specific part of the research problem.

Chapter 1 introduces the research problem, highlighting the origin of the problem, challenges associated with different lung image modalities, and the need for automated lung cancer detection systems. It also presents the objectives, research gaps, motivation, and an overview of the thesis scope.

Chapter 2 provides a detailed literature survey on segmentation and classification methods for lung cancer detection. It reviews conventional, machine learning, and deep learning–based segmentation approaches, followed by traditional and modern deep learning–based classification techniques.

Chapter 3 focuses on segmentation of lung regions using novel UNet variants. It introduces and evaluates methods such as R2UNet with spatial-channel squeeze excitation blocks and BUS-UNet++, along with ensemble frameworks for improved

lung nodule segmentation. Experimental results and discussions are also presented.

Chapter 4 addresses classification of lung cancer images using novel deep learning algorithms. It explores multiple ensemble and optimization strategies, including decision score fusion, fuzzy ensembles, and adaptive genetic algorithm–based feature selection. The chapter also provides detailed experimental results, parameter tuning, and a discussion of comparative strengths, underlying mechanisms, and limitations.

Chapter 5 concludes the thesis by summarizing the contributions and presenting future research directions. It outlines the limitations of the current work and provides recommendations for extending segmentation and classification techniques for more comprehensive and reliable lung cancer diagnosis.

Chapter 2

Literature Survey

2.1 Segmentation Methods Available

Accurate pulmonary nodule segmentation, a primary requirement in CAD systems, helps in the risk assessment of lung cancer. One of the most prevalent and severe diseases affecting people is cancer. If found and identified in its early stages, it can be effectively treated. Older cells that get damaged are often replaced by new cells in a biological process. When this mechanism fails and injured cells are not replaced, cancer develops. These aberrant cells have the potential to spread to other bodily organs. In this section, discuss some of the related research work on lung nodule segmentation, found in the literature. Rukhmini et al. [7] proposed a synergistic combination of DL and shape-driven level sets for automated and accurate lung nodule segmentation. They used a coarse-to-fine solution, with a deep FCN used for coarse segmentation. They achieved an accuracy of 96.48% on the LUNA16 dataset. In order to diagnose patients with asbestosis in segmented CT images, Hyung et al. [8] worked on to develop an algorithm that integrates lung segmentation with deep learning models.
Gao et al. [9] presented a fully automatic method for segmenting lung from three-dimensional images from the LIDC dataset. Chen Zhou et al. [10] developed an automatic segmentation model by integrating (3D) V-Net [11] and spatial transform network (STN) to segment pulmonary parenchyma in CT images and analyze texture features from the segmented pulmonary parenchyma regions to assist the radiologist in COVID-19 diagnosis. They obtained a (dice similarity coefficient) DSC of 95.89% towards the segmentation performance. Tao Zhou et al. [12] presented a parallel UNet network architecture called APUNet. In order to extract the characteristics of various modalities and for extracting feature maps of various scales of decoding path, they utilized a multi-scale feature aggregation block. Their proposed model, APUNet, achieved a DSC 96.86% and recall 97.53%. A parallel DL technique that integrates with hybrid attention mechanism modules and the DenseNet structure was proposed by Hexuan et al. [13]. They implemented a spatial attention mechanism and a channel attention mechanism combining it as a hybrid attention. Their method improved the precision and integrity of tumour picture segmentation and obtained a DSC of 96.89%. Worku et al. [14] proposed two path convolution neural network (DFD-Net). Initially they

removed the noise during the preprocessing step using a residual learning denoising model (DR-Net) and further passed it to a two-path convolutional neural network with the purpose of segmenting lung CT scan images. Other effective CNN-based designs have recently been developed and are employed in a variety of medical image segmentation applications, including UNet [15], SegNet [16], Residual UNet [17], R2UNet [18], Concurrent Spatial and Channel Squeeze & Excitation UNet (scSEUNet) [19] which is an upgraded version of Squeeze & Excitation UNet (SEUNet). These neural network based architectures are most widely used methods for semantic segmentation.

Ronneberger et al. [20] trained UNet architecture using a small sample of the LUNA-16 images. Their network was divided into two parts based on its structure: convolutional encoding and decoding units. In both parts of the network, basic convolution operations were performed, followed by Rectified Linear Activation Unit (ReLU) activation. The UNet approach has various benefits for segmentation tasks, starting with the simultaneous use of global location and context [21]. A number of variations have been put forward to apply the straightforward UNet design to complicated datasets. Siddique et al. [22] reviewed multiple UNet variants for image segmentation tasks over different types of datasets. The balanced skip-connections encoder-decoder network, UNet [10], introduced a new approach to medical image segmentation. Its variants have been explicitly used in very different biomedical image segmentation processes. The goal of the remodelled skip connections in UNet was to close the semantic gap between the feature maps of the encoder and decoder layers. UNet employs regular skip connections, while UNet++ proposed by Zhou et al. [23] used nested and dense skip connections. However, Claw UNet [24] added attention modules and residual blocks to a new network architecture that learned more intricate features to increase the accuracy of segmentation. Hence, this improves the process of filtering the details from the images adequately. The improvement in segmentation results led to the proposition of many other U-shaped models like Res-UNet by Xiao et al. [25], Attention-UNet proposed by Oktay et al. [26], Claw UNet proposed by Chang et al. [27] and H-DenseUNet proposed by Li et al. [28] and many more. The introduction of different variants made a significant enhancement in the study of medical image segmentation.

Transformer and self-attention: Recurrent neural networks (RNNs) were the popular choices for machine translation and language modeling until Transformers based models came into existence. RNN models showed a sequential nature. Hence, computation depended on the symbol positions of the input and output sequences. While RNNs process data sequentially, they cannot parallelize well. Transformers [29] ran multiple sequences in parallel. The encoder in the encoder-decoder structure consists

of layers of a self-attention mechanism and a fully connected feed-forward network followed by a normalization layer. Positional embeddings added to the input embeddings at the bottom of the encoder and decoder stacks provided information about the positions of tokens in the sequence. Transformers superseded the recurrent layers in the encoder-decoder architecture with multiple self-attention heads working in parallel gaining transformers a widespread range of applications in Natural Language Processing (NLP). The attention score in the self-attention mechanism was derived from the same sequence using a query, a set of key-value pairs, and an activation function. The score helps to determine the magnitude of different elements in the sequence for a given token. Self-attention was a central component of the Transformer architecture and allowed the model to capture both local and global dependencies within a sequence. In the realm of computer vision, Hu et al. [30] proposed a novel attention mechanism called Squeeze-and-Excitation to carry out feature re-calibration, in aspects that were relevant for a certain visual task that was prioritized while the rest of the features were considered less significant.

The adaptive attention span model of Transformers proposed by Sukhbaatar et al. [31] allowed the relevant attention span for each token to be calculated dynamically. This led to a linear or sublinear complexity, making Transformers more scalable for long sequences. This reduced computational and memory costs. Hence, Adaptive attention mechanisms overcame some limitations of the standard, fixed-length self-attention mechanism. The adaptive attention mechanism proposed by Shaoqing et al. [32] has improved performance by fusing different kinds of semantic details at the voxel level by learning the context features. Adaptive attention mechanisms have also been used for image captioning by Li et al. [33]. The model used the visual sentinel, introduced as an auxiliary signal to guide the model's attention, to adjust its attention over time adaptively. This enabled it to look at the most relevant parts of the image for generating each word in the caption, rather than having a fixed attention mechanism.

Recently, Transformers attained attention for medical image segmentation. The Transformer mechanism on the encoder part showed better segmentation results than typical UNet and its other variants [34]. A variant proposed by Chen et al. [35] is called TransUNet, which is a sequence-to-sequence framework with the self-attention mechanism. TransUNet overcame the long-range dependency problem of UNet and the limited localization abilities of Transformers due to inadequate obtained low-level features.

Many researchers have proposed state-of-the-art models to achieve better results in image segmentation. The LIDC-IDRI dataset, a prevalent choice for image segmentation, contains 1,018 lesion annotations from four experienced thoracic radiologists. Kumar et al. [36] proposed UNet for segmenting the lung nodules and finding the best opti-

mizer obtained a dice similarity coefficient (DSC) of 82.05% with the Adam optimizer. For small nodules and juxta pleural nodules, the dual branch Residual network proposed by Cao et al. [37] achieved a DSC of 82.74%. A dual-branch CNN architecture based on ResNet can classify each voxel with the extracted multi-view and multi-scale features. The model incorporated the traditional intensity features into the CNN architecture to achieve a performance improvement in the nodule segmentation model. The residual blocks reduce the computation time. The skip connections in the model allow data to flow from the bottom to the top layers without loss. This model has shown improvement in segmentation with a DSC of 87.2%. Swim Transformer by Liu et al. [38] created a transformer whose representation is calculated with shifted windows. Because self-attention is only computed inside each local window, the suggested Swin Transformer creates hierarchical feature maps by combining picture patches in deeper layers.

2.1.1 Conventional Segmentation Methods

Classical segmentation approaches refer to traditional image processing and computer vision techniques used for segmenting objects or regions of interest in an image. Before the development of ML and DL, traditional segmentation methods were often utilized in various applications. But they are still useful in some situations, especially when working with simpler pictures or in combination them with other methods for more difficult segmentation tasks.

Region-based Segmentation

Region-based lung segmentation aims to identify and define lung areas in medical images. This process involves segmentation and merging techniques to divide the lung area into smaller regions, which are later combined to obtain a complete lung segmentation. Different criteria, such as pixel intensity or texture features, are used for region separation, while fusion is based on region size, shape, and feature similarity. Several studies propose segmentation algorithms, including the fast march method and region growing methods, achieving high dice similarity coefficients (DSC) on medical imaging datasets (e.g., DSC of 97.33% [39], 97.49% [40]).

In the lung segmentation task, the graph cut technique [41] [42] [43] is used by many researchers to split the lungs from the surrounding areas of the picture by taking into account the intensity values and pixel locations. In this method, an image is represented as a graph where nodes represent pixels and edges represent similarities between pixels. Using an optimization function, the graph is split into two non-overlapping sets

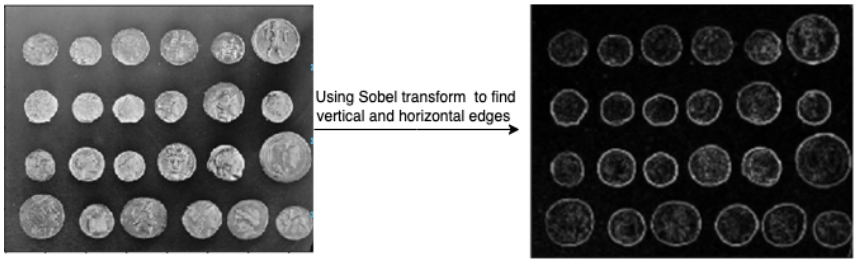


Figure 2.1: Block diagram of a region-based segmentation method using the Sobel transformation

of nodes, foreground and background. The optimization function determines the minimum energy cutoff that separates the foreground from the background.
Region growing is a simple and intuitive method that can produce accurate results in homogeneous regions with well-defined boundaries. Still, it may fail in cases with noise, inhomogeneity, or weak edges. Sometimes the proper delineation of complicated or irregular boundaries may be difficult for region-based segmentation techniques.

Edge-based Segmentation

Edge-based lung segmentation is an image segmentation technique that divides regions of interest using edges or contours, relying on identifying significant discontinuities in the image [44]. It is especially useful when the contrast between lung tissue and surrounding structures is insufficient. Edge detection algorithms like Canny, Sobel, Prewitt, and Roberts-Cross are commonly used to find edges based on image intensity fluctuations [45–47]. The Laplace Gaussian (LoG) operator is also utilized to refine edges and features in edge-based quadratic segmentation methods [48]. The final segmentation is obtained through thresholding or clustering methods on the resulting image after edge detection and refinement.

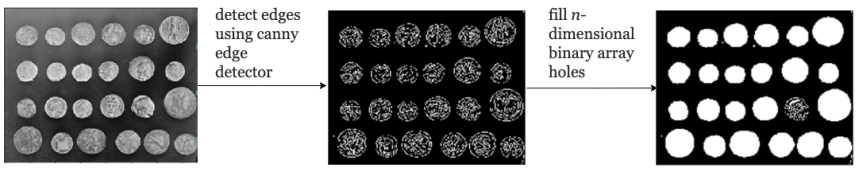


Figure 2.2: Block diagram of an edge-based segmentation method using the Canny edge detector

Figure 2.2 shows how canny edge detectors detects the edges and fills the array holes. Edge-based quadratic methods segment the lung into its constituent parts by

defining boundaries between lung regions and surrounding tissue [49] [50]. Saien et al. [51] use LoG kernel to find the center points of nodules candidate from a 3D blob detector. The multiscale normalized LoG filtering method for detecting pulmonary nodules on CT of the whole lung presented in Fotin et al. [52] provides high sensitivity for both solid and non-solid pulmonary nodules.

Edge-based segmentation is sensitive to noise and it can lead to false edges and incorrect segmentation. In some cases it is challenging to accurately detect and defining the boundaries when the images have weak or incomplete edges and can lead to under-segmentation. Some edge-based segmentation techniques can require a lot of computing power. Because of this, they may not be as useful in real-time or contexts with limited resources particularly when dealing with large or high-resolution images.

Threshold-based Segmentation

Threshold-based segmentation is a method that separates objects from the background based on a threshold value, often determined by pixel intensity. For lung segmentation, the threshold creates a binary mask that isolates the lung region from the rest of the image. Different types of threshold-based methods exist [53], varying in how the threshold is determined. Khan et al. [54] introduce a semi-automatic method that targets the infected regions in lung CT images, while Christie et al. [55] emphasize a fast, semi-automatic technique requiring minimal user interaction. Zhou et al. [56] present a fully automated approach using an enhanced thresholding algorithm based on a texture-aware active contour model. The threshold can be manually set or automatically determined through methods like Otsu's or adaptive thresholding techniques [57]. Adaptive techniques adjust dynamically, as seen in applications to lung cancer X-rays [58, 59].

Threshold-based segmentation can misclassify pixels due to sensitivity to image contrast and lighting variations. When the intensity values of foreground and background pixels overlap or lighting is uneven, it becomes difficult to find a single threshold value that effectively separates objects. This can result in inaccurate segmentation, particularly when dealing with complex or overlapping objects in the image [60].

Watershed Segmentation

Watershed algorithm is applied to obtain a binary mask of the lung area for the segmentation of lungs. A watershed algorithm used to isolate the lung regions and remove small regions, including the binary mask, due to noise or other factors. Various watershed segmentation techniques [61] are available for lung segmentation.

Marker-controlled watershed segmentation is a common method for lung segmentation, relying on markers to guide the process. Studies have combined it with image processing filters like Gabor filter [62] or Laplacian filter [63] for lung cancer detection and staging. Basha et al. [64] use marker-controlled watershed segmentation to resolve over-segmentation. Morphological watershed segmentation, based on morphological gradient and distance transform [65] [66], is another approach used with various lung imaging modalities to enhance segmentation accuracy.

In watershed segmentation, if the image has noise or small variations in intensity across the image, it can lead to over-segmentation, i.e., the algorithm produces too many small regions that are not meaningful objects. Another limitation of watershed segmentation is under-segmentation. Suppose the markers are not placed correctly, or the image has weak boundaries between objects. In that case, the algorithm fails to segment objects with complex shapes or connected to neighboring objects, leading to under-segmentation.

2.1.2 Machine Learning based Image Segmentation

The shift from classical segmentation to ML approaches in medical image segmentation enables more accurate and efficient methods, with supervised learning, especially, playing a vital role in lung cancer segmentation by leveraging labeled data. ML algorithms are commonly used in lung segmentation due to their ability to quickly and automatically solve segmentation problems.

SVM's ability to perform non-linear separation using kernel functions is crucial for accurately identifying complex boundaries and irregular tumor shapes in medical image segmentation. Sathish et al. [67] propose a twin SVM for efficient tumor segmentation on CT images, while Shen et al. [68] combine bidirectional chain coding and SVM to improve lung border smoothing and minimize over-segmentation. SVM can be effective for lung segmentation, but its efficiency may vary depending on the dataset and segmentation task. Ferl et al. [69] combine 3D UNet with SVM classifier to discriminate between lung tumors and non-tumor regions based on shape attributes, texture characteristics, or other relevant information.

K-Nearest Neighbors

KNN, a supervised ML method, is often used for classification and regression, but can be incorporated as a post-processing step in lung tumor segmentation tasks, as shown in Afify et al. [70]. Peng et al. [71] propose a segmentation method that combines KNN with Deep Belief Network for lung segmentation on chest radiographs. The

type of data, segmentation pipeline, and application objectives influence how KNN is employed for lung segmentation, utilizing features extracted from lung images to classify pixels or voxels into different classes based on feature similarity [72, 73].

Decision Tree

A decision tree is created by recursively dividing the data into classes or segments depending on the values of the input features [74]. By recursively partitioning the feature space [75], decision trees can identify relevant image features and their critical thresholds for accurate segmentation. This capability reduces the burden of manual feature selection and simplifies the segmentation process. Yoo et al. [76] propose a deep learning-based decision-tree classifier on the X-ray images. When training a decision tree model on a labeled dataset of lung pictures, input parameters like pixel intensities, texture features, shape features, or other relevant image qualities are employed in segmentation. The automated decision tree-based learning approach for advanced segmentation is founded on [77]. For the purpose of segmenting lung carcinoma, Markel et al. [78] propose a mixed decision tree with KNN classifiers as nodes and combinations of three characteristics and achieve DSC of 60.7%.

Random Forest

Random Forest is an ML algorithm suitable for lung segmentation tasks, like in Liu et al. [79] and [80] with improved accuracy using multi-scale edge detection and texture-based segmentation. Paing et al. [81] use optimized random forests for pulmonary nodule detection in early lung cancer diagnosis. Li et al. [82] suggest a combination of random forest and random walk for interactive pulmonary lobe segmentation. Wang et al. [83] propose GDyBaORF for dynamically balanced training data in interactive image segmentation.

Metaheuristic Approaches

Metaheuristic approaches are optimization algorithms designed to solve complex problems where traditional techniques may struggle. These methods are inspired by natural or abstract phenomena and explore large solution spaces to find good solutions. Lan et al. [84] proposed Group Theoretic Particle Swarm Optimization (GT-PSO) for multi-level thresholding optimization in segmentation, achieving high IoU and Sensitivity on the IQ-OTH/NCCD dataset. Ziyad et al. [85] focused on lung nodule segmentation using LDCT images from the LIDC dataset, using K-means clustering (GAK-means) and a genetic algorithm enhanced Fuzzy c-means clustering (GAFCM) algorithms

to achieve accurate results. [86] introduced Salp Shuffled Shepherd Optimization Algorithm (SSSOA) based GAN for lung nodule segmentation, outperforming existing methods. Wang et al. [87] improved segmentation accuracy using adaptive particle swarm optimization. Parveen et al. [88] proposed FSFCM for automatic lung nodule segmentation, showing better results compared to other algorithms. Zhao et al. [89] used self-generating neural networks and PSO with a Jaccard Index of 88.9%. [90] combined FC and evolutionary computation for pulmonary nodule segmentation on the LIDC dataset. A summary of metaheuristic-based methods for lung nodule segmentation is presented in Table 2.1.

Table 2.1: Metaheuristic based approaches, datasets and results for lung segmentation

Authors	**Year**	**Dataset**	**Method**	**Results**
Lan et al. [84]	2023	IQ-OTH/NCCD	GT-PSO	IoU = 0.902 and Sensitivity = 0.945
Ziyad et al. [85]	2022	LIDC	GAK-means and GAFCM	sensitivity = 99.47 % and accuracy = 99.26%
Jain et al. [86]	2021	LIDC	SSSOA-based GAN	Accuracy = 93.8%, DSC = 79.8%, Jaccard Similarity = 80.2%
Parveen et al. [88]	2015	Private	FSFCM	Accuracy = 97.5%
Zhao et al. [89]	2015	Private	Self-generating neural networks and PSO	Jaccard Index = 88.9%
Badura et al. [90]	2014	LIDC	FC and Evolutionary Computation	Jaccard index = 0.69 using the MRFC-OB method and 0.69 using the Mask variant. TPR100 = 99.9%, TPR50 = 70.2%, and FDR0+ = 91%.

2.1.3 Deep Learning based Segmentation Techniques

This section provides an overview of various DL architectures, such as CNNs, RNN, and their variants, that have been successfully applied for lung image segmentation tasks. discuss the recent advancements in DL-based segmentation, advantages, and challenges highlighting their potential for solving complex segmentation problems and advancing the state-of-the-art in the field.

Convolutional Neural Network

CNNs have become a popular method in the fields of computer vision and pattern recognition due to their ability to share local weights and learn features from images. Pre-trained models like VGG, ResNet, AlexNet, and DenseNet serve as references for designing deep CNNs. In medical imaging, networks like FCN, SNNs, and ResNet have been proposed. Zhao et al. [91] suggest an effective end-to-end FCN for lung segmentation in CT scans, while Chen et al. [92] use a dense deep CNN with multi-instance loss and conditional adversarial loss. Cao et al. [93] propose a dual branch residual network for segmenting lung nodules with multi-view and multi-scale characteristics. Hwang et al. [94] use a multiple-stage training program based on CNN with artous convolution on the JSRT dataset. Faster RCNN [95] is proposed for lung segmentation, and Bu et al. [96] propose a RIDER-based detection method with an IOU of 68.71%. Tang et al. [97] propose a method for nodule detection and segmentation refinement on the LIDC dataset.

To separate lung nodules from diverse CT scans, Wang et al. [98] suggest a data-driven model called the Central Focused Convolutional Neural Network (CF-CNN). This method uses a central pooling layer that retains a lot of knowledge about the voxel patch center, followed by a multi-scale patch learning approach. The researchers also propose an MV-CNN [99] model to capture a wide range of node detection capabilities in axial, coronal, and sagittal CT scans.

Limitations : Drawbacks of CNNs in lung segmentation include those related to data augmentation, overfitting, class mismatch, limited ability to capture spatial connectivity, and interpretability issues. Especially for complex lung structures, these variables can reduce the accuracy and reliability of segmentation results. However, CNNs remain an effective method for lung segmentation and are often applied in academic and medical applications.

UNet and Its Variants

UNet is a convolutional network [100] designed to achieve accurate and fast segmentation of medical images. It is capable of working effectively with limited training data by utilizing extensive data augmentation techniques. UNet has also demonstrated successful applications in 3D imaging segmentation. The network consists a contracting path, a expansive path and skip connections with 23 convolutional layers. The overall shape of the network resembles a large U. In recent research experiments, several variants of the UNet have been developed, each with its own unique features and improvements. Table 2.2 shows the different variants of unet based models, datasets

and their results in lung segmentation. In UNet, the encoder captures the contextual information by downsampling the input image, while the decoder upsamples the feature maps to produce a pixel-wise segmentation map. There are various types of UNet variants which are as follows:

Residual UNet: Enhanced versions of UNet are frequently used for lung nodule segmentation because numerous researchers have suggested considerable improvements on it like Residual UNet [18]. This variant incorporates residual connections, inspired by the ResNet architecture, into the UNet model. The residual connections enable the direct propagation of information from earlier layers to later layers, helping to mitigate the vanishing gradient problem and facilitating the training of deeper networks. The Residual UNet combines the benefits of skip connections in UNet and residual connections in ResNet, resulting in improved gradient flow and feature reuse.

Dense UNet: The Dense UNet incorporates dense connections, as introduced in the DenseNet architecture, into the UNet framework. Dense connections establish direct connections between all layers, allowing for better information flow and feature reuse throughout the network. This facilitates the efficient propagation of gradients during training and helps alleviate the issue of information loss across layers. Dense UNet, introduces dense connections within the encoder and decoder networks, allowing for direct and dense feature reuse across different depths. This promotes better information flow and gradient propagation, leading to more efficient learning and improved segmentation accuracy.

Attention-Guided UNet: Another popular variant is the Attention UNet, which introduces attention mechanisms to selectively highlight relevant features during the encoding and decoding process. This enables the network to focus on important regions and suppress noise, resulting in more precise segmentation maps. Additionally, the Recurrent UNet incorporates recurrent connections, such as LSTMs or GRUs, to enable the model to capture temporal dependencies in sequential data, such as videos or time-series images. This variant is beneficial for tasks that require understanding the temporal context, such as video segmentation. This variant integrates attention mechanisms with UNet to guide the network's focus on important regions. It uses both spatial [101] and channel-wise attention modules [102] to capture fine-grained details and emphasize informative features, enabling more accurate segmentation results. The Attention-Guided UNet selectively attends to relevant regions while suppressing irrelevant or noisy information, enhancing the model's discriminative power.

Attention-based models for lung segmentation in medical imaging involve using attention mechanisms [103] to focus the model on relevant parts of the image during the segmentation process.

Zhou et al. [12] propose a hybrid attention method that added two UNet (UNet A and UNet B) models parallelly. The attention system processes incoming features from the UNet A and UNet B encoder paths and forwards them to the UNet A and UNet B decoder paths. This attention model has inspired Gu et al. [104] to make significant use of multiple attentions in a CNN architecture and suggest a thorough attention-based CNN (CA-Net) for more precise and understandable medical picture segmentation that is simultaneously aware of the most crucial spatial locations, channels, and scales. A residual attention UNet is put up by Chen et al. [105] to learn robust feature representations by including aggregated residual transformations and soft attention mechanisms.

UNet++: UNet++ [106] is an extension of the original UNet architecture that incorporates a nested and densely connected pathway. It features a series of nested UNet sub-modules at multiple scales, enabling the network to capture multi-scale contextual information. By aggregating features from different levels of the network hierarchy, UNet++ enhances the representation power and improves the localization accuracy of the segmentation model. UNet++ enhances the original architecture by incorporating skip connections at different depths of the encoder and decoder networks, allowing for better information flow and gradient propagation. This helps in capturing both local and global contextual information, leading to improved segmentation accuracy.

Hybrid UNet: Hybrid UNet [107] combines the strengths of different architectures, such as UNet and DeepLab, by incorporating dilated convolutions in the encoder network to capture multi-scale contextual information. This helps in segmenting objects of different sizes accurately.

Graph Convolution Networks

Lung segmentation can be approached as a graphical signal processing problem, where each pixel in a lung image is viewed as a node in the graph, and the relationship between adjacent nodes is modeled as edges. GCNs can then be used to process the structure of the graph and extract useful features for lung segmentation. Several studies have investigated the use of GCNs for lung segmentation with promising results. A recent study [124] proposes a GCN-based lung segmentation method that achieves state-of-the-art performance on publicly available datasets. The method uses GCNs to model the relationship between lung pixels and extracts features from the graph structure to

Table 2.2: UNet based methods, datasets and results for lung segmentation

Authors	Year	Dataset	Method	Results
Gautam et al. [108]	2023	LIDC	Ensemble of BUS-UNet and UNet++	DSC = 0.95, IoU = 0.84, Accuracy = 97%
Xu et al. [109]	2023	DSB2017	Improved UNet 3+	Accuracy=98.72%, DSC=97.22%, IOU=94.76%, Sensitivity=96.60%,Precision=98.06%
Wu et al. [110]	2023	LIDC-IDRI and CT dataset	RAD-UNet	IOU=87.76% and 88.13%, F1-Score=93.56% and 93.72%
Wang et al. [111]	2023	LIDC	CA-UNet and Swin Transformer	DSC = 89.86%, Sensitivity = 92.44%, IOU = 82.42%, Precision = 89.07%
Zhang et al. [112]	2022	NSCLC-Radiomics and LIDC	M-SegSEUNet-CRF	DSC = 0.85±0.071, Sensitivity of 0.82±0.10, IoU = 0.74±0.10
Zhang et al. [113]	2022	TCIA, LIDC and Private	Improved 3D dense connected UNet	DSC = 0.83 for TCIA, LIDC and 0.81 for the private dataset
Yang et al. [114]	2022	LUNA16, MC and JSRT	Shape-guided deep residual network	DSC=97.9%, 97.0% and 97.2%
Salama et al. [115]	2021	NLST	UNet	DSC=98.75%, Sensitivity=99.09%, Accuracy=99.90%
Kumar et al. [116]	2021	LIDC	UNet	Accuracy = 0.98, DSC = 0.82 and IOU = 0.75
Chen et al. [117]	2021	Private	MAUNet	DSC = 95%
Chen et al. [118]	2021	LUNA16	UNet with dilated convolution	DSC = 0.97, IoU = 0.96
Xiao et al. [119]	2020	LUNA16	3D-UNet and Res2Net	DSC = 95.30%, and Sensitivity = 99.1%
Jalali et al. [120]	2020	LIDC-IDRI	Modified UNet	DSC = 97.31%
Rocha et al. [121]	2020	LIDC	UNet and SegUNet	DSC = 0.830 for UNet and 0.82 for SegUNet
Usman et al. [122]	2020	LIDC	Deep Residual UNet	DSC = 87.5%
Tong et al. [123]	2018	LUNA 16	Improved UNet	DSC = 0.736

classify the pixels.

HybridGNet [125] is a neural network that combines standard CNNs with GCNs for the task of anatomical segmentation in chest x-ray images. is proposed in the year 2021. Recently, GCNs[[124], [126], [127], [128]] have been developed to model the relationships among different data points in a graph-structured data. By encoding both graph structure and node features, GCNs can capture more informative representations of the data and make more accurate predictions. These experiments suggest that GCNs are a powerful tool for working with graph-structured data and can be applied to a wide range of real-world problems.

For segmenting lung images GNNs have some challenges, such as limited interpretability, computational complexity, graph construction, limited training data, and sensitivity to hyperparameters. However, GCN is still a promising approach for lung segmentation and may address some limitations of traditional CNN-based methods. Further research is needed to address these limitations and explore the full potential of GCNs in lung segmentation.

Encoder Decoder Based Models

The traditional structure of an auto-encoder (AE) consists of an encoder and a decoder. The encoder maps the input vector to a hidden representation, which is usually a lower-dimensional space, through deterministic mapping. The decoder, on the other hand, maps the hidden representation back to the original input space in order to reconstruct the input. The main idea behind this approach is that by forcing the network to reconstruct the input, it will learn to extract useful features from the input and compress them into the hidden representation. Many variants of AE have been proposed since the first AEs is presented in the 1980s for the use of dimensionality reduction. Some popular variants are Sparse Auto-Encoders (SAEs) [129], Denoising Auto-Encoders (DAEs) [130], Contractive Auto-Encoders (CAEs) [131], Convolutional Auto-Encoders (CAEs) [132], Variational Auto-Encoders (VAEs) [133], LSTM Auto-Encoders [134], and Adversarial Auto-Encoders (AAEs) [135].

All of these modifications enhance the original AE model and increase its adaptability, giving researchers more options to tailor it to specific problems. Ullah et al. [136] propose a dual encoder-decoder CNN for anatomical structure segmentation in chest radiographs, achieving a DSC of 97.67%. Luo et al. [137] introduce an encoder-decoder architecture utilizing shadow mapping. The shadow mapping encoder uses deterministic mapping to translate the input into a concealed representation, which the decoder then maps back to the original input space for reconstruction.

Another method for lung segmentation is based on multi-scale and pyramid network models [138]. These models often employ a multi-scale architecture that analyses the input image at different scales, such as at the original resolution and also at lesser resolutions. This enables the model to accurately segment the lungs by capturing both fine and coarse information in the image. A particular kind of multi-scale architecture called pyramid networks employs a pyramidal structure to interpret input images at various stages, with coarser scales feeding into finer scales. This helps the model transfer information from the coarser scales to the finer scales, which enhances the segmentation's accuracy. Ibrahim et aL. [139] suggest a combination of VAE and UNet on JSRT and MC dataset.

In most cases, CNNs serve as the framework for these models, which send the input image through numerous convolutional layers in order to extract features. Then, using a pyramid-like structure, the features are cycled through several scales to capture both fine and coarse information in the image. The final segmentation mask is created by running the features through a decoder network. Multi-scale and pyramid network-based models have been shown to be effective in lung segmentation, particularly in CT images. They are able to effectively leverage the multi-scale information present in the images and improve the accuracy of the segmentation.

DeepLab and Artous Convolutional Models

Atrous convolution is a type of convolution operation used in deep learning models such as CNN. It extends the standard convolution operation by adding "holes" or "gaps" between the values in the convolution kernel. These gaps or dilation rates allow the convolution operation to have a larger receptive field, capturing more contextual information from the input feature map. The Atrous Convolution-based CNN (ATCNN) [140] framework utilizes a technique called atrous convolution, also known as dilated convolution, to capture multi-scale features. Block diagram of ATCNN is shown in Figure 2.3. Atrous convolution allows the model to increase the receptive field of each convolutional filter without increasing the number of parameters, which helps the model to capture larger context and more detailed information in the images. The dilated convolution has been widely used in various computer vision tasks, especially in semantic segmentation [141] [142] [143] [144], where the goal is to classify each pixel in the image. By using dilated convolution, the network can learn multi-scale features, which helps in better understanding the context of the image and improving the segmentation accuracy. However, the simple dilated convolution in CNNs may lead to grid effects, which might significantly impair performance. A popular solution

to this issue is to employ dilated convolutions on many scales, such as by combining dilated and non-dilated convolutions or by utilizing atrous convolution [145]. Utilizing various pooling scales concurrently, another strategy is to employ pyramid pooling modules [146], which enable the network to efficiently gather information about the global context. By combining both multi-scale and atrous convolution [147], it has been shown to achieve good performance in semantic segmentation tasks. The equation of atrous convolution is taken from [148] is shown in Equation 2.1.

$$y[i] = \sum_{k=1}^{K} x[i + r \cdot k]w[k]. \tag{2.1}$$

where y[i] represents the output at position, x[i+rk] denotes the input signal at position i+rk, with r being the rate, and w[k] is the filter weight at position k. In this equation, the rate parameter r determines the stride at which the input data is sampled. When r = 1, it corresponds to normal convolution, and when r > 1, it indicates atrous convolution.

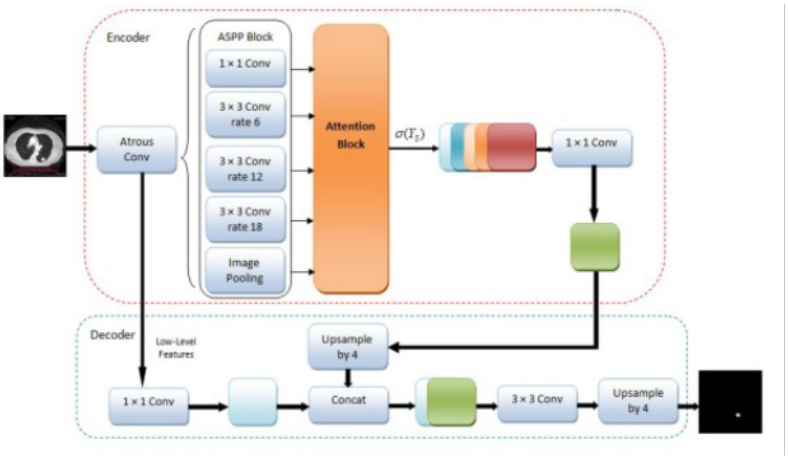


Figure 2.3: Block diagram of ATCNN for segmentation method [140]

In the lung segmentation task, DeepLab models show good results due to their ability to gather information on many details. The DeepLab model can extract fine-grained data and global content from lung images by creating precise segmentation boundaries. This is done using extended convolutions and ASPP. In order to segment lungs quickly and accurately, [149] uses an end-to-end Deeplabv3+ architecture combined with dilated convolution and artous spatial pyramid pooling for automated segmentation of the lungs and achieve DSC of 96.85% in JSRT dataset and 94.19% in MC dataset. They use ASPP to collect multi-scale contextual information and achieve a DSC of 97.20%. Chang et al. [150] compare a 3-class approach and 2-class approach on NIH chest X-ray dataset for semantic segmentation of lung cancer using

Deeplabv3+ architecture and achieve 94:9% mIoU and 92% mIoU respectively.

There are many methods where both DeepLab and dilated convolutions are working together [151]. Zhao et al. [152] propose an UNet-based dual attention strategy and hybrid dilated convolutions to create bigger receptive fields, to the model decoder, which improves the decoding process. By extracting multi-scale characteristics from the HRCT images, Halder et al. [140] propose ATCNN architecture that can segment and describe lung nodules.

Overall, although performing well on a variety of semantic segmentation tasks, DeepLab models still have issues and restrictions that need to be resolved to increase their generalizability and usefulness.

Recurrent Neural Networks

RNNs are a type of DL neural network that are particularly well-suited for sequential data analysis, such as time series data or sequential image data which also leads to the application in segmentation. The RNN design integrates recurrent connections into the CNN framework, often in the form of Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) layers.

Naeem Abid et al. [153] proposed an innovative approach called Multi-view Convolutional Recurrent Neural Network (MV-CRecNet). This method leverages shape, size, and cross-slice variations to improve the accuracy of identifying lung cancer nodules from CT scans. The utilization of multiple views as inputs to the model ensures better generalization and the learning of robust features. The model's ability to process multiple perspectives enables it to extract more comprehensive and generalized features, thereby improving its overall performance. Lung segmentation process can be enhanced combined with UNet and LSTM [154] or GRU [155]. For example, Alom et al. [18] provide a better UNet architecture with the same number of network parameters and recurrent residual convolutional layers that can give better results for medical image segmentation, called RUNet and R2UNet. Jalali et al. [156] used Bidirectional Convolutional LSTM as a concatenator in UNet architecture, and in the encoder part, they use pre-trained ResNet-34 network, and finally they reach the DSC of 97.31%. Another Bidirectional ConvLSTM is proposed by [157]. For the improvement of lung cancer, Gunjan et al. [158] propose an RNN with the Grey Wolf Optimization (GWO) algorithm, which increases the chance of better accuracy. Different RNN-based methods, datasets used in this method, and their corresponding results are shown in Table 2.3.

Table 2.3: RNN-based methods, datasets and results for lung nodule image segmentation.

Authors	Year	Dataset	Method	Results
Agnes et al. [154]	2022	LUNA16	Pyramid dilated convolutional LSTM	Accuracy=0.96±0.01, Precision=0.97±0.03, Recall=0.95±0.02, Accuracy=0.96±0.02, F1-Score=0.96±0.02
Xuan et al. [155]	2022	MSD, Private	Convolutional bi-directional GRU and Spatial enhanced attention	MSD Dataset: DSC=78.4%, IOU=70.3%, HD=34.429mm ; Private Dataset: DSC=68%, IOU=56%, HD=49.645mm
Kanipriya et al. [159]	2022	LIDC-IDRI	ICSA-optimized hybrid LSTM-CNN	Accuracy=98.231%, DSC=91.16±0.11, HD=4.34±11.04
Jalali et al. [156]	2021	LIDC-IDRI	ResBCDUNet	DSC=97.15%, Accuracy=97.58%, Precision=99.12%, Recall=97.01%
Xu et al. [160]	2020	NIH14	CNN-Attention-LSTM network	AUC, Precision, Recall, F-1 Score=85.4%
Asadi et al. [157]	2020	DSB2017	UNet, Squeeze and Excitation block, Bi-directional ConvLSTM	IOU=99.72%, Sensitivity=99.04%, F1 Score=99.04%
Gao et al. [161]	2020	NLST	Distanced LSTM	Accuracy=78.96%, AUC=82.55%, F1-Score=70.85%, Recall=61.61%

Generative Models and Adversarial Training

In the subject of semantic segmentation, Generative Adversarial Networks (GANs) have been used. In these models, a generator network is trained to produce accurate segmentation masks, and a discriminator network is taught to differentiate between genuine and created masks. GAN architecture is introduced by Goodfellow et al. [162] in the year 2014. In lung segmentation, GANs can be used to generate synthetic data that can be used to improve the accuracy and robustness of segmentation models. The generated simulated images can be used to augment the lung segmentation model training dataset after GAN training. Therefore, segmentation models can achieve

more diverse feature sets and perform better on real datasets. GANs can also be used to create synthetic images to test and validate segmentation models. These images can then be used to assess the reliability of the segmentation model. An overview of the GAN architecture is shown in Figure 2.4. Equation 1 is the general equation of the GAN architecture where the generator network "G" generates synthetic samples from random noise "z", while the discriminator network "D" distinguishes between the generated samples "G(z)" and real samples "Y.

$$\min_{G} \max_{D} V(G, D) = \mathbb{E}_y[\log D(y)] + \mathbb{E}_z[\log(1 - D(G(z))] \tag{1}$$

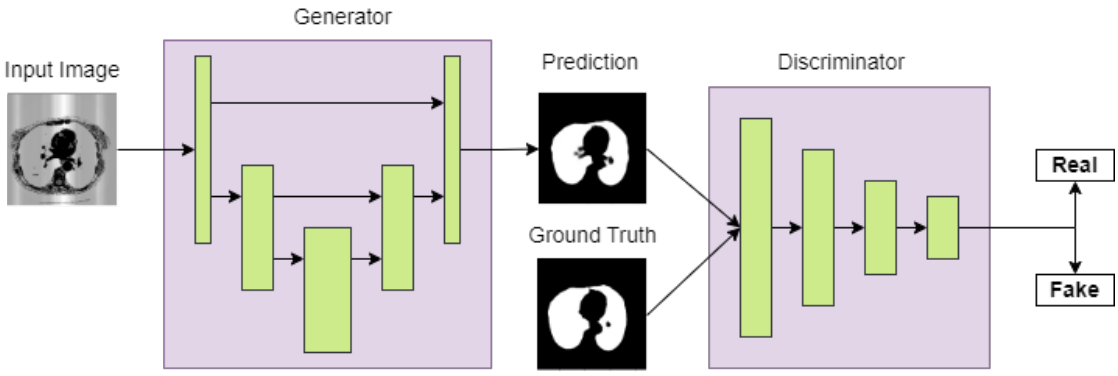


Figure 2.4: The UNet based GAN architecture for the segmentation of lungs on CT images

Several studies have shown that GAN can effectively improve the accuracy of lung segmentation. For example, a study published in 2019, which uses GANs [163] to generate synthetic lung CT images and achieves peak performance on public datasets. Another study published in 2021 uses GANs [164] to generate synthetic lung CT images and shows that training with synthetic images improved the accuracy of a segmentation model compared to using only real pictures. GANs are often used to create accurate data by learning a mapping from one region to another. In Munawar et al. [165], the generator of a GAN learns to create a segmentation mask of a specific CXR input. Using an adversarial loss metric, the discriminator refreshes the generator by distinguishing the false mask from the ground truth. The objective is to generate an input chest-X ray mask as realistic as possible compared to the ground truth mask. There are several types of GANs models used in lung segmentation. The most commonly used GAN architecture cGAN is used by Pawar et al. [166]. Tyagi et al. [167] also use cGAN to overcome the problem of overfitting and bad performance by learning the data distribution. Nishio et al. [168] propose a technique based on

transfer learning from learned models and using GANs. A GAN-based strategy for lung nodule separation is described by Supiksha et al. [169] that uses SSSOA (Salp Shuffled Shepherd Optimization Algorithm) to optimize the network and is able to obtain an accuracy of 0.9387 and dice coefficient of 0.7986. Overall, GANs are powerful in improving the accuracy and robustness of lung segmentation models, and their efficiency can be further improved by combining them with other techniques such as attention mechanisms, transfer learning, and ensemble methods.

Semi-supervised Approaches for Lung Image Segmentation

Semi-supervised learning has gained significant attention in the field of lung cancer research as a promising approach to leverage both labeled and unlabeled patient data for improving predictive models. Traditional supervised learning methods often suffer from limited availability of labeled data, as obtaining high-quality annotations for large-scale datasets can be time-consuming and costly. However, the field of lung cancer research is rich in unlabeled patient data, such as medical imaging scans, genomic profiles, and clinical records. Semi-supervised learning techniques aim to harness this vast unlabeled data to augment the limited labeled data, thereby enhancing the performance and generalizability of machine learning models.

Cui et al. [170] propose a semi-supervised deep learning framework called semi-supervised deep linear discriminant analysis to address the challenge of limited labeled data in histopathology image classification. The framework combines deep neural networks (DNNs) and graph-based approaches to leverage semantic information from labeled and unlabeled images simultaneously. By replacing the DNN loss function with linear discriminant analysis, features minimizing intra-class distance and maximizing inter-class distance are obtained to construct a robust graph Laplacian. The proposed framework outperformed state-of-the-art methods in classification tasks on skeletal muscle and lung cancer images, showcasing its effectiveness in leveraging unlabeled data for improved performance. Gao et al. [171] propose a classifier method with backbone feature extraction and weight control step. The backbone feature extractor is shared between the two task-specific classifiers, allowing for capturing the interaction between the CRD (Cancer Risk Detection) and subtyping tasks. The weight control mechanism was to ensure the sequential relationship between the tasks and facilitate error back-propagation from the subtyping task to the CRD task within the multi-task learning (MTL) framework. The framework was trained in a semi-supervised setting using a minimal point-based annotation strategy. Extensive experiments on The Cancer Genome Atlas (TCGA) datasets with various cancer types were performed

to demonstrate the effectiveness of the proposed framework in terms of accuracy and generalization.

Table 2.4: Semi-supervised learning methods, datasets and results for lung nodule segmentation

Authors	Year	Dataset	Method	Results
Gao et al. [171]	2023	TCGA	MTL framework	-
Xui et al. [172]	2023	LIDC-IDRI	Reverse GAN	Accuracy=93.21%
Fredriksen et al. [173]	2022	MSD, NSCLC	End-to-end automated tumour using pseudo annotation	DSC=71% and 53.89%, Recall=100% and 84.56%, Precision=83.33% and 60.44%
Shi et al. [174]	2022	Clinical Dataset	SDTL framework	Accuracy=88.30%
Liu et al. [175]	2021	Clinical and Genomic data	Random Survival forests	DSC=94%
Cui et al. [170]	2018	LC25000	DNN	-
Khosravan et al. [175]	2018	LUNA-16	3D deep multi-task CNN	DSC=91%

2.2 Lung Cancer Classification Methods

2.2.1 Traditional Machine Learning based Approaches

Accurate analysis of lung cancer disease (LCD) is an essential process to determine the appropriate treatment for lung malignancy sufferers. The scarcity of publicly available data is one of the limitations in this domain. In this section, review some research articles that used traditional machine-learning approaches for LCD detection.

In the research conducted by [176], a novel approach was introduced to classify lung nodules and non-nodules in CT images using texture features. The methodology involved three distinct strategies for texture measurement: rose diagrams (RD), artificial crawlers (ACs), and a hybrid model that combined ACs and RD. To differentiate between nodules and non-nodules during the classification process, the authors employed a support vector machine (SVM) classifier with a radial basis kernel. The study utilized the Lung Image Database Consortium (LIDC-IDRI) image database and reported promising results with a mean specificity of 94.78%, mean sensitivity of 91.86%, and mean accuracy of 94.30%.

Nevertheless, it was essential to acknowledge a notable limitation of the study, which lay in the comparison with previous methods. Due to differences in databases and test scenarios, the comparison was only approximate, making it challenging to draw precise conclusions. The study could only be directly compared to the work of [177], revealing that the proposed methodology achieved a lower specificity score. This indicated a relatively lesser ability to accurately classify non-nodule cases when compared to that specific previous study.

[178] presented a methodology for automated assistance in medical examinations using a nature-inspired algorithm for lung segmentation in chest radiographs. The proposed method helps to detect points in X-ray images, where tissues differ from regular ones, and therefore it needed an expert examination from the doctor. The results showed that the proposed methodology gave positive results both for the detection of unhealthy tissues and healthy tissues and can potentially improve the efficiency of examinations in pulmonary clinics.

[179] proposed a methodology for analyzing lung lesions using temporal evaluation, which helps diagnose indeterminate lesions during treatment. They used the modified quality threshold clustering technique to assign voxels to clusters and evaluated lesion changes by analyzing voxel movement over time. Statistical features were extracted to differentiate between benign and malignant lesions. With two databases of pulmonary lesions, they achieved a 98.41% accuracy in identifying lung lesions, showing the potential of their method in diagnosing indeterminate cases.

2.2.2 Deep Learning based Approaches

Unlike machine learning algorithms, deep learning methods are capable of performing end-to-end classification as they automatically extract translationally invariant features. CNNs are widely preferred for image classification as they are highly effective at extracting features from an image through the convolution of input images and kernels. When compared to standard machine learning approaches, the accuracy of several prominent CNNs such as InceptionV3 [180] and GoogleNet [181] for object recognition and picture categorization is quite high. Hence, CNN-based models are widely used by researchers for different complex image classification tasks. Additionally, transfer learning-based models solve the data scarcity problem in biomedical image classification tasks, wherein knowledge gained from a large dataset like ImageNet [182] is used to fine-tune the model on a small dataset under consideration. In this section, review some research articles using deep learning approaches on LCD detection.

For the classification of lung cancer, Shahid Mehmood et al. [29] suggested an

image contrast enhancement-based method incorporating class selective image processing (CSIP), which makes use of an AlexNet model that has already been trained. The Lung and Colon Cancer Histopathological Images (LC25000) dataset was utilized. The accuracy of the suggested model was 89.9%; the precision of the lung-aca, lung-n, lung-scc, and colon-aca were 98.8%, 99.7%, 67.4%, 99.9%, and 96%, respectively. The recall percentages for lung-aca, lung-n, lung-scc, colon-aca, and lung-n are 99.1%, 99.9%, 97.1%, 99.8%, and 53.3%, in that order. Masud et al. [34] used the LC25000 dataset and image improvement techniques to improve certain features, extract meaningful information, and minimize noise in images, making them more appropriate for learning. 2D Fourier and 2D Wavelet features were extracted by them as part of the feature extraction process, which is followed by the development of feature sets. They then passed the generated feature set into a single channel CNN that has a batch normalisation layer with a dropout rate of 0.3, two max pooling layers, and three interleaved 2D convolution layers. The maximum testing accuracy was 96.33%, and the highest training accuracy was 98.91%.

Hage et al. [183] implemented lung cancer classification on lung histopathology image dataset using XGBoost model and achieved an accuracy of 99%. S. Mangal et al. [31] used the LC25000 dataset and a single channel CNN composed of a convolution layer, a pooling layer, and a flatten layer to turn the output from the preceding layer into a 1D tensor. A fully connected layer was then inserted into the flattened tensor. On data related to lung cancer, the model produced training accuracy of 97.9216%, validation accuracy of 97.8987%. The CNN, utilized by B.K. Hatuwal et al. [24] to classify and recognize images, were built using a linear stack of layers. Convolutional layers with fully linked layers, max pooling, and kernel filters were applied to training and test images. The given object was classified using the softmax function. employed a neural network with three hidden layers, one input layer, and one fully connected layer. Each convolutional layer employed a (3, 3) kernel matrix with the activation function ReLU(x) = max (0, x). To reduce the computing parameters in the following convolution layer, a max pooling size of (2, 2) was used. The model's dropout value was set to 0.1. The model has a training accuracy of 96.11% and a validation accuracy of 97.20% in the final epoch.

Another study used a CNN to classify lung cancer on CT images. The CNN was trained on a dataset of over 1,000 CT images and was able to accurately classify lung cancer with a 93.3% accuracy. Deepak et al. [184] used a DL algorithm called deep belief network (DBN) to classify lung cancer on chest radiographs. The DBN outperformed traditional methods of manual analysis in classifying lung cancer, achieving an accuracy of 94.5%. DL algorithms, in addition to image classification, have been

used to predict the aggressiveness of lung cancer. Another study used a DL algorithm known as a recurrent neural network (RNN) to analyse gene expression data and predict the aggressiveness of lung cancer disease.

Previous research in the field of lung cancer classification has explored various approaches leveraging advanced machine learning techniques. Notably, Safiyari and Javidan [185] proposed a method for predicting lung cancer survivability using ensemble learning methods. Their study demonstrated the effectiveness of ensemble techniques in improving the accuracy of survivability predictions, highlighting the potential for personalized treatment strategies in lung cancer management. Cai et al. [186] introduced a classification framework for lung cancer utilizing ensemble-based feature selection and machine learning methods. Their research focused on identifying relevant features from complex datasets extracted from lung cancer patients, aiming to enhance the accuracy and interpretability of classification models. The study contributes to advancing our understanding of lung cancer biomarkers and offers valuable insights for clinicians in diagnosing and prognosticating this critical disease. Raza et al. [187] proposed "Lung-EffNet," a unique approach for lung cancer classification using the EfficientNet architecture and CT-scan pictures. Their findings indicated that transfer learning using EfficientNet can accurately distinguish between malignant and non-cancerous regions in lung tissue. The study demonstrates that deep learning approaches may accurately detect malignant spots in lung tissue, leading to improved diagnostic tools for lung cancer. Humayun et al. [188] presented a transfer learning approach with a CNN for the classification of lung carcinoma. Their research showcased the effectiveness of transfer learning in leveraging pre-trained neural network architectures for accurate classification of lung cancer from medical imaging data. The study underscores the importance of leveraging pre-trained models and transfer learning techniques to address the challenges of limited data availability in medical image analysis.

Walker et al. [189] combined a DL model known as a CNN with radiomics to improve the detection of lung cancer. The results showed that combining radiomics and DL improved lung cancer detection accuracy by about 10%. Some review studies [190] [191] [192] [193] summarize recent research on the application of DL techniques to the detection of lung cancer. The researchers conclude that DL has the potential to improve lung cancer detection accuracy, but for that more research is needed to optimize the performance of DL models for this task. Tyagi et al. [194] suggested a multi-level strategy for lung cancer stage classification utilizing 3D dense CNN with concurrent squeeze-and-excitation module on a publically available lung CT image dataset, achieving 97% classification accuracy.Basra et al. [195] conducted their ex-

periment with histopathological pictures. To identify lung cancer, they developed a hybrid model using CNNs with machine learning methods such as support vector classifier (SVC), random forest, and Xgboost. Their strategy attained an overall accuracy of 99.13%.

Gopi Kasinathan et al. [196] proposed a Cloud-based Lung Tumor Detector and Stage Classifier (Cloud-LTDSC) as a hybrid technique. The employed dataset is the LIDC-IDRI dataset which consists of CT or Positron Emission Tomography (PET) scan images. The model achieved 97% accuracy. Halder et al. [35] put forward an approach where the LIDC-IDRI dataset was used to train a 2-pathway morphology-based CNN (2PMorphCNN). The proposed model attained accuracy of 96.10Abraham et al. [197] used three 2D-CNN models (AlexNet, VGG16, and SilNet) to classify lung nodules and improve early lung cancer prediction networks.Based on the findings of the inference, a new network model was then created. The aforementioned investigation revealed that while the 2D-CNN model offers the benefits of quick computation and low network complexity, it overlooks some spatial information. This was primarily because most current CNN-based techniques employ a 2D model, which cannot capture the spatial information between slices, and CT scans are 3D pictures.

Shen et al. [198] proposed a hierarchical learning framework called multiscale CNN (MCNN) in order to account for the heterogeneity of lung nodules. Their framework used discriminatory features extracted from alternating stacked layers to classify types of lung cancer for the lung histopathology image dataset. This network provided excellent tolerance to noisy input in addition to improving classification accuracy. It is important to note that the MCNN's architecture, with alternating stacks of convolutional and max-pooling layers, made the process of feature extraction very time-consuming. They were able to obtain an accuracy of 92% with this model. Another technique for classification of lung cancer as benign or malignant was suggested by Gonzalez and Ponomaryvo [199]. They performed segmentation of the images on lung histopathology images followed by classification. The shape features such as area, eccentricity, fractal dimension, and textual information were extracted using the thresholding technique and morphological operations. Finally, they passed the features to a support vector machine (SVM) for training and classification. This set of experiments was done in order to determine the nodule's malignant and non-cancerous status. Their experiment showed a significant reduction in the number of false positives.

Most of the methods found in this field have used the dataset in RGB color format [200]. In our proposed model, extracting features from various color models provide complementary information about the input. Further details are provided in the following sections of the paper. The novelty that this work brings up is to get a better

prediction model for lung cancer using histopathology images. In doing so, apply a three-way feature extraction step, where the same CNN architecture (i.e., the NasNet-Large model) is fed with different color-transformed input images of lungs. Features obtained by each model are stacked and fed to the KNN classifier for the final classification. Using this approach, have performed experimentation on a standard dataset, and achieve better results than the existing methods found in the literature. The comparison of our proposed model with the existing state-of-the-art models is competitive. Most of the models in the field already perform significantly well and coming up with a viable solution that performs better than the existing models is really a hurdle to overcome.

[10] proposed an ensemble method called Weight Optimized Neural Network with Maximum Likelihood Boosting (WONN-MLB) for accurately diagnosing LCD in big data using Artificial Neural Networks (ANN). They achieved a precision value of 93% and a recall value of 91%, resulting in an F1-score of 92%. [201] introduced an innovative automated diagnosis method for lung CT images. It employed Optimal Deep Neural Network (ODNN) and Linear Discriminant Analysis (LDA) to analyze the CT scans. The deep features from CT lung images were extracted and reduced using LDR to classify nodules as malignant or benign. The ODNN was then optimized with a Modified Gravitational Search Algorithm (MGSA) for lung cancer classification. The proposed classifier achieved a sensitivity of 96.2%, specificity of 94.2%, and accuracy of 94.56%.

[202] used the DenseNet-based CNN model with an adaptive boosting algorithm for classification, and they considered 541 lung cancer CT images from the LIDC dataset. They achieved an 89.85% accuracy rate, which was better than DenseNet without Adaboost, ResNet, VGG16, and AlexNet. They used various image translation and transformation tools to balance the training data.

[203] developed an advanced computer-aided decision support system for lung nodule detection, utilizing a 3D Deep Convolutional Neural Network (3DDCNN) to assist radiologists in lung cancer diagnosis. The system leveraged 3-dimensional information from CT scans by applying median intensity projection and multi-Region Proposal Network (mRPN) for automatic region-of-interest selection. Their CAD system was trained and validated using LUNA16, ANODE09, and LIDC-IDR datasets, demonstrating superior performance with 98.4% sensitivity, 92% specificity, 96% AUROC, and 98.51% accuracy at 2.1 false positives per scan. However, the study acknowledged a limitation in the accuracy of detecting micronodules with a diameter of less than 3 mm, which warranted further investigation and improvement.

In [204], researchers used Deep Convolution Network (DCN) to learn the features

of CT images of lung nodules and improved the model's representation by using the MLP, which enhanced lung nodule image classification accuracy.

[205] devised a simple AlexNet-based CNN architecture for the classification of lung cancer CT images. In order to compensate for the data scarcity, they used a few data augmentation techniques. They achieved a 91.68% accuracy rate on the LIDC dataset. Although data augmentation provided a limited quantity of extra information from which CNNs could learn features, it did not result in a considerable performance gain.

[206] proposed a lung cancer classification framework based on Generative Adversarial Networks (GANs) and discriminator networks wherein they further optimized the parameters using a novel technique. Their architecture was comprised of a deconvolution layer, batch normalization, and a leaky Rectified Linear Unit (ReLU) activation function to obtain a 64×64×3 image. Their research mainly focused on dealing with sparse image data to produce robust predictions. Their experimental results revealed that the average accuracy reached up to 99.86% after image augmentation and parameter optimization.

[207] proposed a deep-learning framework for automated lung nodule detection in CT images. They combined VGG-SegNet-based nodule mining, pre-trained deep learning-based classification, and handcrafted features like Grey Level Co-Occurrence Matrix (GLCM), Local-Binary Pattern (LBP), and Pyramid Histogram of Oriented Gradients (PHOG). The proposed approach achieved a 97.83% accuracy with the SVM-RBF classifier on experimental datasets (LIDC-IDRI and Lung-PET-CT-Dx).

[208] introduced a computer-aided detection method aimed at improving the detection of pulmonary nodules on CT scans. The proposed approach employed a 3D Residual UNet model in combination with a multi-branch classification network. Through multi-task learning, this method achieved a high detection sensitivity of up to 94.0% and an impressive competition performance metric (CPM) score of up to 0.959 on the Lung Nodules Analysis 2016 dataset.

[209] proposed LCD-CapsNet, a novel deep learning framework that combined a CNN and a Capsule Neural Network (CapsNet) for LCD and classification using CT images. The method aimed to develop algorithms capable of classifying and analyzing images from a dataset to determine the presence or risk of developing lung cancer. The deep learning model was evaluated using the LIDC dataset, consisting of 4335 images used in the training and testing pipeline. The experimental results demonstrated that LCD-CapsNet outperformed CapsNet, achieving an average Precision of 95%, Recall of 94.5%, F1-Score of 94.5%, Specificity of 99.07%, Area Under the Curve of 0.989, and Accuracy of 94% for both benign and malignant data. **Table ??** tabulates recent

work in LCD detection

The previous methods in LCD detection, utilizing both traditional machine learning and deep learning approaches, have made valuable contributions to the field. However, these studies have identified several key limitations that motivate the current research. One significant challenge is the accurate detection of small nodules, including micro nodules. Another limitation in the previous works is the lack of exploration and utilization of ensemble approaches. Ensemble models, which combine multiple individual models to make predictions, have shown promising results in various image classification and pattern recognition tasks. However, the majority of the reviewed studies focused on individual models rather than ensemble methods. This gap highlights the need for research that specifically investigates and harnesses the potential of ensemble approaches in improving the accuracy and robustness of LCD detection. Addressing the issue of data imbalance and variability is crucial, and ensemble models offer a promising solution. By combining multiple individual models, ensemble approaches can effectively handle imbalanced datasets [210] and capture diverse perspectives, reducing the risk of overfitting and bias. The initial segment presents a comprehensive summary of previous studies conducted on lung cancer classification. Further segment concentrates on examining ensemble techniques, specifically in the realm of medical image processing.

[176] presented an approach for the classification of lung nodules and non-nodules in CT images, which involved the utilization of texture features. They used three strategies to extract texture measurements: rose diagrams (RD), artificial crawlers (ACs), and a hybrid model that incorporates ACs & RD. The authors employed a support vector machine (SVM) classifier with a radial basis kernel to differentiate between nodules and non-nodules in the candidate categorization process. The study used the Lung Image Database Consortium (LIDC-IDRI) image database and achieved a mean specificity of 94.78%, mean sensitivity of 91.86%, and mean accuracy of 94.30%. However, the main limitation of this work is that the comparison with previous methods is only approximate due to differences in databases and test scenarios. The only precise comparison is with [177], and the proposed methodology achieved a lower specificity score, indicating its lesser ability to classify non-nodule cases accurately.

Netto et al. [211] proposed a methodology for analyzing lung lesions using temporal evaluation, which can aid in the diagnosis of indeterminate lesions during treatment. The modified quality threshold clustering technique was employed to assign each voxel of the lesion to a cluster, and the alteration in the lesion was evaluated by analyzing the movement of voxels to other clusters over time. To differentiate between benign and malignant lesions, statistical features are extracted. The authors employed two

databases of pulmonary lesions, one for malignant lesions under treatment and the other for undetermined cases, to develop their proposed methodology. By analyzing the density changes of lesions over time, the researchers achieved an accuracy of 98.41% in accurately identifying lung lesions. This demonstrates the potential of their proposed methodology in assisting with the diagnosis of indeterminate lung lesions.

Xie et al. [212] proposed a distinctive approach for lung nodule classification, named Fuse-TSD. This method combines information on texture, shape, and deep model-based learning at the decision level. The approach utilizes a texture descriptor derived from a gray-level co-occurrence matrix (GLCM), a Fourier shape descriptor that captures the heterogeneity of nodules, and a deep convolutional neural network (DCNN) that learns the feature representation of nodules automatically, slice-by-slice. The obtained characteristics are trained with an AdaBoosted back propagation neural network (BPNN), and the judgments of three classifiers are combined to distinguish benign from malignant nodules. This work has some limitations, including significant computational costs, particularly during offline training.

An automated diagnosis classification method for CT lung images was introduced by [201]. The method employs an Optimal Deep Neural Network and Linear Discriminate Analysis to extract deep features and reduce dimensionality. To optimize the Optimal Deep Neural Network for classifying lung nodules as benign or malignant, the authors used a modified gravitational search algorithm. The automated approach not only enhances the classification accuracy but also reduces the time required for manual labeling and prevents human errors in recognizing normal and abnormal lung images.

Masood et al. [203] developed a novel decision support system for CT scan-based lung nodule detection. Their approach utilized a 3D deep convolutional neural network, multi-Region Proposal Network, and median intensity projection to automatically identify regions of interest. The performance of the system was evaluated on LUNA16, ANODE09, and LIDC-IDRI datasets, and was found to outperform existing systems. The paper's main shortcoming is that the accuracy of detecting micronodules with a diameter of less than 3 mm is relatively low.

Lin et al. [206] put forth a framework for classifying lung cancer using Generative Adversarial Networks (GANs) and discriminator networks. They utilized a deconvolution layer, leaky Rectified Linear Unit (ReLU) activation function, and batch normalization to obtain a $64\times64\times3$ image. The main objective of their research was to overcome the challenge of sparse image data and generate precise predictions. They also proposed a novel technique for parameter optimization.

Yuan et al. [208] proposed a computer-aided detection method to enhance the

detection of pulmonary nodules on CT scans. The proposed method utilizes a 3D Residual UNet model along with a multi-branch classification network to achieve a high detection sensitivity of up to 94.0% and a competition performance metric (CPM) score of up to 0.959 on the Lung Nodules Analysis 2016 dataset through multi-task learning.

Halder et al. [140] presented an end-to-end system for detecting and classifying lung nodules from high-resolution CT images using atrous convolution. They achieved high-performance indices, with the proposed architecture, ATCNN2PR, consisting of a two-layer atrous pyramid and residual connections, demonstrating the highest classification accuracy. The results showed that the system outperformed other competing frameworks with an accuracy, specificity, and sensitivity of 95.97%, 96.89%, and, 95.84%, respectively.

Chapter 3

Segmentation of Lung Regions using Novel UNet Variants

Accurate segmentation of lung regions plays a pivotal role in computer-aided diagnosis and treatment planning for pulmonary diseases, particularly lung cancer. As imaging modalities such as CT become more prevalent in clinical workflows, automated lung segmentation has become a critical preprocessing step for subsequent analysis tasks like nodule detection, tumor classification, and disease progression monitoring. Traditional rule-based methods and early machine learning techniques often fall short in handling the complexity and variability of real-world CT images, especially in the presence of noise, pathological abnormalities, or anatomical distortions.

In recent years, deep learning has emerged as a powerful alternative, offering end-to-end learnable frameworks capable of extracting hierarchical and spatially meaningful features from large-scale medical datasets. Among these, CNNs have demonstrated substantial success in medical image segmentation. However, standard CNNs typically rely on local receptive fields, which may limit their ability to model long-range dependencies and global context—key factors when segmenting anatomically coherent but spatially dispersed regions like the lungs. To address this limitation, attention mechanisms have been integrated into deep learning models to enhance their focus on relevant regions of the input image. These mechanisms enable networks to dynamically weigh features based on contextual importance, improving both precision and consistency in segmentation outcomes. Variants such as self-attention, spatial attention, and channel attention have shown promising results in highlighting tumor boundaries and suppressing irrelevant background noise.

This chapter explores the application of attention-enhanced architectures for lung region segmentation, analyzing their performance, robustness, and adaptability across diverse imaging conditions. The goal is to design an effective segmentation framework that not only captures local structural detail but also leverages global contextual cues—ultimately contributing to more accurate and clinically usable outputs in lung cancer detection pipelines.

3.1 Method 1

3.1.1 Segmentation of Lung CT Images using R2UNet with Concurrent Spatial Channel Squeeze and Excitation Blocks

One of the novel contributions of this thesis is a methodology for improved version of UNet for segmentation of Lung CT Images. It draws ideas from the recurrent residual based UNet, traditional recurrent networks, and scSE as well [213]. The UNet encoder and decoder's Conv with ReLU structures are updated to the scSE module with recurrent residual block structures. The residual structure of R2UNet is used as a result of the increased network depth to prevent the vanishing gradient issue brought on by the increased network depth. Encoder and decoder, consisting of several blocks, are two of the network's components. Recurrent residual block, scSE block, and max pooling layer are the three components that make up the encoder block. Following every convolution layer, batch normalization and ReLU processing are applied. With the exception of using a de-convolutional layer in place of the pooling layer, the decoder block's structure is quite similar to that of the encoder block. The feature maps from the encoder layers to the matching decoder layers are concatenated using skip connections. Each decoder receives spatial information from the skip-connection architecture so that it may efficiently recover fine-grained details while generating output masks.

The architecture shown in Figure 3.1 consists of 4 pooling layers, 4 scSE blocks, 9 recurrent residual blocks, and 4 deconvolution layers. The input image is 224 x 224 x 1, the pooling layer is 2 x 2, and the convolution kernel is 3 x 3. Each image is subjected to several processes, including convolution, feature extraction, and pooling, before the binary segmented image with a resolution of 224 $\times$ 224 is generated. Recurrent residual learning is used to replace the ordinary unit of traditional UNet across the entire network architecture. In Figure 3.1 each convolution, batch normalization and ReLU activation processes are carried out in the recurrent residual unit. The recurrent residual structure, in contrast, adds shortcut connections based on a single forward link, allowing a deeper network to be trained without degrading while extracting more discriminative features.

3.1.2 Recurrent Residual Block

A Recurrent Residual Block (RRB) is a neural network component that combines the strengths of residual learning and recurrent connections to improve feature extraction and gradient flow. In this block, convolutional layers are applied recurrently—the

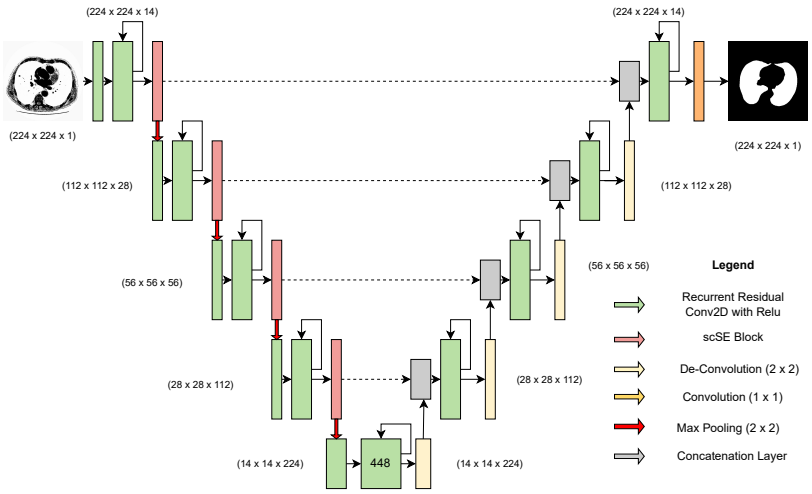


Figure 3.1: Network architecture of the lung CT image segmentation method using R2UNet

same weights are used across multiple iterations—to refine feature maps progressively. The residual connection then adds the original input back to the recurrently updated output, helping the network learn both low- and high-level representations while mitigating vanishing gradients.

Two consecutive recurrent convolution blocks make up each recurrent residual block, as seen in Figure 3.2. The final output is obtained by combining the output from the second recurrent unit with the original input using the residual connection. Information can be stored over time via the recurrent network using feedback connections. As time steps increase, the network makes more use of contextual information and neighbourhood data. The network can evolve over time even with static input by using recurrent convolution layers. Each unit is influenced by the neighbouring unit, which contains the image context information.

In the context of lung CT segmentation, where spatial and contextual information is vital, the feature reduction problem refer to the challenge of maintaining meaningful features related to the structure and context of lung images as they are processed through the network. The sequential and spatial nature of CT scans can make it difficult for traditional feedforward neural network to retain relevant contextual information across different regions of the image. By introducing recurrent connections with residual connections in the architecture can help mitigate the feature reduction problem by allowing the network to retain and propagate important information more effectively.

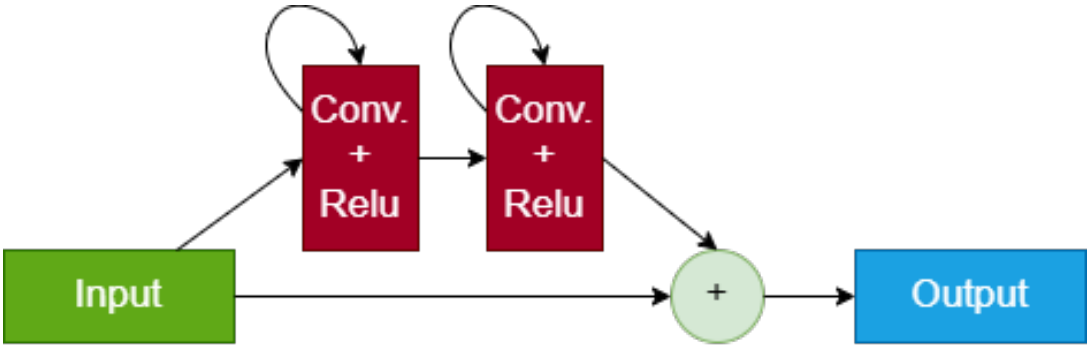


Figure 3.2: A typical recurrent residual block used in R2UNet based deep CNN

Channel squeeze-and-excitation and Spatial squeeze-and-excitation Blocks

Figure 3.3(a) illustrates how the channel squeeze-and-excitation (cSE) block improves convolutional feature learning by directly simulating the channel interdependency. Subsequent adjustments has been made to the network to increase its sensitivity and extract important characteristics. On the other hand, the spatial squeeze-and-excitation (sSE) block spatially constricting the feature maps and re-calibrated the model. The salient spatial locations have been improved through this procedure. The sSE block is shown in Figure 3.2 (b).

Concurrent Spatial and Channel Squeeze and Excitation Block

The channel and spatial excitation components are added one at a time while merging the cSE and sSE blocks to produce a scSE block, as seen in Figure 3.3(c) [213]. The goal of the scSE block is to concurrently promote the network's acquisition of more relevant channel and geographical characteristics. In order to concurrently re-calibrate the input's spatial and channel properties, this block combines the two Squeeze and Excitation (SE) blocks. The concurrent spatial and channel SE, also known as the concurrent spatial and channel squeeze & excitation (scSE) block, is created by sequentially adding the channel and spatial excitation components. When both channel re-scaling and spatial re-scaling provide high relevance, an area of the input feature map is given increased activation. In both the spatial and channel domains, this re-calibration encourages the network to learn more important feature mappings.

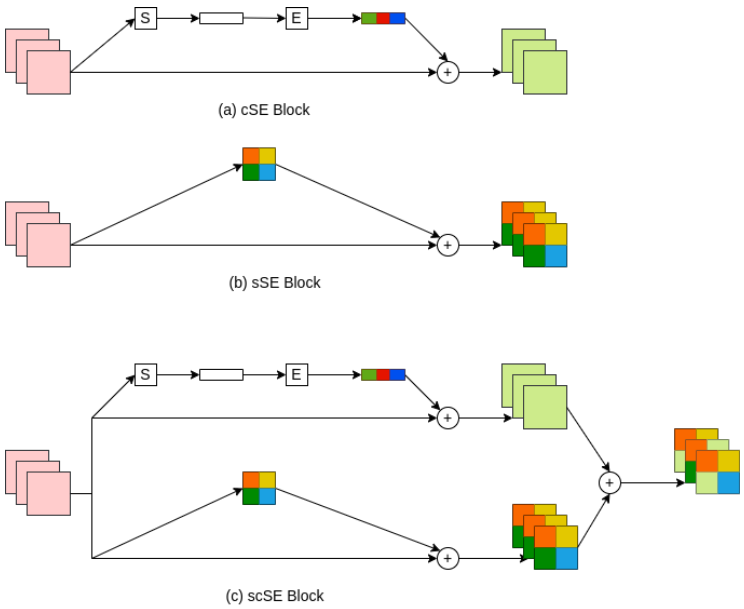


Figure 3.3: (a) Channel Squeeze & Excitation Block (cSE), (b) Spatial Squeeze & Excitation Block (sSE), (c) Spatial and Channel Squeeze & Excitation Block (scSE)

3.2 Method 2

3.2.1 An Ensemble of UNet Frameworks for Lung Nodule Segmentation

In this work, a technique for lung nodule segmentation using an ensemble model named BUS-UNet++ is developed, which combines BUS-UNet and UNet++ architectures. For evaluation, we use the publicly available LIDC-IDRI dataset. A sample image from this dataset is shown in Fig. 3.4. BUS-UNet++ has been implemented as an extension of the BUS-UNet and UNet++ architectures for medical image segmentation. The BUS-UNet++ model has 68 layers. Four steps have been added in the contracting path, each containing two convolution layers with 3×3 filters, followed by 2×2 max-pooling and a ReLU activation function. At each step, the number of feature maps has been doubled. In the contracting path, image features are extracted progressively, increasing the representation depth layer by layer.

In BUS-UNet++, direct skip connections have been introduced between the corresponding layers of the encoder and decoder, while also maintaining skip connections from the encoder to the aggregating blocks. The skip connections from aggregating blocks to decoder layers and to other aggregating blocks have been extended from the UNet++ architecture. The aggregating blocks have been used to concatenate features from different scales, followed by convolution operations before passing them to the next layer. Each aggregation block has used max-pooling, convolution layers, and

dropout layers to prevent overfitting.

During each up-sampling step, batch normalization has been applied to standardize the inputs to the next layer. After that, the data has been passed through a bidirectional ConvLSTM layer to process the input through both forward and backward paths.

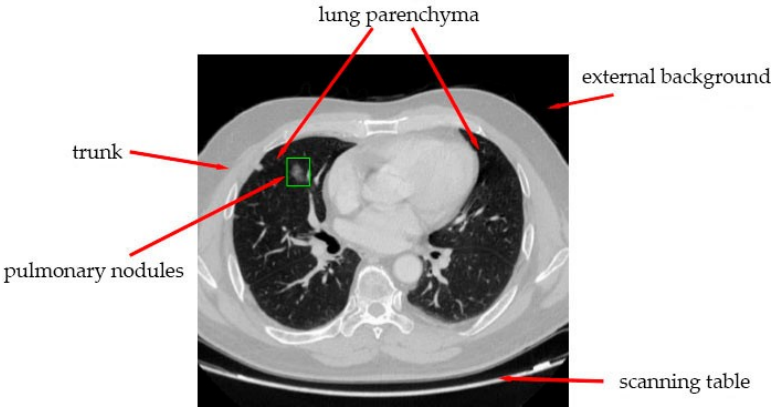


Figure 3.4: A sample lung nodule image from the LIDC-IDRI dataset

A combination of UNet++ and BUS-UNet is demonstrated in **Fig 3.5**. The overall pipeline of the conducted methodology is shown in **Fig 3.6**, and the model architecture is presented in **Fig 4**.

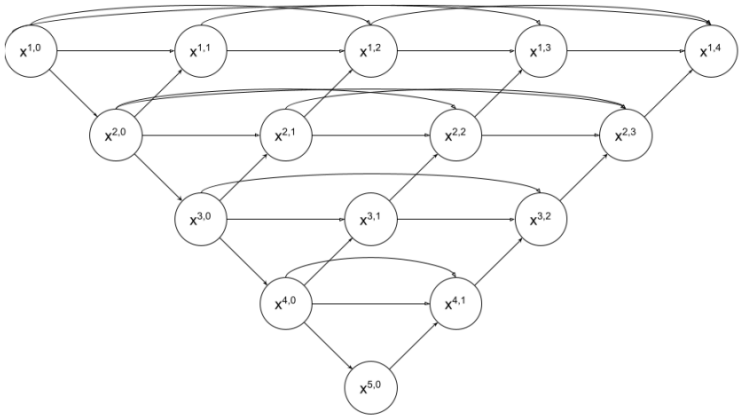


Figure 3.5: A diagram showing the UNet++ architecture

Table 3.1: Existing methods on lung nodule segmentation on LIDC dataset.

Authors	No. of nodules	Model	DSC (%)	Sensitivity (%)
Mamtha et al. [214]	1018 images	Integratingwater cycle algorithm and Bat algorithm	82.08	_
Yeganeh et al. [156]	1714 images	Res BCDUNet	93.83	_
Weihua et al. [215]	1186 images	UNet DCNN	91.97	_
Guilherme et al. [216]	888 images	Automatic Detection with Gaze Information	_	69
Rocha et al. [217]	2653 nodules	2DU-NET 2DSegUNet	83.0 84.3	89.8 85.8
Sunyi et al. [218]	1018 images	MIP-CNN	_	95.4
Usman et al. [219]	893 nodules	2D Residual UNet	87.55± 10.58	91.62± 8.47
Amorim et al. [220]	1018 images	2D Modified UNet	83	_
Shi et al. [221]	700 nodules	2DVGG 16+ SVM	_	90.00
Aresta et al. [222]	1012 images	3D iW-Net	75	_
Qin et al. [223]	1182 nodules	3D CGAN+3D CNN	84.33	85.11
Liu et al. [224]	3556 images	2D Mask R–CNN	_	_
Hongyang et al. [225]	1018 images	Multigrouppatch-based learning using Frangi Filter	_	94
Wu et al. [226]	1404 images	3D multi-task and InterpretableCNN	74.05	_
Wang et al. [227]	547 images	3D nodule R–CNN	64	_
Tong et al. [228]	1245 nodules	2D improved UNet	73.6	_

3.2.2 UNet++

UNet++ [229] is an advanced architecture for medical image segmentation, structured as a deeply-supervised encoder-decoder network. In this design, the encoder and decoder are linked through multiple nested and densely connected skip pathways. The aggregating blocks in these pathways are intended to merge features at different semantic levels. Beyond the main skip connections, additional links exist from the encoder to the aggregating blocks and from the aggregating blocks to other feature blocks or the decoder layers. The aggregating blocks connect highlights from various scales and afterward apply convolutions to them prior to passing them further. The quantity of aggregating blocks is proportional to the square of the network depth. UNet++ also

has a prunable architecture with deep supervision.

UNet++ consists of UNets of different depths, with their decoders densely connected at the same resolution through redesigned skip pathways. These architectural modifications provide several advantages. First, by embedding UNets of varying depths, UNet++ eliminates the need to select a specific network depth. The UNets partially share an encoder, and their decoders are interconnected. All constituent UNets are trained simultaneously, leveraging a shared image representation through deep supervision. This approach enhances overall segmentation performance and also enables model pruning during inference.

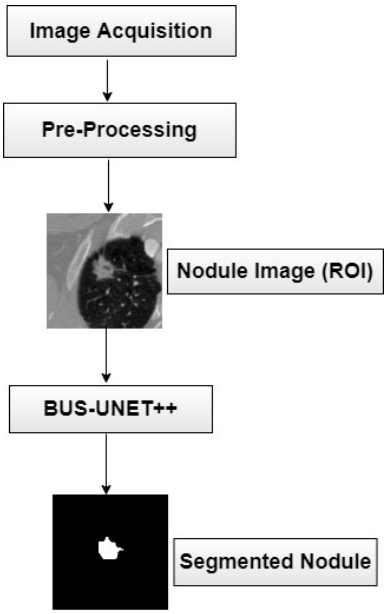


Figure 3.6: Overall pipeline of the developed methodology

Secondly, UNet++ is not hampered by overly restrictive skip connections that allow only the same-scale feature maps from the encoder and decoder to be fused.

3.2.3 BUS-UNet

The original model, Big-U Small-U Net (BUS-UNet) [230], consists of 108 layers formed by connecting two BCDUNets, with the first (Big-U) being deeper than the second (Small-U). The Big-U is deeper than the original BCDUNet, while the Small-U matches the size of the original network.

BCDUNet, an extension of UNet, has demonstrated superior performance compared to other state-of-the-art segmentation models. Its architecture features a contracting path composed of four steps, each containing two 3×3 convolutional layers followed by 2×2 max-pooling and a ReLU activation. The number of feature maps

doubles at each step. Along the contracting path, image representations are progressively extracted, increasing the depth of the feature maps layer by layer.

Densely connected convolutions address the UNet's tendency to learn redundant features in successive layers. Each step in the decoding path begins with an up-sampling operation on the output of the previous layer. Unlike the original UNet, where feature maps from the contracting path are cropped and concatenated onto the decoding path, BCDUNet processes these feature maps using bidirectional convolutional LSTMs (BConvLSTM). After each up-sampling operation, outputs are batch-normalized to improve network stability by standardizing layer inputs through mean subtraction and division by the batch standard deviation. Batch normalization also accelerates training. The BConvLSTM layer then processes the normalized outputs through forward and backward paths, capturing dependencies in both directions to make a decision for the current input. In contrast, the original ConvLSTM only considers dependencies in the forward direction.

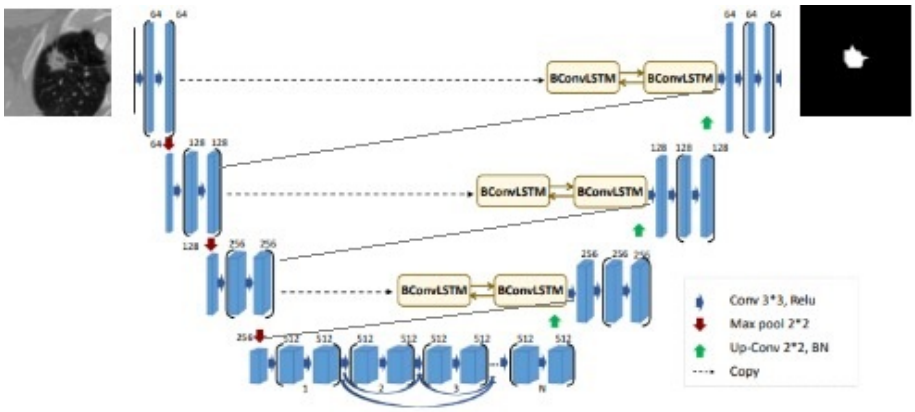


Figure 3.7: Architecture of the developed model, BUS-UNet++.

3.3 Experimental Results for Method 1

In light of the above methods, certain ablation studies have been conducted, and comparative analysis of developed and existing methods has also been mentioned. The box-plots for the comparison is also demonstrated.

3.3.1 Experiment Setup

The entire set of experiments has been run on a Jupyter Notebook equipped with a 12 GB NVIDIA Tesla T4 GPU that has been made available by Google's collaboratory environment. This environment is open for usage, as it is supplied by Google for

research. For the experiments, have implemented our developed method using Python 3, and the primary open-source modules used are: Tensorflow, Keras, Matplotlib, Scikit, Numpy, Pandas.

3.3.2 Dataset Description

The Lung Nodule Analysis (LUNA) competition at the Kaggle Data Science Bowl in 2017 was the first to employ a lung segmentation dataset [231]. This data includes images labelled for lung segmentation. For our assessment, use 70% of the data as the training set and the remaining 30% as the test and validation sets. The resolution of each image is 224 by 224 pixels. In this collection are thousands of DICOM-formatted low-dose CT scans from high-risk patients. The information required for the patient ID and scan parameters, such as slice thickness, are contained in the header of the DICOM files. A number of axial slices of the chest cavity are included in each picture. The subject and scanning instrument affect how many 2D slices there in each picture. have enhanced the training data in this study to improve the segmentation results of the suggested model.

3.3.3 Data Augmentation

The quantity of labelled training data limits the ability to segment medical images. Data augmentation helps in avoiding over-fitting of the model and improves the network's capacity to generalize to the data outside the training dataset. It is essential for building reliable deep learning networks. In medical imaging, the augmentations are applied equally to the images and labels. Effective sampling techniques or data augmentations are essential for the stable operation of neural networks.

The suggested model's inputs are CT images with a size of 224 x 224. Instead of using a sampling method, a data augmentation strategy has been used to increase the developed model's resilience and generalizability. The suggested network uses the data augmentation techniques of scale, center-crop, horizontal and vertical flip, random shift, and rotation, as shown in Figure 3.8. To further increase the generalizability of the suggested network, grid-distortion has been applied to the input slice of the CT image. Random shear, zoom, and rotation on the input image have also been implemented for greater variability. The training dataset now has more variations as a result of these image level changes, which also increases the resilience of the suggested model.

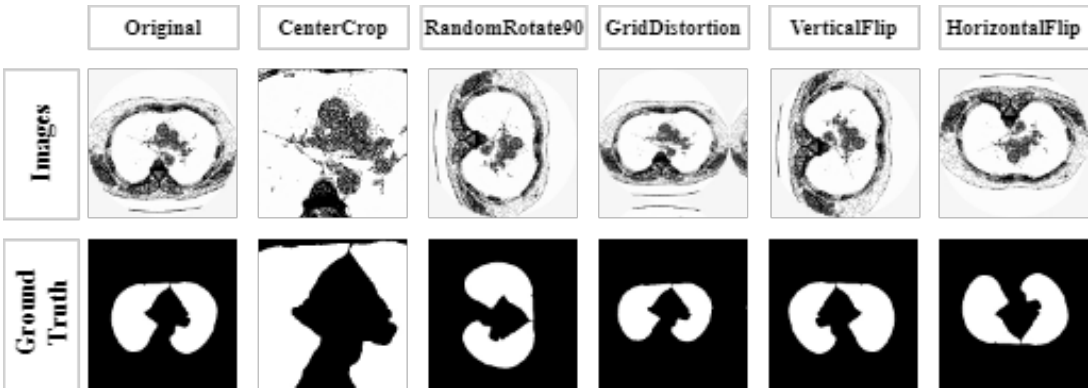


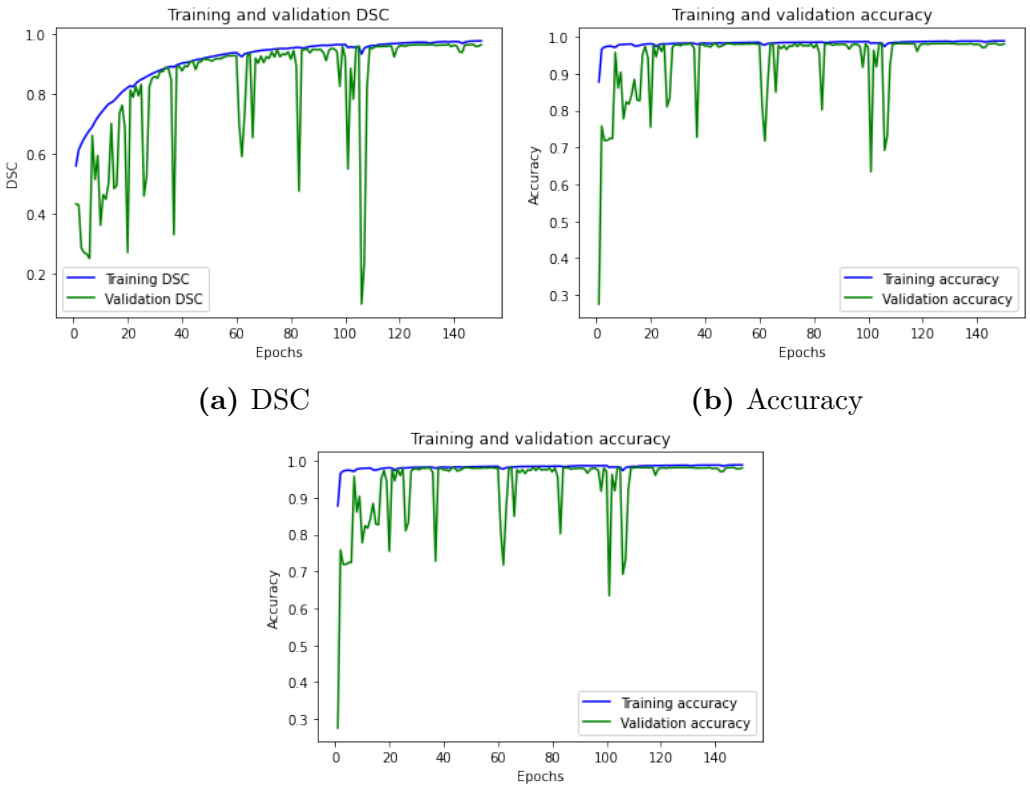
Figure 3.8: Some examples of different augmented images and their ground truth

3.3.4 Ablation study on the LUNA16 dataset

Five comparative ablation tests is conducted to confirm that the suggested network architecture is appropriate. The modifications are made one at a time, testing each one to see how it affects the overall framework, using UNet as the baseline. Recurrent residual structure is used in place of all convolution units in the UNet encoder and decoder components. Then, on the basis of the R2UNet model, which followed the convolution operation of the encoder, the scSE block is introduced. Finally, on the basis of R2UNet, scSE blocks are used to adaptively readjust the feature maps so as to highlight important traits while reducing the impact of less important ones. On the LUNA16 dataset, all of the models have been trained and assessed. The performance of the UNet model is noticeably improved by the combination of scSE and recurrent residual block when compared to a single utilization, as shown in Table 3.3. For comparison, the DSC, precision, recall, and accuracy produced by the network of different topologies are shown in Table 3.3. The developed method's DSC and accuracy are 97.36% and 98.55%, respectively. Training and validation results of different metrics of our developed model are given in Figure 3.9, Figure 3.10, and Figure 3.11.

3.3.5 Comparison with past segmentation methods

The DSC, precision, recall and accuracy generated by the network of various architectures are displayed in Table 3.4 for comparison. The final DSC and accuracy for the UNet are both 97.40% and 98.32%. The accuracy increases to 98.73% and the network's DSC to 97.07% for the R2UNet. The DSC of the network and the accuracy are both 97.03% and 98.17% for the SegNet model, respectively. But in the case of SEUNet and scSEUNet, DSC decreases to 90.03% and 91.48%, respectively. Residual UNet network architecture's DSC and accuracy increase as compared to the SEUNet and scSEUNet. Recall increases for R2UNet to 97.91% but in UNet, precision is 98.05%



(a) DSC **(b)** Accuracy

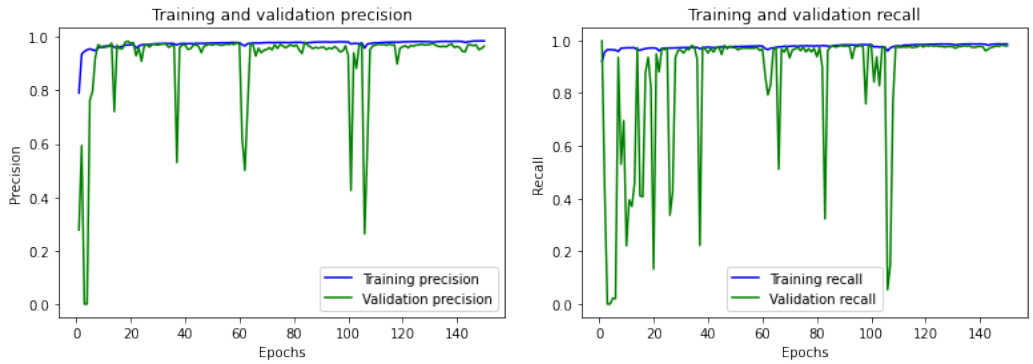
Figure 3.9: Training and validation DSC and accuracy scores of the developed method

Table 3.2: Comparison of the suggested method's performance utilising various optimisers and loss functions on the LUNA16 test set. The performance metrics' values are displayed as percentages %

Optimizer	Loss Function	DSC	Precision	Recall	Accuracy
Adam	Binary Cross entropy	**97.36**	**97.72**	98.23	98.55
	Binary Focal Loss	86.64	96.40	**98.66**	98.42
	Dice Loss	97.24	97.02	98.62	98.37
SGD	Binary Cross entropy	85.47	89.85	87.98	96.40
	Binary Focal Loss	81.57	90.96	94.86	**98.66**
	Dice Loss	80.15	85.21	83.03	85.12
RMSprop	Binary Cross entropy	97.30	97.15	97.77	98.23
	Binary Focal Loss	95.69	97.58	97.08	97.86
	Dice Loss	95.15	94.37	95.44	96.33

which is higher than other existing methods. Some visualization of the segmentation results and comparing them with ground truth segmentation of our developed model on the test images are shown in Figure 3.12.

Compared to other segmentation frameworks, DSC, recall and accuracy improves to 97.36% , 98.23% and 98.55% when use R2UNet with scSE block. The combination of this structures has a favourable effect on the performance improvement of the other



(a) Precision **(b)** Recall

Figure 3.10: Training and validation precision and recall scores of the developed method

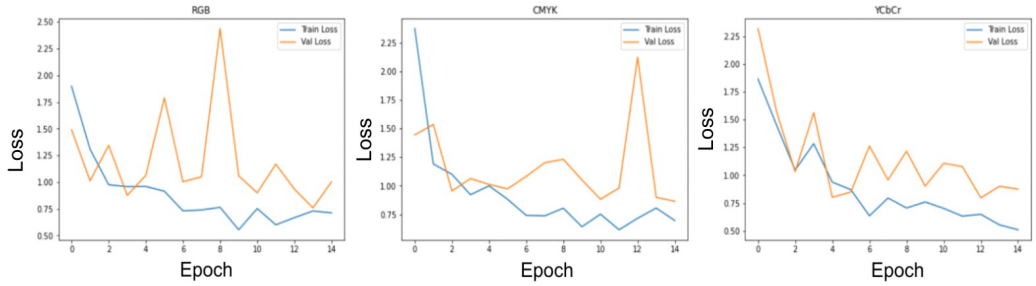


Figure 3.11: Training and validation loss curves of the developed method

Table 3.3: Results of the ablation study on the LUNA16 dataset. Values of the performance metrics are shown in terms of %

Method	DSC	Precision	Recall	Accuracy
ResidualUNet with cSE	96.07	96.40	96.57	97.80
ResidualUNet with sSE	97.24	97.44	**98.46**	98.50
ResidualUNet with scSE	97.06	97.84	96.84	98.16
R2UNet with cSE	96.76	97.48	97.41	98.26
R2UNet with sSE	97.18	**98.25**	97.22	98.35
R2UNet with scSE Block (developed)	**97.36**	97.72	98.23	**98.55**

UNet based approaches. In Figure 4.1 results of different architectures compared with developed method and images with ground truth are shown. It is shown that in Figure 4.1, our model gives better segmentation results than the other segmentation model like SegNet, SEUNet , Residual UNet and R2UNet.

The DSC scores for each model during 150 epochs of training, are demonstrated in Figure 3.14 on the LUNA16 dataset. The suggested model and the model with deep supervision both exhibit higher segmentation performance for the lung dataset than other models, as shown in Figure 3.14. The results for SegNet, SEUNet, Residual

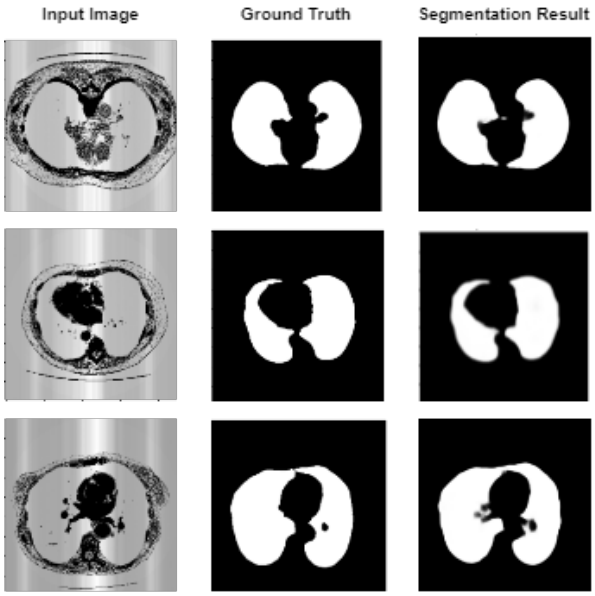


Figure 3.12: Segmentation results of the developed model with their ground truth

Table 3.4: Comparative score of different methods on the LUNA16 dataset. Values of the performance metrics are shown in terms of %

Method	DSC	Precision	Recall	Accuracy
UNet	94.40	**98.05**	97.18	98.32
SegNet	97.03	97.53	97.01	98.17
SEUNet	90.03	90.61	91.74	94.67
scSEUNet	91.48	88.12	96.39	95.51
Residual UNet	96.99	96.95	97.51	98.06
R2UNet	97.07	97.56	97.91	98.43
developed	**97.36**	97.72	**98.23**	**98.55**

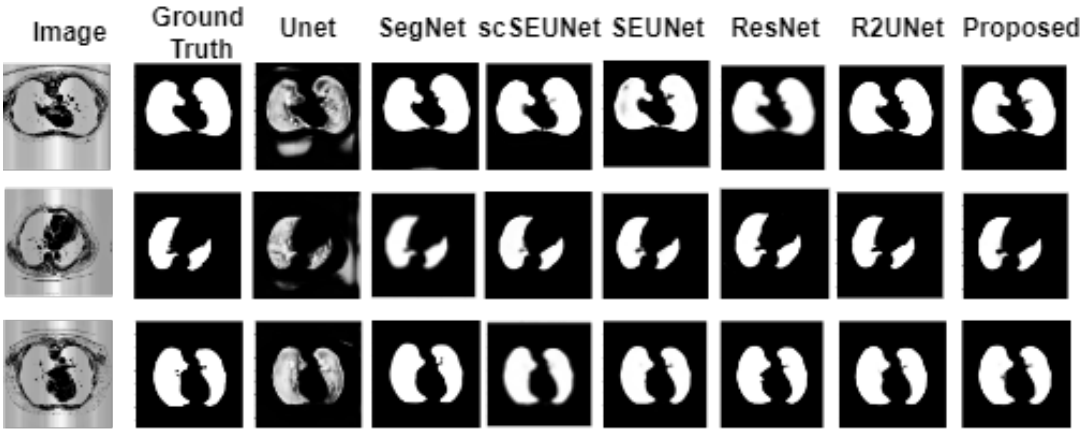


Figure 3.13: Segmentation results of existing methods and the developed method

UNet, R2UNet, and the developed method on the LUNA16 dataset are shown in box plots. This plot shows that the DSC of our developed model is 97.36%, which is higher

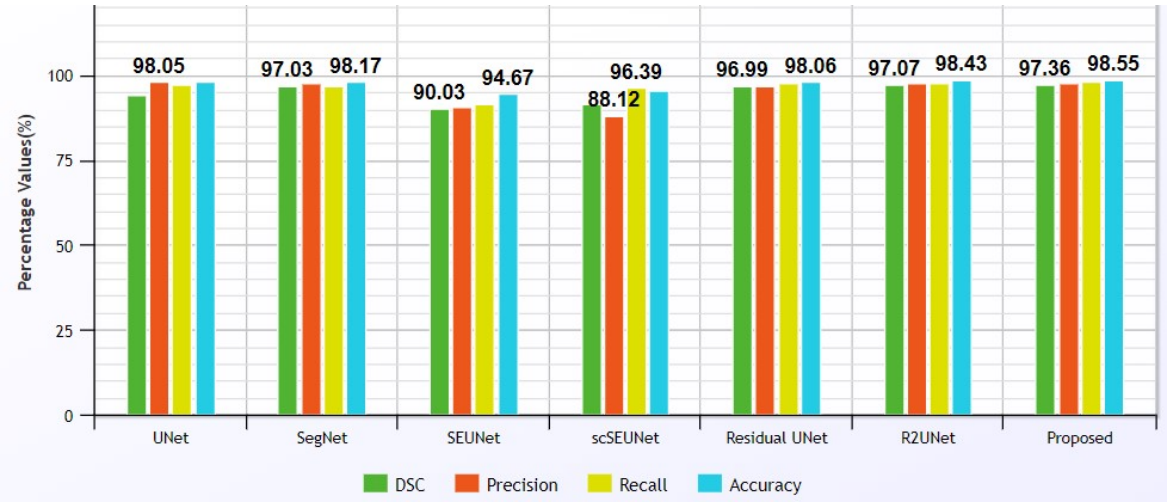


Figure 3.14: Bar chart showing the performance comparison of the developed method with various other methods

than the other existing models. The second best performance is obtained for R2UNet with DSC score of 97.07% followed by SegNet with DSC of 97.03%.

The suggested model's findings are given along with those from the residual and recurrent residual block and the scSE block. In order to compare the outcomes of our suggested model with those of the standard UNet, the traditional UNet models are also built. Initial results are used to create the performance metrics recall, precision, dsc, negative predictive score (NPV), and accuracy. In Table 3.4, the findings from developed and traditional UNet models that were highest are shown in the entries in bold text. make sure that the developed method is not skewed toward the deep learning model being considered here in order to ensure the effective use of the this model. The performance comparison of the suggested method with other widely-used deep learning models, is one efficient way to verify this. employ some of the deep learning models for this purpose. In Table 3.5, present the significant results from this investigation. It is evident from Table 3.5 that the suggested method performs admirably good with all deep learning models.

3.3.6 Data Pre-processing

The training of the model essentially requires the medical image data. There is an imbalance in the positive and negative sample categories in the dataset that might have an impact on the performance of the neural network. This is mainly due to the fact that the network is unable to reach its optimal value of the weights due to class imbalance. Therefore, utilize the Augmented Generative Adversarial Network (AugGAN) for the data enhancement. It solves the problem of limited positive samples to some extent in

Table 3.5: Comparison with state-of-the-art methods. Values of the performance metrics are shown in terms of %

Method	DSC	Precision	Recall	Accuracy	Year
TSDID29 [232]	-	-	89.53	96.65	2018
BCDUNet [233]	96.32	99.02	98.03	97.21	2019
ResNet34-UNet [234]	95.28	97.32	98.35	96.73	2020
AWEUNet [235]	89.79	-	-	91.32	2021
ResBCDUNet [156]	97.31	99.93	97.45	97.83	2021
Dense-UNet-Ds [236]	85.56	85.67	85.26	-	2022
DB-Net [237]	88.89	-	-	-	2022
DS-CMFSF [238]	96.00	93.47	97.38	-	2022
DDUNet [239]	96.90	-	-	98.30	2022
AgresNet with CNN [240]	-	-	-	98.00	2024
DL based Segmentation [241]	80.00	-	92.00	-	2025
Pyramidal Attention Network [242]	-	98.50	-	98.50	2025
CapsNet with UNet [243]	-	97.90	-	98.00	2025
Modified UNet [244]	85.30	-	-	-	2025
developed	**97.36**	**97.72**	**98.23**	**98.55**	**-**

the dataset. Also, it reduces the possibility of model overfitting. Some samples images generated by the AugGAN model are shown in Figure 3.15.

Figure 3.15: Lung nodules as generated by the AugGAN model

Region of Interest (ROI)

The ROI slice is extracted with a dimension of 512x512 from the scan in the LIDC-IDRI database. have used the centroid information for this purpose. Further, have cropped the slice into a 64x64 sized image with the same centroid. In this way, the RoI images are obtained as shown in **Fig 6** and **Fig 7**.

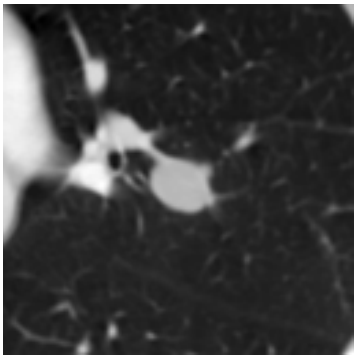


Figure 3.16: RoI image [LIDC-IDRI-0014]

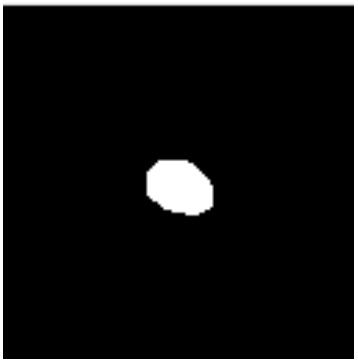


Figure 3.17: Mask image [LIDC-IDRI-0014]

3.4 Experimental Results for Method 2

3.4.1 Dataset Description

In this study, the LIDC-IDRI dataset was used to evaluate the developed model. This dataset has a total of 1186 lung nodules in a set of 888 CT images. For the experimentation, have randomly divided the data into three parts, wherein 70% is used as training data, 20% is used as testing data, and 10% of the data is used for validation data. In order to judge a nodule in the LIDC-IDRI dataset, the annotations are done by four radiologists, where a nodule is determined by its radius. Nodules with a radius of greater than 3 mm are identified as nodules belonging to the RoI, whereas nodules with a radius of less than 3 mm are annotated as non-nodules.

3.4.2 Evaluation Metrics

This section describes the metrics which are used to evaluate the developed method implemented in this paper. The first one is the dice score coefficient (DSC) index, which refers to the degree of fit between the original target and the segmented target.

The DSC value is higher based on how much the two objects fit and simultaneously lower the loss function value, and it indicates that the segmentation model of the lung nodules is more accurate. DSC is calculated using **Equation** (1):

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{1}$$

Here, true positive (TP) represents the area where the lung nodule exists as well as its correctly segmented area, false positive (FP) represents the area where the lung nodule exists but is not correctly segmented, and false negative (FN) represents the area where the lung nodule does not exist and is not segmented. When the loss function

Table 3.6: Lung nodule segmentation results of BUS-UNet++ with varied optimizers.

Optimizer	Accuracy	DSC	IoU
Adam	0.9771	0.9580	0.8439
SGD	0.9529	0.8594	0.7521
AdaGrad	0.9470	0.6771	0.7289
AdaDelta	0.9180	0.5591	0.6011

is infinitely close to 0, the invalid dice score coefficient reaches close to 1. In an ideal scenario, the model segmentation result matches the real result at this point. The relationship between DSC and loss coefficient is defined in **Equation** (2):

$$loss = 1 - DSC \tag{2}$$

IoU: It is another standard metric for semantic image segmentation that have used in this study to evaluate our model performance.
have obtained an IoU of 0.84 and a DSC of 0.958 with our developed method as shown in **Table 3.6** and **Table 3.7**.

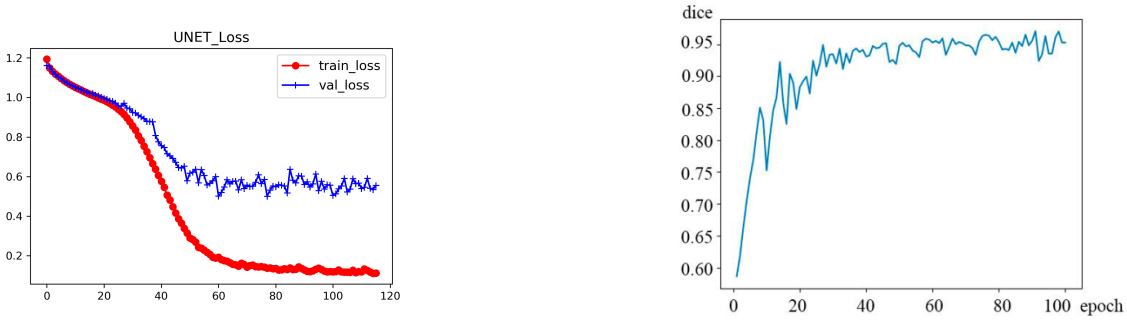


Figure 3.18: Training and validation loss values (left) and the network DSC value (right) of the developed model.

The above suggested method describes the performance of the various standard models along with the developed model through a set of experiments. All the evaluation metrics mentioned are calculated and it is observed that the developed model

Table 3.7: Performance comparison of the developed model with state-of-the-art methods on the LIDC-IDRI dataset

Work Ref.	**Method**	**DSC (%)**
Badrinarayanan et al. [16]	SegNet	84.21
Ronneberger et al. [15]	UNet	82.90
Zhou et al. [106]	UNet++	83.32
Huang et al. [245]	UNet-3+	89.72
Yeganeh et al. [156]	Res BCDUNet	93.83
Khoong et al. [230]	BUS-UNet	94.25
developed	**BUS-UNet++**	**95.80**

gives the best result with Adam optimizer, with an accuracy rate of 0.9771, as the loss function stabilizes more quickly as compared to other optimizers tested for the same purpose. The BUS-UNet++ helps to improve the segmentation accuracy of the lung nodules and has achieved good results with a dice score coefficient of 0.9580 as shown in **Fig 8**. The model also outperforms several state-of-the-art networks such as UNet DCNN and BUS-UNet that have generated a dice score coefficient of 0.91 and 0.94 respectively. However, there are a significant number of false positives which can be reduced by improving the feature extraction pipeline before training the network. Lung cancer is one of the most common types of cancer in today's world. Research reveals that accurate early identification of malignant lung nodules is essential for the prevention of this disease. In this paper, have developed a UNet based model, called BUS-UNet++, for segmenting the lung nodules in the LIDC-IDRI dataset. The BUS-UNet++ network is a combination of UNet++ and BUS-UNet architectures. The CT images for this purpose are extracted from the LIDC public repository and further preprocessed to obtain the nodule image along with its mask image. A set of experiments has been performed with several UNet variants such as UNet++, UNet-3++ and BUS-UNet to generate the segmentation maps. The performance metrics such as accuracy, DSC and IoU for BUS-UNet++ have shown that it outperforms the other three variants used in this study.

Chapter 4

Classification of Lung Cancer Images Using Novel Deep Learning Algorithms

4.1 Method 1

4.1.1 Lung Cancer Detection from Thoracic CT Scans using an Ensemble of Deep Learning Models

An ensemble architecture with three classifiers is implemented in this study namely, ResNet-152 [246], DenseNet-169 [247], and EfficientNet-B7 [248], using a weighted average ensemble method, wherein use a novel scheme for weight optimization before allocating those weights to the classifiers, as explained in detail in the following sections. The overall pipeline of the developed model is shown in **Fig.** 4.1. The developed model utilized for LCD detection from CT scan images has three primary stages: the generation of probability scores from multiple CNN base models, the fusion of the scores using the developed ensemble scheme for making the predictions, and optimizing the weights distributions of the three base models in the ensemble model using a novel approach to get the final predictions. These three phases are described below.

4.1.2 Generation of Probability Scores

At first, three transfer learning-based CNN models, ResNet-152, DenseNet-169, and EfficientNet-B7 are utilized to generate the probability scores on the sample images. The dataset is split into an 80%-10%-10% ratio of train, validation, and test sets, and the same sets are used for all the models. The Adam optimizer, along with Cross Entropy loss is used to fine-tune the networks for 10 epochs each on top of ImageNet weights. Class Weighting has also been performed to tackle the problem of class imbalance in the dataset used in this study. The hyperparameter values used to train the learning algorithms (base models) are set empirically which are shown in **Table 4.1**.

The ensemble learning model integrates the discriminative data (here confidence scores) of all its constituent models, and consequently, it has been found that its

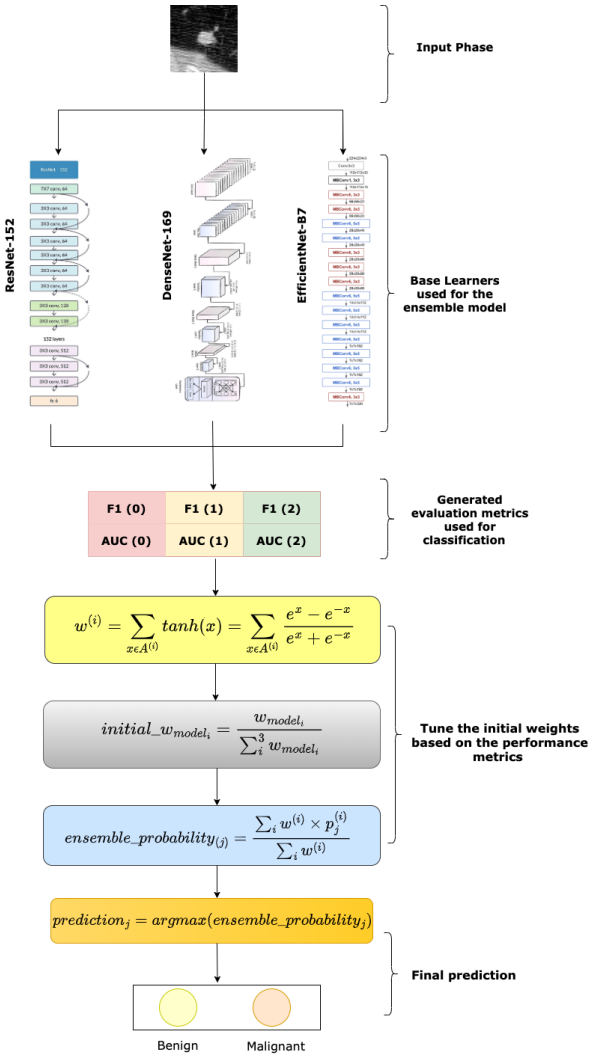


Figure 4.1: Developed methodology for ensemble based classification model

predictions outperform those of any of its constituents. models. The weighted average ensemble is a strong classifier combination procedure. Nonetheless, the decision of the weights to be designated to the separate base models assumes a vital part in guaranteeing the outcome of the ensemble. Most methodologies in the literature set the weights tentatively or depending exclusively upon the exactness (accuracy) of the models. However, this may not be a useful measure when deal with a class-imbalanced dataset. The dataset distribution for the dataset used in this study has been shown in **Table 4.3**. The utilization of other assessment measures, like accuracy, recall (sensitivity), F1-score, and ROC-AUC score, may give generally strong data for measuring the predicting capability of the base models for a class-imbalanced dataset. To this end, in this paper, have contrived a novel methodology for weight designation and optimization, which made sense in the accompanying segments, along with a dry run.

Table 4.1: Hyperparameters and their associated values utilized to train CNN base models.

Hyperparameter	Value
Learning Rate Scheduler	StepLR
Initial Learning Rate	0.0001
Loss Function	Cross Entropy
Optimizer	Adam
Number of Epochs	10
Batch Size	16

Table 4.2: Comparison of the weights of the selected base models using different combinations of the evaluation metrics.

Base Model	Weight distribution	
	Using all five metrics	**Using F1 and ROC-AUC**
ResNet-152	0.3316	0.3317
DenseNet-169	0.3345	0.3344
EfficientNet-B7	0.3337	0.3338

Firstly, the weights assigned to each base model using the developed technique are determined using the base models' probability scores which are obtained throughout the training phase. Then an ensemble is created using these produced weights on the test set. This procedure is executed to guarantee that the test set stays autonomous for predictions.

In order to generate the corresponding array $A^{(i)} = \{Accuracy^{(i)}, Precision^{(i)}, Recall^{(i)}, F1^{(i)}, ROC - AUC^{(i)}\}$, the predictions of the i^{th} model $\hat{y}^(i)$ are generated and compared with the true labels (y).

For the purpose of our experiment, have considered all five performance metrics, namely accuracy, precision, recall, F1 score, and ROC-AUC score. The metrics are defined in the result section. The weights of the selected base models generated with all five metrics are almost the same as the values of the weights generated using only F1 and ROC-AUC scores, as demonstrated in **Table 4.2**. have considered only the F1-score and the ROC-AUC score, as it is known that these two metrics are good measures for the class-imbalanced dataset. This is explained in more detail using a Dry Run in the **Appendix A** section.

This results in an array $A^{(i)} = \{F1^{(i)}, ROC - AUC^{(i)}\}$.

4.1.3 Fusion of Decision Scores

The hyperbolic tangent function is then used to compute the weight $w^{(i)}$ assigned to each classifier, as shown in **Equation 1**. As x represents an evaluation metric with a value in the range $[0, 1]$, the hyperbolic tangent function's range is $[0, 0.762]$. In this region, it rises monotonically. As a result, the *tanh* function rewards a measure if the value of x is high with a high value by giving it a high priority; otherwise, the function penalizes it.

$$w^{(i)} = \sum_{x \epsilon A^{(i)}} tanh(x) = \sum_{x \epsilon A^{(i)}} \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4.1}$$

These weights $(w^{(i)})$ computed by **Equation 4.1** to determine the weighted average probability ensemble, are multiplied by the decision scores of the associated base models, as shown in **Equation 4.2**, where the probability array (for a binary class dataset) of the j^{th} test sample by the i^{th} base classifier is, $p_j^{(i)} = \{a, 1-a\}$, where $a \leq 1$ and the ensemble probability for the test sample is $ensemble\_probability_{(j)} = \{b, 1-b\}$ for (benign, malignant) classes.

$$ensemble\_probability_{(j)} = \frac{\sum_i w^{(i)} \times p_j^{(i)}}{\sum_i w^{(i)}} \tag{4.2}$$

Finally, the class prediction is made by the ensemble as defined by **Equation 4.3**, where $prediction_j$ denotes the predicted class of the sample.

$$prediction_j = argmax(ensemble\_probability_j[benign, malignant]) \tag{4.3}$$

4.1.4 Weight Optimization

Initially, all three models have been given equal weights while calculating the ensemble probability. The weight distribution should depend on the individual performance of the models while classifying the samples. Thus propose a novel weight optimization technique wherein the distribution of the weights is done on the basis of test accuracies of the base models. Thus, take the ensemble probability and class prediction as a function of the performance metrics of the models, and weights are assigned based on that. propose a mathematical approach to adjust the weights of the base models instead of intuition-based or random adjustment of the weights.

The weights $(w^{(i)})$ computed by **Equation 1** return an array of the weights of the base models. Let us consider the array to be

$w_{(i)} = \{w_{model1}, w_{model2}, w_{model3}, ..., w_{modeln}\}$

These weights are normalized using the formula as shown in **Equation 4.4**.

$$initial\_w_{model_i} = \frac{w_{model_i}}{\sum_i^n w_{model_i}} \tag{4.4}$$

Equation 4.4 also gives us the distribution of the weights, which forms the ensemble model. To calculate the ensemble, use F1-score and ROC-AUC and for weight optimization use recall. In medical science, the false negative reduction is a major task and thus in order to meet the precision-recall trade-off, optimize weights based on the recall metric.

To optimize the weight distribution which is calculated using **Equation 4.4**, have extensively experimented with the various performance metrics and observed that the recall metric showed a varied difference among the base models' performance. Thus, utilize the test recall metric of the base models to distribute the weights. have calculated their respective increase as shown in **Equation 4.5** with respect to the recall score of the base models. The *Increase* represented in the numerator is calculated by subtracting the recall value of each base model with the lowest generated recall value. A sample example is given in **Appendix A** of this paper.

$$Percentage\_Increase_i = (Increase/Recall_i) \times 100 \tag{4.5}$$

then increase the weights of the base models with their percentage increase value and decrease the weight of the lowest base model. As the model with the highest accuracy should be assigned the highest weight, therefore after weight optimization using the recall metric, the respective weight assignment to the base models is done according to its accuracy (i.e., the largest weight is assigned to the model with the highest test accuracy, and so on).

4.1.5 Deep Learning Architectures Utilized for the developed Methods

ResNet-152

The ResNet-152 model is a residual learning framework as developed by [246]. It increases the training efficiency of a deep neural network. In monotonically progressive convolutions, unlike the original unreferenced mapping, the residual blocks in ResNet models facilitate overall network optimization, which improves model accuracy. These residuals are essentially "skip connections" that accomplish identity mapping without adding parameters or increasing computational complexity. The ResNet-152 model's architecture is shown in **Fig 4.2**.

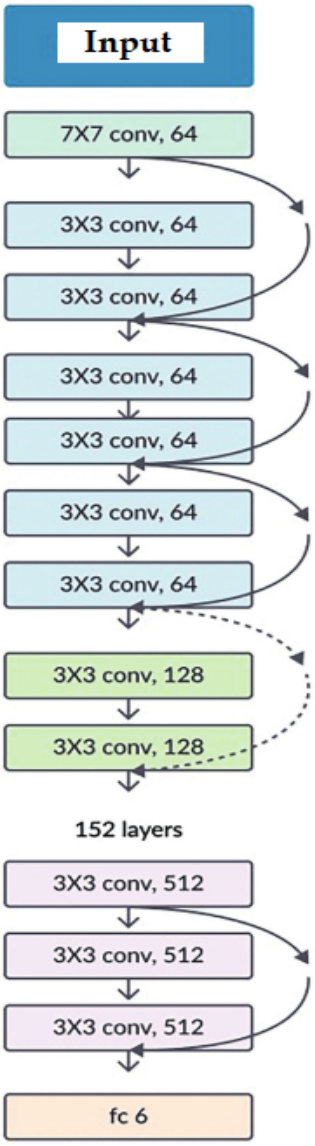


Figure 4.2: The ResNet-152 model architecture used in this study.

DenseNet-169

The DenseNet architecture developed by [247] provides a detailed feature representation that is also computationally efficient. As demonstrated in **Fig 4.3**, the feature mappings in the current layer of the DenseNet model are concatenated with those from all preceding levels. The convolutional layers accommodate fewer channels, which reduces the number of trainable parameters and makes the model more computationally efficient. Furthermore, the feature representation is improved by concatenating feature maps from previous layers with the current layer.

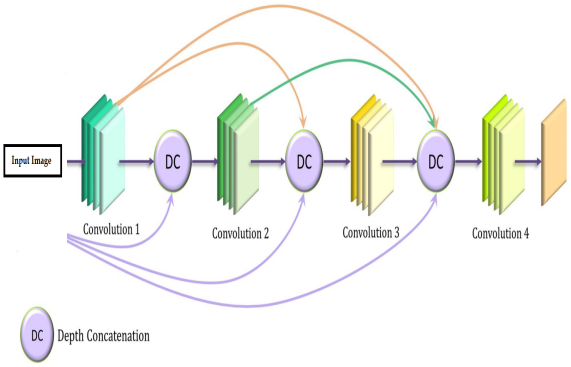


Figure 4.3: Basic architecture of the DenseNet model.

EfficientNet-B7

EfficientNet is a CNN architecture and scaling method developed by [248]. It utilizes compound scaling to balance depth, width, and resolution for better performance with fewer parameters as shown in **Fig 4.4**.

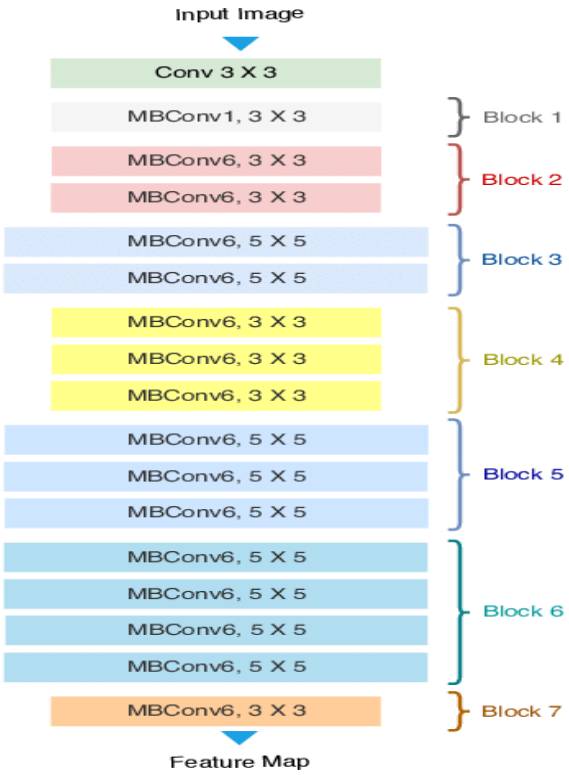


Figure 4.4: Architecture of the EfficientNet-B7 model.

Unlike traditional practice, which arbitrarily scales these factors, the EfficientNet scaling method uniformly scales network width, depth, and resolution using a fixed set of scaling coefficients. EfficientNet employs a compound coefficient to scale network width, depth, and resolution uniformly. EfficientNet B7 is the largest and most powerful model in the EfficientNet family of neural networks.

NasNet

NasNet is a CNN architecture, which consists of multiple blocks of convolutional and pooling layers. The blocks are designed to be highly modular and can be stacked on the top of each other to form a deep network. The architecture of each block is determined by a set of operations, such as convolution, pooling, and activation function, that are chosen from a pre-defined set of options. NasNet [249] uses reinforcement learning to optimize the selection of operations in each block, so that the network can automatically adapt to the task at hand. The key strength of NasNet is its ability to achieve high accuracy with relatively few parameters. This is important because it reduces the amount of computation required to train the network, making it faster and more computationally efficient. Additionally, NasNet has a highly flexible architecture that can be easily adapted to different computer vision tasks and input sizes. This makes it a versatile architecture that can be applied to a wide range of problems. A simple architecture of NasNet is shown in Figure 4.5.

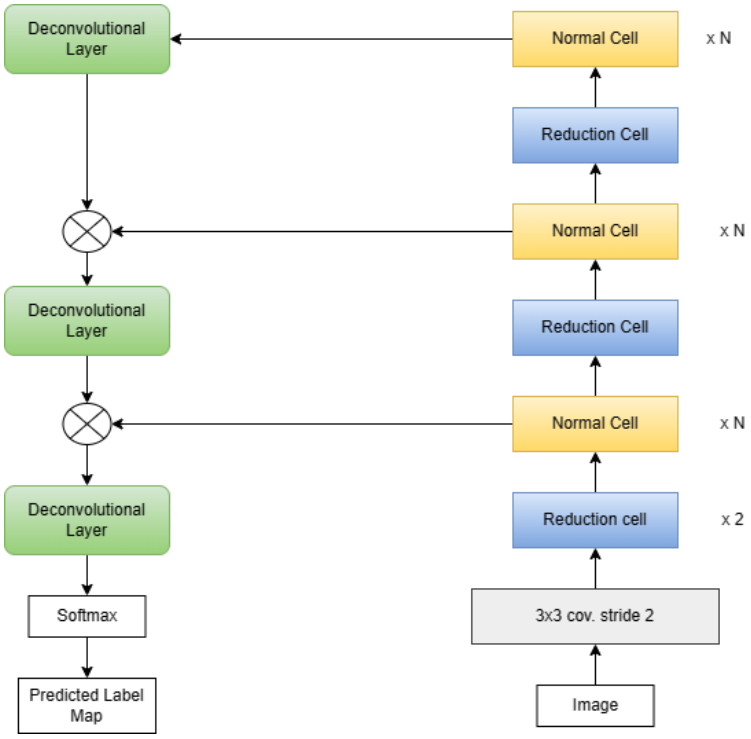


Figure 4.5: Architecture of the NasNetLarge [250] model used in this study

Keras offers two different NasNet models, namely NasNetLarge [251] and NasNetMobile. NasNetLarge has a higher number of trainable parameters compared to NasNetMobile. From the available performance comparison in Keras, it can be observed that NasNetLarge yields higher accuracy than NasNetMobile. Therefore, have chosen NasNetLarge [252] for our work.

4.2 Method 2

4.2.1 Mistcherlich Function based Ensemble

Another study is conducted for the classification of lung nodule. This approach is ensemble-based lung cancer classification model (MENet) described in-depth in this section. In order to determine the correct class of the test data, first provide a brief description of every base learner that produces scores of confidence for an incoming lung image. These scores are then further fused via the suggested ensemble methodology, according to rewarding and loss characteristics.

Deep Neural Based Classifiers

CNNs are preferred over other types of machine learning algorithms for image classification tasks, including the detection of lung cancer from medical imaging datasets. One of the main reasons for this is that CNNs are specifically designed to handle spatial data such as images. They are able to learn features from the input data by using convolutional layers and pooling layers combination, which can effectively capture local patterns and spatial relationships between pixels. This makes them well-suited for analyzing medical images, which often contain complex structures and patterns. Additionally, CNNs are able to automatically learn and adapt to the features of the input data, without the need for manual feature engineering. Overall, these characteristics make CNNs a powerful tool for accurately classifying medical images, including those of lung cancer and have the potential to increase diagnosis precision and speed. Hereafter an exhaustive set of experiments, have successfully settled down with Xception, InceptionResNetV2, and MobileNetV2 models, as these models have produced a satisfactory level of accuracy at the initial stage of the entire process.

Xception

Xception is a convolutional neural network architecture developed by [253]. It is an extension of the Inception architecture and is named "Extreme Inception" because it uses depthwise separable convolutions instead of standard convolutions used in Inception. The depth-wise separable convolution layer blocks make up the bulk of the Xception architecture, which is then followed by batch normalization and ReLU activation. The depthwise separable convolution layers are made up of two distinct layers: a pointwise convolution layer that applies a 1x1 convolutional filter to combine the output channels of the depthwise convolution and a depthwise convolution layer that applies a single convolutional filter to each input channel. A global average pooling

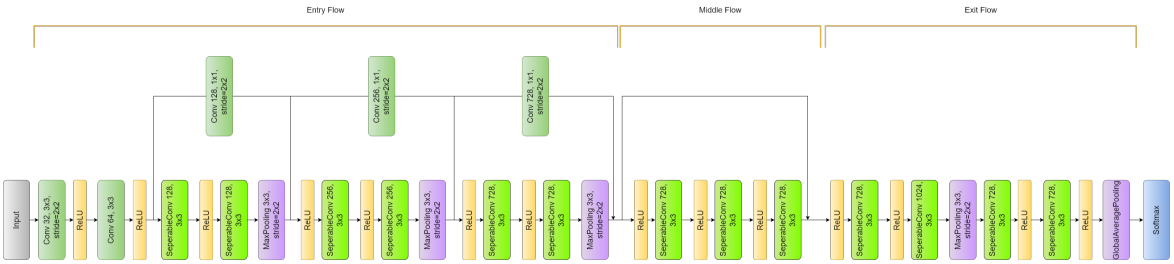


Figure 4.6: Architecture of the Xception model

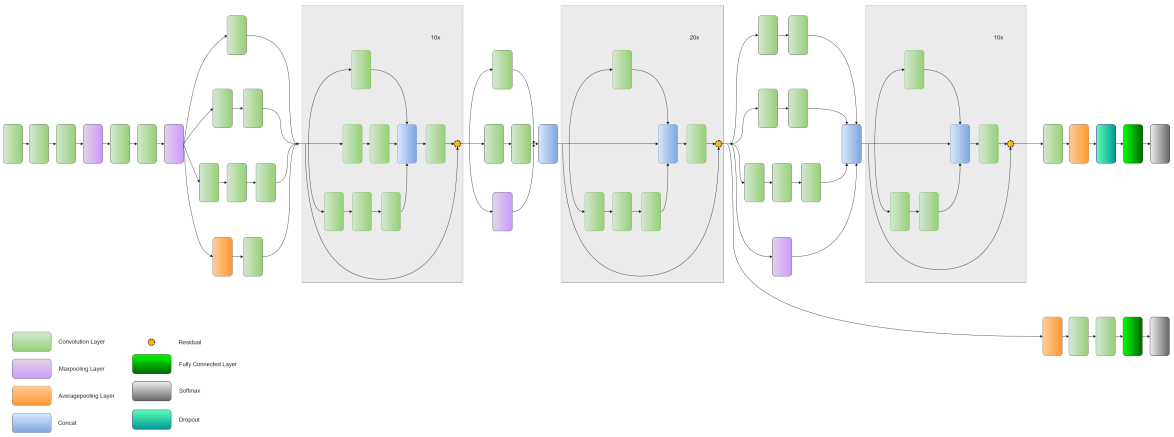


Figure 4.7: Architecture of the InceptionResNetV2 model

layer and a fully connected layer for classification, constitute the final layer of the network. The depthwise separable convolutions are more computationally efficient than conventional convolutions and aid in lowering overfitting, and the residual connections enable deeper network topologies while addressing the vanishing gradient issue. This makes Xception a powerful and effective model for image classification tasks. The Xception architecture is illustrated in Figure 4.6.

InceptionResNetV2

InceptionResNetV2 [254] combines the strengths of Inception and ResNet architectures. It utilizes residual connections with Inception modules, which not only address the vanishing gradient problem but also enable more complex network designs. Additionally, the model requires fewer parameters to achieve high accuracy and can quickly learn and extract features from input data. Overall, the InceptionResNetV2 architecture is a powerful tool for image classification tasks, offering both high accuracy and efficient computation. The InceptionResNetV2 architecture is shown in Figure 4.7.

MobileNetV2

MobileNetV2 [255], is a convolutional neural network designed for mobile and embedded vision applications. It uses depthwise separable convolution, inverted residuals, and linear bottlenecks to improve efficiency while preserving representational power. It includes two types of blocks: residual blocks with a stride of 1 and blocks with a stride of 2 for downsizing. Both types of blocks consist of three layers. The first layer in each block is a 1×1 convolution with ReLU6 activation. The second layer is a depthwise convolution, which applies a separate convolutional filter to each input channel. The final layer in each block is another 1×1 convolution but without any non-linearity. A width multiplier is used in this network to optimize the network for different hardware and resource constraints. MobileNetV2 is a highly efficient and lightweight model, making it suitable for deployment on mobile devices and other resource-constrained environments. Despite its small size, MobileNetV2 achieves high accuracy on a wide range of vision tasks, including image classification, object detection, and semantic segmentation. The architecture of MobileNet is shown in Figure 4.8.

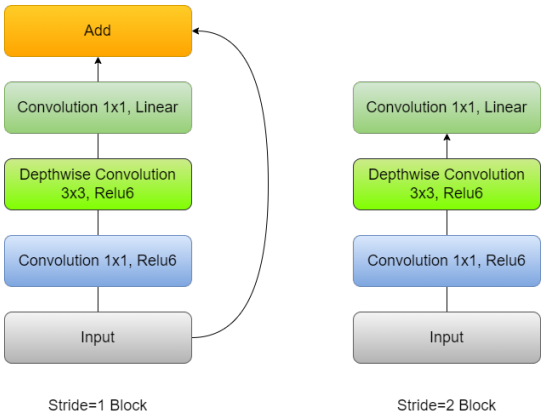


Figure 4.8: Architecture of the MobileNetV2 model

4.2.2 Fuzzy-Ensemble Method

The main goal of our developed approach is to provide greater adaptability and freedom in handling datasets with varying degrees of complexity. Using a fuzzy ranking-based approach, can consider the uncertainty of each classifier's predictions and assign different levels of importance to each classifier based on its performance on a particular test case. In the developed methodology, the three CNN-based classifiers (Xception, InceptionResNetV2, and MobileNetV2) are used to detect lung cancer cases from CT scans, and fuzzy ranks are generated for each of them using the re-parameterized Mitscherlich function. To increase the overall accuracy of the classification, these fuzzy ranks are

then combined using an ensemble method. The Mitscherlich function, which has been applied to the field of machine learning by [256], and [257], is based on the idea of crop yield response [258] to various levels of fertilizer application [259]. Here, it is used to combine the outputs of different models with various strengths and weaknesses.

Significance of Mitscherlich function

In the Mitscherlich function, a dose-response curve that depicts each model's performance as a function of input strength is used to illustrate each model's performance. The Mitscherlich function measures the relationship between the weighted sum of the predictions and the actual outcomes, and it returns a value that represents the fit between the predictions and the actual outcomes. This involves generating a fuzzy rank of a class for each model in the ensemble and combining the individual models to make predictions on a validation set for each rank inside the top K ranks. The Mitscherlich function is applied to the confidence scores of a class by each base classifier to generate fuzzy ranks for that class. This method enables the development of an adaptable ensemble model that can modify the ranking of individual models in response to the unique properties of each input instance. As the decision score of a class accurately predicted by a classifier mostly approaches one, the Mitscherlich function's steeply dropping nature in the domain range from 0 to 1, is helpful to form an ensemble of the learning models' decision scores.

Implementation of the Ensemble Model

For each input image P, there are M decision scores (also referred to as classifier confidence scores), denoted as $CoF^{(1)}, CoF^{(2)}, \ldots, CoF^{(M)}$. Since three different CNN-based models were employed to generate the confidence scores for the dataset, in our case $M = 3$. As shown in Equation 4.6, the decision scores are normalized, where C denotes the total number of classes in the dataset under consideration.

$$\sum_{c=1}^{C} CoF_c^{(i)} = 1; \forall i,\ i = 1, 2, 3, ..., M \tag{4.6}$$

Fuzzy ranks are generated using the scores of confidence of each sample in the dataset, which are divided into three different classes. The Mitscherlich function produces the fuzzy rank for a class c using the i_{th} classifier's scores of confidence as shown Equation 4.7.

$$R_c^{(i)} = 2(1 - 2^{CoF_c^{(i)} - 1}); \forall i, c,\ i = 1, 2, 3, ..., M; c = 1, 2, ..., C \tag{4.7}$$

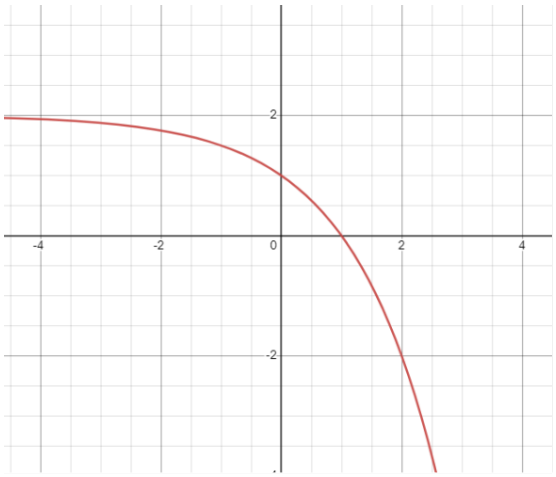


Figure 4.9: Graphical representation of the modified Mitscherlich function used in the present work.

The Figure 4.9 shows a pictorial depiction of the modified Mitscherlich function graph, as mentioned in Equation 2.

The value of $R_c^{(i)}$ is in-between 0 and 1, with the lowest value 0 being equivalent to rank 1 (best rank), i.e., a greater score of confidence results in a lower (better) value of rank. The fuzzy rank sum (FRS_c) and complement of the confidence factor sum ($CCFS_c$) are computed as follows if $K^{(i)}$ is used to represent only the top k ranks, that is, rank 1, 2,..., k, that corresponds to class c are as follows:

$$FRS_c = \sum_{i=1}^{M} \begin{cases} R_c^{(i)}, & \text{if } R_c^{(i)} \in K^{(i)} \\ P_c^R, & \text{otherwise} \end{cases} \tag{4.8}$$

$$CCFS_c = \frac{1}{M} \sum_{i=1}^{M} \begin{cases} CF_c^{(i)}, & \text{if } R_c^{(i)} \in K^{(i)} \\ P_c^{CF}, & \text{otherwise} \end{cases} \tag{4.9}$$

If class c does not appear within the top-k ranked classes, penalty terms P_c^R and P_c^{CF} are applied. According to the Mitscherlich function described earlier, P_c^R is assigned a value of 1 by setting $CF_c^{(i)} = 0$, while P_c^{CF} is assigned a value of 0. These penalties ensure that class c is unlikely to be selected as an improbable winner. The final decision score is obtained by combining FRS_c and $CCFS_c$ to generate the ensemble model's prediction. The final decision score (FDS) is computed using Equation 4.5.

$$FDS_c = FRS_c * CCFS_c \tag{4.10}$$

Finding the class with the lowest FDS value yields the final projected class for data

instance **I**, which is provided as shown in Equation 4.11.

$$class(\mathbf{I}) = \underset{c=1,2,...,C}{\arg\min} FDS_c \tag{4.11}$$

For a better understanding of this fuzzy rank-based ensemble method, have shown a dry run of the entire procedure under 4.5.4.

4.3 Method 3

4.3.1 Adaptive Genetic Algorithm based Deep Feature

GA is one of the most popular optimization algorithms used in various research fields over the years. Liu et al. [260] used GA to train neural network for land cover classification. Their approach included a real coded GA strategy, hybrid with Back Propagation (BP) algorithm. The model was able to achieve 97% accuracy on the SPOT-4 XS imagery of Jiangning County, Jiangsu, China. Singh et al. [261] developed a semi supervised method for satellite image classification based on GA and Radial Basis Function Neural Network (RBFNN). The Landsat 8 OLI dataset was used for this work. Initially the features were extracted from preprocessed images by spectral indices. The extracted features were fed into GA using which the RBFNN was trained till the termination conditions were met. The developed method managed to achieve an accuracy of 94.92%.

Tsai et al. [262] employed Gaussian-distributed fuzzy membership functions (GDMFs) for their study. The GDMFs were initially generated using various texture-based features obtained from reference images. Subsequently the shapes of GDMFs were optimized by a GA learning process. After optimization, the classifier is used for disease discrimination. The developed model aachieved an average accuracy of 96% for myocardial heart disease and accuracy of 88.5% at 100% sensitivity level for microcalcification on mammograms. Zhang et al. [263] introduced a novel approach that combines a neural-genetic algorithm with a neural network classifier for feature selection. Their research focused on the classification of various small breast abnormalities by incorporating computer-extracted statistical features from mammograms along with manually extracted features. The results demonstrated promising accuracy rates of 90.5% for calcification cases and 87.2% for mass cases using different feature subsets.

Ali et al. [264] developed a hybrid approach that combines filter-based feature

selection methods and a genetic algorithm to improve cancer classification accuracy on high-dimensional microarray datasets. The method applies filter feature selection techniques, such as information gain, information gain ratio, and Chi-squared, to identify significant features. A genetic algorithm then optimizes and enhances the selected features, showing superior performance in Accuracy, Recall, Precision, and F-measure compared to conventional machine learning methods on breast, lung, central nervous system, and brain cancer datasets. Sarkar et al. [265] introduced Diversification of Population in GA (DPGA), an FS algorithm that reduces the number of features extracted by a modified texture-based feature descriptor, achieving good outcomes on a 7-class microstructural image dataset. Lin et al. [266] developed a GA-based FS method for image retrieval and classification, reducing features and increasing accuracy at the expense of computational cost using adaptive motifs co-occurrence matrix (AMCOM), gradient histogram for adaptive motifs (GHAM), and adaptive color histogram for K-means (ACH).

Sarkar et al. [205] developed a model that encoded sensor data time series as multi-channel images using continuous wavelet transform. A spatial attention-aided CNN extracted higher-dimensional features, and a novel FS method combined with Mutual Information, Relief-F, and mRMR filters was used to identify important features for human activity recognition. Guha et al. [202] introduced the cooperative genetic algorithm (CGA), incorporating game theory concepts to enhance classification accuracy and time efficiency in human activity recognition. CGA showed significant improvement in overall classification accuracy using a small fraction of the original feature vector. GA has also been used successfully in various fields such as numerical recognition system for Devanagari , Bangla and Roman scripts by Chowdhury et al. [206], proposal of a new method called Deluge based GA (DGA) by Guha et al. [267], proposal of a novel model, named binary genetic swarm optimization (BGSO), which initially allows GA and Particle Swarm Optimization (PSO) to run independently by Ghosh et al. [268], handwritten word recognition by Malakar et al. [269], cancerous gene identification by Ghosh et al. [270], gene selection from microarray data by Marjit et al. [271], breast cancer detection in thermograms by Pramanik et al. [272]. Some more methods include stock prediction based on GA based FS by Chen et al. [273], stock market prediction using GA-optimized multi-channel CNN by Chung et al. [274], parameter optimization of solar PV cell/module using GA based on non-uniform mutation by Saadaoui et al. [275], GA based optimized leach protocol for energy efficient wireless sensor networks by Bhola et al. [276]. The success of these methods as well as the varied applications of GA makes it a promising algorithm for FS.

4.3.2 GA based method for lung cancer classification

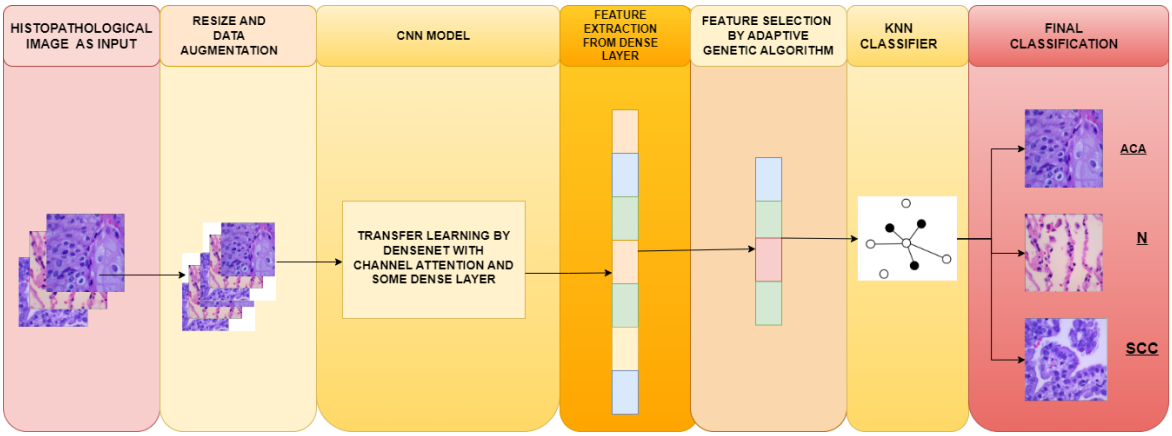


Figure 4.10: Overall workflow of developed method for lung cancer classification using histopathological images.

In this work, have developed a lung cancer classification model combining the concepts of deep learning and GA. Figure 4.10 shows the overall workflow of the developed method. Each of the steps has been discussed in details in the following sub-sections.

Data preprocessing

In this study, the training dataset is divided into training and validation subsets. To enhance the model's performance and generalize better, apply several preprocessing techniques. In the training data, first rescale the pixel values by dividing them by 255. This step ensures that all images have consistent pixel intensity values. Also, different data augmentation techniques have been implemented, including random rotations of up to 20 degrees and horizontal flipping of the images. These augmentations aim to increase the diversity of the training data and help the model learn robust features.

The training set is then organized into batches of 16 images and resized to a target size of 224x224 pixels. The images are represented in RGB color mode to capture the color information. The corresponding labels are encoded in a categorical format, enabling multi-class classification. For the validation data, apply a similar preprocessing pipeline, including rescaling and resizing to the same target size. However, do not apply data augmentation to the test set to ensure an unbiased evaluation of the model's performance. By performing these preprocessing steps, prepare the dataset in a standardized and augmented manner, enabling effective training and evaluation of the model for lung image classification tasks.

Feature Extraction using CNN

A CNN is a substantial deep neural network that models and recognizes stimuli as the brain's visual cortex processes them. One might see a basic CNN model like a mix of the features extractor and the classifier components. The CNN's feature extractor, which consists of a sequence of convolution layers followed by pooling layers, is one of the hidden layers. It aims to discover intricate characteristics and patterns specific to a image class after convolving using a variety of filters. Finally, the classification component makes use of these features.

DenseNet121

The DenseNet121 model has 121 layers and uses a dense block structure, where each layer is connected to all previous layers in a feedforward manner. DenseNet121 is composed of six dense blocks, each consisting of an equal number of sub-layers. These sub-layers include batch normalization, ReLU activation function, and convolution layers. The dense blocks are then followed by transition layers, which utilize 1×1 convolutions and 2×2 average pooling. Global average pooling is applied for classification, followed by a softmax classifier. This leads to a compact model architecture. DenseNets consistently improve accuracy as the number of input parameters increases, without any signs of performance degradation or overfitting. Furthermore, DenseNets achieve good accuracy with significantly fewer parameters and computational resources. This effectiveness stems from their ability to extract essential features by leveraging condensed internal representations and reducing feature redundancy. Its architecture, as shown in Figure 4.11, has been widely adopted, and serves as a basis for many other deep learning models.

Figure 4.11: An illustration of the DenseNet121 model showing four dense blocks. The omitted blocks follow the same structural pattern as the illustrated ones, excluded for space considerations.

Channel Attention

Attention plays a key role in human perception, especially within the visual system, where instead of attempting to process an entire scene at once, humans focus on specific, important areas to effectively capture visual details. This concept has been adopted in deep learning models, where attention mechanisms allow the network to concentrate on the most significant parts of the input, such as certain regions of an

image. In the code provided, incorporate this idea using a channel attention mechanism. Figure 4.12 illustrates the steps involved in this process.

The channel attention mechanism works by utilizing the inter-channel dependencies of the feature maps. Since each channel functions as a detector of specific features, applying channel attention enables the model to enhance the most relevant features in the input image. In the developed solution, after extracting features using DenseNet121, the channel attention layer emphasizes critical channels, guiding the network to focus on meaningful information. To efficiently compute channel attention, the spatial dimensions of the input feature map are reduced through a global pooling operation, ensuring that the model processes only the most essential features from each channel.

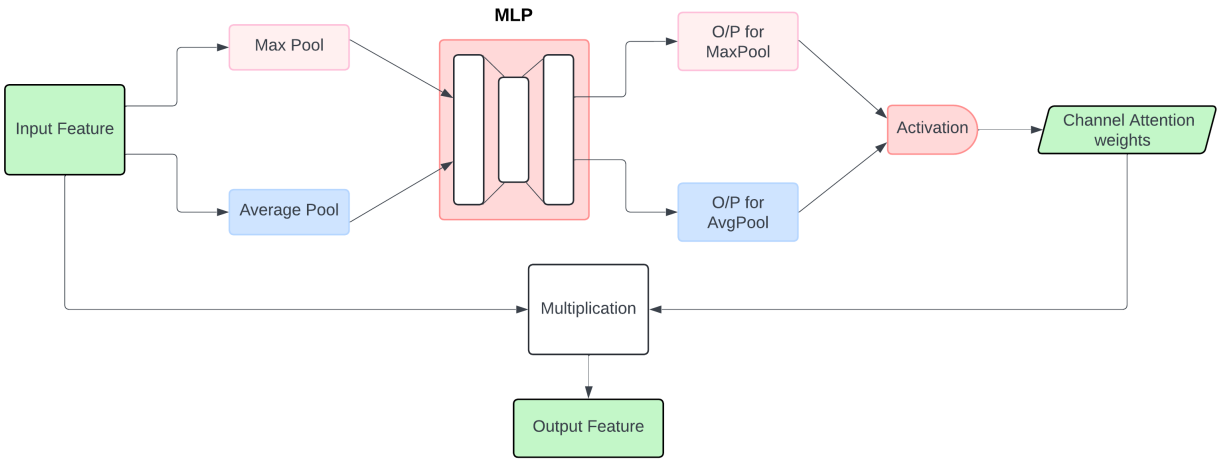


Figure 4.12: An illustration of the channel attention mechanism.

begin by extracting spatial information from the feature map through two pooling operations: average pooling and max pooling. These processes yield two separate spatial context descriptors, denoted as F^c_{avg} for average-pooled features and F^c_{max} for max-pooled features. Both descriptors are then passed through a shared network to generate a channel attention map $M_c \in \mathbb{R}^{C\times1\times1}$.

The shared network consists of a multi-layer perceptron (MLP) with one hidden layer. To minimize the parameter overhead, the hidden layer size is set to $\mathbb{R}^{C/r\times1\times1}$, where r is the reduction ratio. Once the descriptors have passed through the MLP, the resulting output feature vectors are combined using element-wise summation. Finally, the channel attention is computed as:

$$A_c(F) = \sigma(\text{MLP}(\text{GlobalAvgPool}(F)) + \text{MLP}(\text{GlobalMaxPool}(F)))$$
$$= \sigma(V_1(V_0(F_{avg}) + V_1(V_0(F_{max})))$$

In the expression, $A_c(F)$ represents the channel attention map generated from the input feature map F. The terms F_{avg} and F_{max} denote the feature maps obtained from global average pooling and global max pooling, respectively. MLP refers to the multi-layer perceptron used for processing the pooled feature maps. V_0 and V_1 are the learnable weight matrices of the MLP. Finally, σ represents the sigmoid activation function, while GlobalAvgPool and GlobalMaxPool indicate the global average pooling and max pooling operations.

For our work, use DenseNet121 [23] model pre-trained on the ImageNet [25] weights. The layers in DenseNet121 have been kept frozen and the trainable layers include the channel attention module, flattened layer, two dense layers and the output softmax layer. The attention layer is applied after the final convolutional layer of the pre-trained DenseNet121 model. The weights obtained after applying channel attention on the input feature is multiplied with it to obtain the output feature map. The output of the channel attention layer is then flattened, and fed into two fully connected dense layers following which obtain the classification results.

train the model for 25 epochs with a learning rate of 0.01. Adam optimizer is used along with the calculation of cross-entropy loss. After training the models for 25 epochs, the model weights are saved and later loaded in evaluation mode to generate feature space for validation and test sets. For feature generation extract the pre-final layer dense layer.

Feature Selection

A vast dimension of features may be generated by the feature extraction process of any CNN model, which the classifier must process. However, all these features may not be useful for the classification tasks. The remaining elements are superfluous or irrelevant and simply serve to lengthen computation times and take up more physical space. Furthermore, the classification accuracy is lowered by the existence of these non-informative features. On the set of features generated from the aforementioned CNN model, feature selection has been conducted in order to eradicate the said issues. In the developed method, the fitness of each chromosome in the GA population is

evaluated using two distinct filter approaches, making the GA an adaptive feature selection algorithm.

Filter Method

To calculate the fitness the individual chromosomes, rely on the filter-based method, namely Minimum redundancy maximum relevance (mRMR).

Minimum redundancy maximum relevance: mRMR [26] is a filter ranking strategy, which ranks features according to correlation to the class and itself. Preferably, features having a low correlation among themselves and a strong correlation with the class (output) are picked. The F-statistic values [277] for continuous features can be used to assess correlation with the class (relevance), while PCC values can be used to assess correlation between features (redundancy). The F-statistic is calculated by comparing the variance between classes (numerator) with the variance within each class (denominator). It evaluates how well a feature separates class labels, with a higher F-statistic indicating stronger relevance. Conceptually, this compares a "full model," where the feature is included, to a "reduced model," where it is excluded. The between-class variance represents the variability explained by the feature (full model), while the within-class variance reflects the residual variance (reduced model). Thus, the F-statistic measures the added value of including the feature in terms of class separation.

The features are chosen one at a time using a greedy search strategy in order to maximise the objective function, which is based on relevance and redundancy. The objective function that indicates the difference between relevance and redundancy, or the quotient of relevance and redundancy is computed via the formula -

$$score_i(f) = \frac{F(f, target)}{\sum_{s \in f'(i-1)} |corr(f, s)|/(i-1)}$$

Where i represents the i-th iteration, f is the feature under evaluation, F signifies the F-statistic [277], f'(i-1) corresponds to the features chosen through (i-1) iterations, and "corr" indicates the Pearson's correlation.

Overview of the Genetic Algorithm

GA [28] is a well-known meta-heuristic evolutionary algorithm, which is utilized to tackle challenging optimization issues. It is a biologically based algorithm featuring selection, crossover, and mutation properties. Initial population development, parent selection, crossover, mutation, and generation of child chromosomes are all processes in GA. A random population is first created, each chromosome having a finite number of random values of a set length. From this collection of chromosomes, parent chromosomes are chosen, and after crossing and mutation, they are used to produce the offspring chromosomes. In the process of GA, the quality of the solution obtained in each iteration is evaluated by its fitness scores. This process is iteratively repeated, and new sets of chromosomes are generated using selection, crossover, and mutation operations. Individuals with lower fitness are eliminated to make room for new offspring in the subsequent generations.The flowchart consisting of various steps of the GA is shown in Figure 4.13 . The application of genetic algorithm results in a solution that is close to optimal after a predetermined number of iterations. In FS, a binary version of GA is employed, where each chromosome is represented as a vector consisting of '0's and '1's. A '0' indicates the feature is not selected, while a '1' indicates the feature is selected.

GA Variant

GA is an evolutionary algorithm inspired by nature, and it is one of the oldest and most widely used algorithms in the field of FS and optimization. Researchers have utilized GA to achieve near-optimal subsets of features from a given feature space. The algorithm employs key operators such as crossover and mutation to explore and exploit the search space. These operators enable GA to provide a balanced approach between exploration and exploitation. In order to achieve near-optimal solutions, numerous modifications have been suggested by researchers to improve GA. One challenge in GA is the randomness of the mutation probability. Additionally, the fitness of each candidate solution is determined by a learning algorithm, which can be computationally expensive. To address these issues, propose an adaptive version of GA that estimates fitness by aggregating the results of three filter-based methods. This modification significantly reduces computational time. In addition to the fitness estimation modification, propose a different fitness score value calculation method that improves the fitness of individual candidate solutions. use a multi-point crossover and Roulette wheel parent selection to achieve better exploitation. Algorithm 4.1 provides the pseudo code for this fitness score value calculation technique.

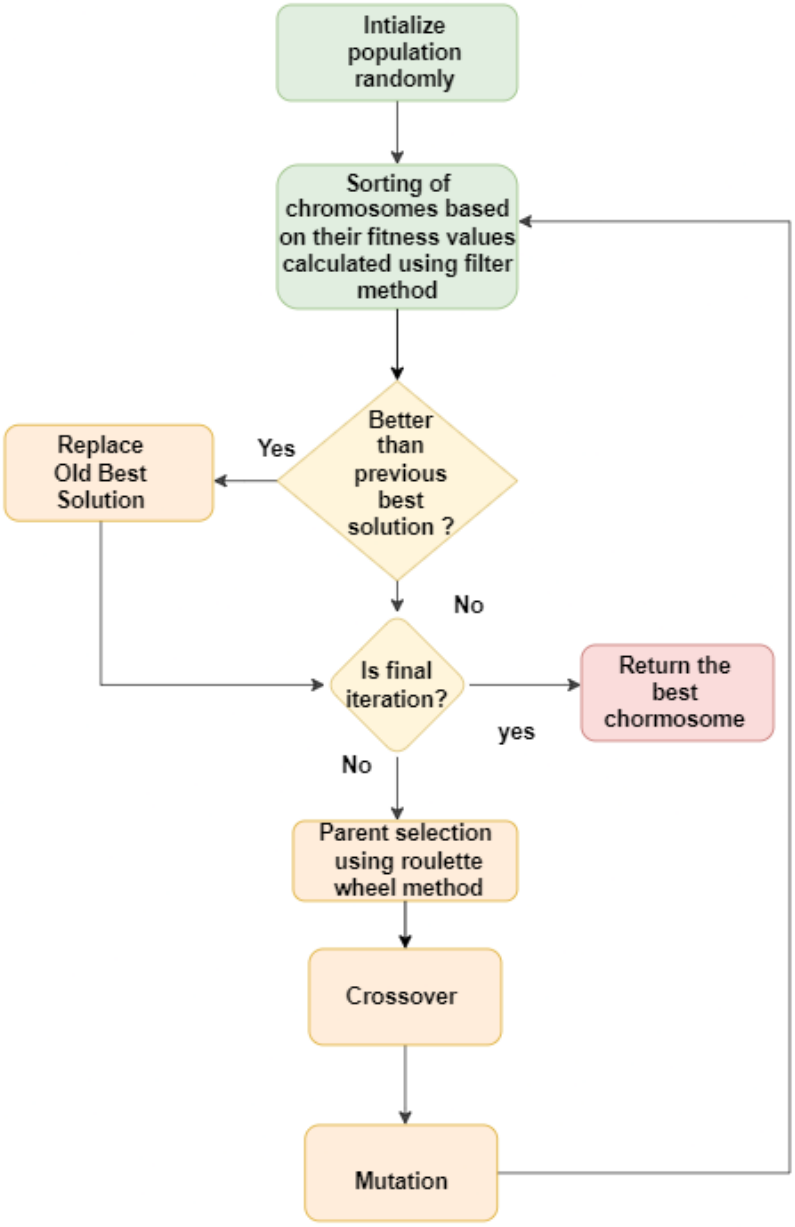


Figure 4.13: Various stages of the Genetic Algorithm

Algorithm 4.1: Score Calculation (using mRMR)

Input: *data*, *target*

Output: *scores*

Compute mRMR scores of features $\rightarrow mrMr$;

Initialize $scores \leftarrow []$;

for $i \leftarrow 1$ **to** b **do**

  $scores_i \leftarrow mrMr_i$;

  Append $scores_i$ to *scores*;

return *scores*;

Fitness Function

Wrapper-based FS methods typically employ a learning algorithm, such as a classifier, to evaluate the fitness of chromosomes. Although GAs are widely used as wrapper-based methods, they tend to increase computational time. To address this issue, in the present work, the use of classifiers is substituted with filter methods to determine the score of each feature vector (i.e., a chromosome). This approach helps assess the strength of each chromosome using a less expensive way, thereby reducing the overall computational complexity of the feature selection method. The algorithm for calculating the mRMR score for a chromosome, provided in Algorithm 4.1, is utilized by the fitness computation module (Algorithm 4.2) which is explained below . This module uses the mRMR scores to evaluate the relevance and redundancy of each selected feature.

Here, provide the population size, maximum number of iterations, and random states (random seed for reproducibility) as parameters. Then from the feature set, take an initial population on a random basis with varied dimensions. Next, pass this initial population no of generations which is numerically equivalent to the max number of iterations. Inside the loop, use a method that sorts the population on the basis of the fitness value of every chromosome which is previously determined for every chromosome by using the filter methods. This step ensures that solutions with better fitness values are given higher priority. A chromosome is basically a binary vector, where '0' denotes that the feature should not be included, and '1' signifies that the feature should be included for fitness calculation. Accuracy is used as part of the fitness value calculation to assess how well a specific subset of features contributes to the model's predictive performance. Agents are evaluated using the fitness function, which considers their accuracy (and potentially other factors) to determine their quality. The population of agents evolves through selection, crossover, and mutation to optimize feature subsets over generations. By employing filter method, obtain a value (i.e., a score) for each feature. It can be said that the feature column with the highest value is the most important, while the one with the lowest value is the least important. Consequently, to compute the score of each individual chromosome, take the mean of these values for all features marked as '1'. The pseudo code for the fitness value calculation is provided in Algorithm 4.2.

Algorithm 4.2: Compute Fitness

Input: $agent, train\_X, test\_X, train\_Y, test\_Y, weight\_acc, score$

Output: $fitness$

$weight\_feat \leftarrow 1 - weight\_acc$;

$num\_features \leftarrow length(agent)$;

`// Compute accuracy based on selected features`

$fitness\_acc \leftarrow 0$;

$number\_selected \leftarrow 0$;

for $i \leftarrow 1$ **to** $num\_features$ **do**

  if $agent[i] == 1$ **then**

    $number\_selected \leftarrow number\_selected + 1$;

    $fitness\_acc \leftarrow fitness\_acc + score[i]$;

if $number\_selected == 0$ **then**

  $acc \leftarrow 0$;

else

  $acc \leftarrow fitness\_acc/number\_selected$;

`// Compute feature importance`

$feat \leftarrow \frac{num\_features - \sum_i agent[i]}{num\_features}$;

`// Combine accuracy and feature importance to get overall fitness`

$fitness \leftarrow weight\_acc \times acc + weight\_feat \times feat$;

return $fitness$;

In FS, our goal is to simultaneously enhance the classification accuracy of the problem at hand and reduce the number of selected features. To achieve this, introduce a single objective function that evaluates the overall fitness of each chromosome (feature subset). This objective function is defined in the following equation-

$$\text{fitness} = \alpha \cdot \text{acc} + (1 - \alpha) \cdot \left(\frac{\text{num\_features} - \sum_i \text{agent}_i}{\text{num\_features}} \right)$$

where the term α represents the weight assigned to the accuracy component of the fitness score, while $(1 - \alpha)$ represents the weight assigned to the feature importance component. The variable acc refers to the accuracy achieved by the agent, which is calculated using both the training and test sets. num_features denotes the total number of features being considered by the model, equivalent to the length of the agent vector. The sum $\sum_i \text{agent}_i$ corresponds to the number of features selected by the agent, where agent is a binary vector indicating feature selection (1 for selected features, 0 for unselected features).

The following are the major stages of GA:

1. Selection: In GA, the selection process determines which individuals from the current population will become parents for the next generation, prioritizing those with higher fitness values for increased chances of selection. Here, the selection process is performed using Roulette Wheel Selection. The fitness values of the solutions are normalized, and selection probabilities are calculated based on the fitness values. A random number is then generated, and a solution is selected with a probability proportional to its fitness value. This process is repeated to select pairs of parents for crossover.

2. Crossover: In GA, crossover combines genetic information from two parent solutions, exchanging features to create new offspring and explore different regions of the solution space. Here, the crossover operation is performed between pairs of parent solutions ('parent1' and 'parent2'). A crossover probability ('crossoverprob') determines the likelihood of performing the crossover operation. For each feature in the solutions, a random number is generated, and if it is less than the crossover probability, the corresponding feature values are swapped between the parents, creating two new child solutions ('child1' and 'child2').

3. Mutation: In GA, mutation introduces random changes to individuals, adding diversity to the population and preventing premature convergence by exploring new areas of the solution space. Here, the mutation is applied to each solution in the population after the crossover operation. The mutation probability ('mutationprob') increases linearly with the current iteration. For each feature in the solution, a random number is generated, and if it is less than the mutation probability, the feature value is flipped (from 0 to 1 or from 1 to 0). The mutation process introduces random changes to the solutions, allowing the algorithm to explore new solutions that may have been missed through selection and crossover operations.

By combining selection, crossover, and mutation, the GA explores and exploits the population of solutions over multiple generations. Solutions with better fitness values have a higher chance of being selected, and genetic material is exchanged and mutated to explore different regions of the solution space. If better than the previous best solution is found in any solution then the old best solution is replaced by the new best solution. This iterative process continues until the maximum number of iterations is reached or a termination condition is met, ultimately converging toward an optimal or

near-optimal solution. After iterations, the best solution (i.e., the best chromosome) is considered as the optimal feature subset, which is then fed to the KNN classifier.

Classification

The feature selection step is followed by classification for which have used a k-Nearest Neighbors (KNN) [278] classifier.

KNN Classifier

The KNN classifier is a non-parametric, instance-based learning algorithm that classifies a data point based on the majority class of its nearest neighbors in a feature space. Given a positive integer k and a query sample, the k-NN algorithm identifies the k training examples that are closest to the sample based on a distance metric (typically Euclidean distance). The query sample is then assigned the label that is most frequent among the k nearest neighbors. The k-NN classifier assumes that similar data points exist in close proximity in the feature space, making it suitable for pattern recognition and classification tasks where decision boundaries are complex and non-linear. This method requires no explicit training phase, and its performance is dependent on the value of k, the choice of distance metric, and the underlying distribution of data.

For our work have set the k value to 5. The classifier was fitted using the train data generated by the GA applied on the extracted features, following which the fitted model was tested based on the test dataset.

4.4 Experimental Results for Method 1

This section provides a comprehensive evaluation, comparing the performance of the developed ensemble model with the base models. It starts with a detailed dataset description, followed by evaluation metrics, an analysis of base model performance, and finally, a showcase of the developed ensemble mode's advancements.

In this segment, have provided the results that have acquired after working on the LIDC-IDRI dataset. The distribution of images in the LIDC-IDRI dataset used in this study is shown in **Table 4.3**. A detailed explanation of the evaluation metrics, performance of the base transfer learning models, and the final ensemble model is provided.

Table 4.3: Distribution of images in the LIDC-IDRI dataset used in this study in the training, testing, and validation sets.

	Division	Class	No. of Images
	Train	Benign	4342
		Malignant	845
	Test	Benign	1340
		Malignant	282
	Validation	Benign	1073
		Malignant	224

4.4.1 Performance of Base Models

The ImageNet dataset has been used for pretraining each of the three base models. All model weights, with the exception of the classification layers, have been frozen. The output layer of the classification layers initially had softmax activation of size (1, 1000), which is fine-tuned to (1, number of classes).

On the LIDC-IDRI dataset, each model is trained for exactly 10 epochs, after which the best validation accuracy is considered. Adam optimizer is used for gradient descent with a learning rate of 0.0001. The final values for the developed ensemble model have been empirically selected after testing the training process with various learning rates, batch sizes, and epoch counts. The ResNet-152 model has achieved 94.88% accuracy, the EfficientNet-B7 model achieved 95.93% accuracy, and the DenseNet-169 model achieved 96.86% accuracy on the LIDC-IDRI dataset.

4.4.2 Performance of the developed Ensemble Model

The developed ensemble model predicts the labels from the corresponding testing dataset with 97.16% accuracy and after optimizing the weights distributions of the base models it predicts with 97.23% accuracy, which refers to the ensemble approach's ability to reduce each CNN model's errors and provides more accurate classification results.

Fig 4.16 shows the confusion matrix obtained on the LIDC-IDRI dataset. The class-wise metrics of the developed ensemble model are given in **Table 4.4**

The model is trained for 30 epochs and then use the early stopping criterion by monitoring the model's performance on a validation set during training. Early stopping involves halting the training process when the validation loss or accuracy starts to

Table 4.4: Class-wise Metrics of the developed Ensemble Model. Metrics are in %.

Metrics	Benign	Malignant
Accuracy	98.58	90.78
Precision	98.07	93.09
Recall	98.58	90.78
F1	98.33	91.92

deteriorate or reaches a plateau. This prevents the model from overfitting to the training data. Figure 4.14 shows the learning curves for the LIDC-IDRI dataset using the ResNet-152, DenseNet-169 and EfficientNet-B7 models. The graphs indicate that the models do not acquire any new information and reach its peak accuracy around the tenth epoch. Thus, train our model for 10 epochs.

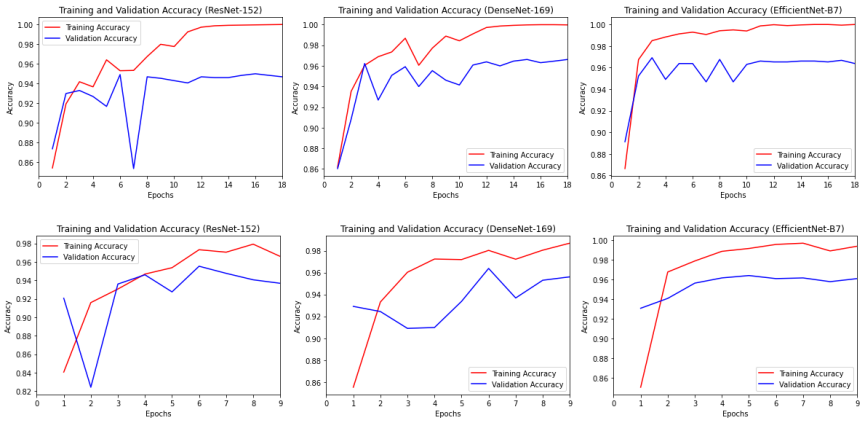


Figure 4.14: Training accuracy vs. Validation accuracy graphs plotted for 18 epochs (upper) and 9 epochs (lower) on the LIDC-IDRI dataset.

The receiver operating characteristic curve, or ROC curve, is used to visualize the classification performance of our ensemble model. The receiver operating characteristic curve, or ROC curve, is used to show how well our ensemble model performs in categorizing data. The measure of a model's ability to differentiate across classes is the ROC curve. The accuracy with which a class is correctly categorized is inversely correlated with the area under the ROC curve. In the ROC curve, the True Positive Rate is pitted against the False Positive Rate. **Fig 4.15** shows the ROC curve for our ensemble model on the LIDC-IDRI dataset. The area under the ROC curve is 94.68%, which is implied by the accuracy of 97.23% achieved on the LIDC-IDRI dataset.

4.4.3 GradCAM analysis

GradCAM++, an enhanced gradient-weighted class activation map [279] is used in this study to provide a visual representation of the model predictions' explainability. The

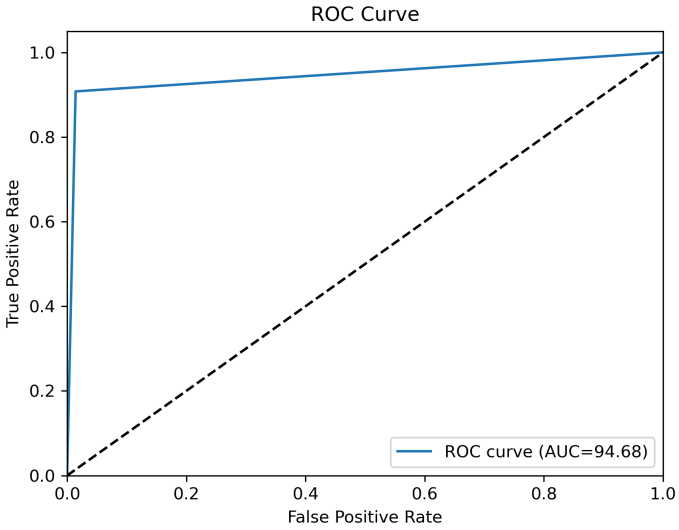


Figure 4.15: ROC curve after ensemble-based classification and weight optimization for the LIDC-IDRI dataset.

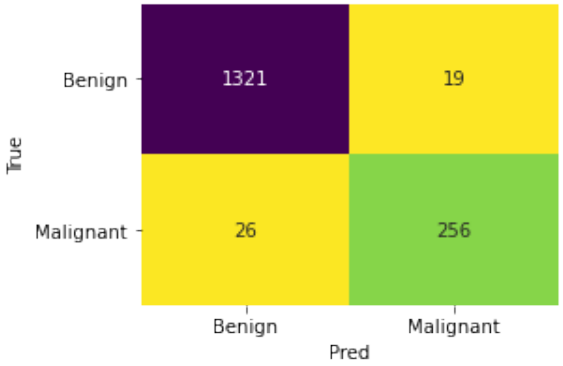


Figure 4.16: Confusion matrix obtained on the LIDC-IDRI CT scan dataset by the developed method.

GradCAM++ explanations provide a visual representation of how a neural network architecture makes decisions. **Fig 4.17** and **Fig 4.18** show GradCAM++ decision visualization of the Malignant and Benign lung nodule CT images respectively, taken from the LIDC-IDRI dataset with the three chosen base models that are used to build the ensemble.

The accuracy rates attained by the transfer learning-based base models using various optimizers on the LIDC-IDRI dataset are shown in **Fig 4.19**. The Adam optimizer is chosen to train the base models for the ensemble framework because it produces the best results for all three base models. **Table 4.5** demonstrates the outcomes of the several ensemble combinations of three different base models (including recently developed framework), GoogLeNet, ResNet-101, ResNet-50, EfficientNetB7, ResNet-152, DenseNet-169, EfficientNetV2L, MobileNet v2, InceptionV3 and NasNetMobile, on the

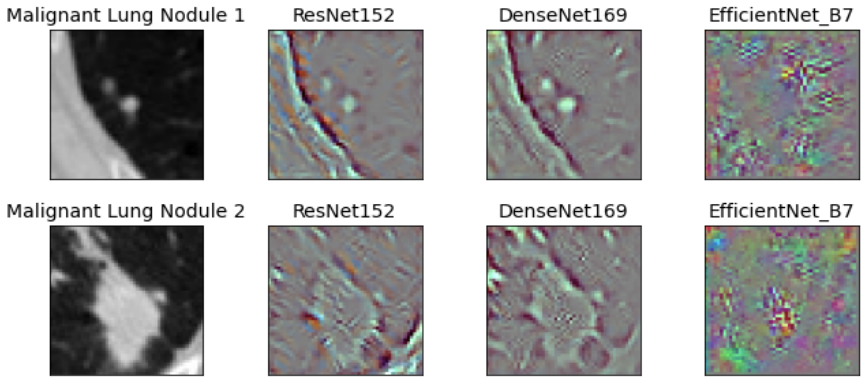


Figure 4.17: GradCAM++ decision display of Malignant lung nodule CT images from the LIDC-IDRI dataset using the three base models chosen for the ensemble.

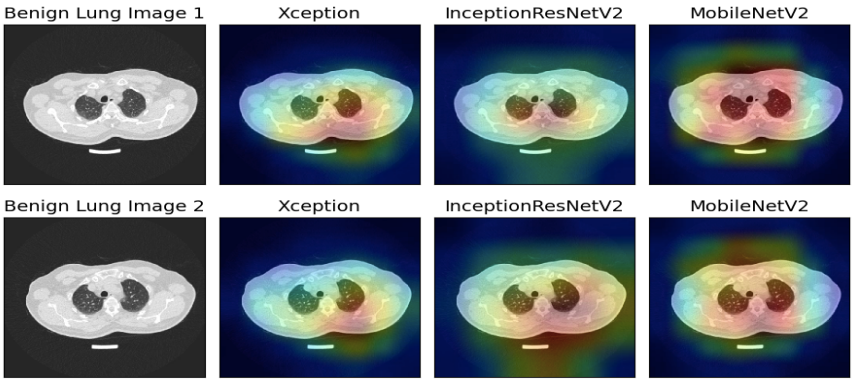


Figure 4.18: GradCAM++ decision visualization of the Benign lung nodule CT images taken from the LIDC-IDRI dataset with the three chosen base models that are used to form the ensemble.

LIDC-IDRI dataset. The results support the choice of the combination of base models used in this study, EfficientB7, ResNet-152, and DenseNet-169. The accuracy rate for the ensemble combination is 97.23%. The combination of DenseNet-121, ResNet-101, and GoogLeNet yields the second-best result, with an accuracy rate of 97.17%.

4.4.4 Comparison with state-of-the-art methods

Table 4.6 displays a performance comparison between the developed ensemble method and well-known ensemble methods. from the literature, using the LIDC-IDRI dataset. The ensembles in the comparison employ the same base models, namely ResNet152, EfficientNetB7, and DenseNet169.

Table 4.7 displays the performance comparison of existing methods in the literature with the developed ensemble model on the LIDC-IDRI lung cancer dataset. It should be highlighted that the developed framework outperforms all existing approaches. Additionally, it is noted that some of these earlier techniques ([264], [280],

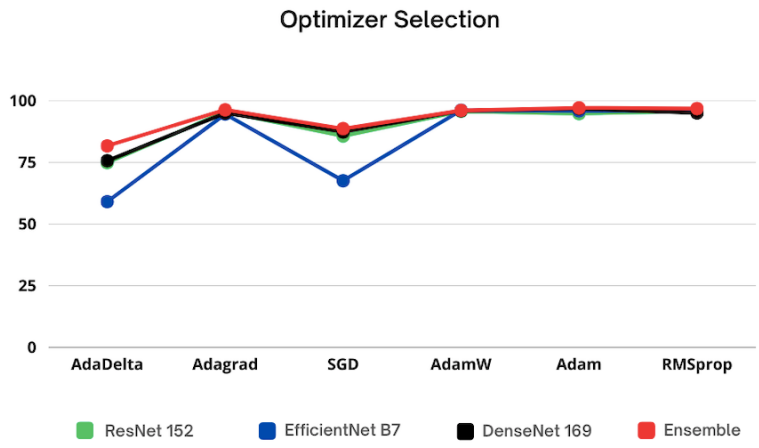


Figure 4.19: Variation in accuracy rates of EfficientNet-B7, DenseNet-169, ResNet-152, and their ensemble on the LIDC-IDRI dataset with different fine-tuning optimizers.

[281], and [26]) focused on employing just one CNN model for identifying the malignant lung tissue on CT scans, the suggested ensemble technique outperforms them, demonstrating that the ensemble method developed in this study is a dependable strategy for the image classification problem under consideration. have compared the performance of the developed method to that of several baseline CNN models for the given dataset.

The comparison of several standard CNN models with the developed ensemble framework on the LIDC-IDRI dataset used in this study is shown in **Table 4.8**. The results clearly demonstrate the superiority of the developed ensemble method over individual models, including AlexNet, GoogLeNet, DenseNet-121, ResNet-152, and EfficientNet-B7. Across multiple performance metrics, such as Recall, AUC, Precision, F1-score, and Accuracy, the developed ensemble method consistently outperforms the individual models. The developed ensemble method achieves the highest overall performance, exhibiting a Recall of 98.07%, AUC of 94.68%, Precision of 98.58%, F1-score of 98.33%, and Accuracy of 97.23%. These findings provide strong evidence for the effectiveness of the ensemble approach in significantly enhancing the classification accuracy of lung nodules on the LIDC-IDRI dataset, surpassing the performance of individual models, and indicating its potential for improving diagnostic accuracy in the field of biomedical imaging.

4.4.5 Statistical Analysis

The McNemar's statistical test which is a non-parametric test has been performed, to find the statistical significance of the outcome of the developed ensemble scheme. The base models, DenseNet-169, ResNet-152, and EfficientNetB7, are compared to the

Table 4.5: Tabular view of experimental results to determine base classifiers to form the ensemble

Model-A	**Model-B**	**Model-C**	**Acc(%)**	**Rec(%)**
EfficientNetB7	MobileNetv2	ResNet-152	96.57	96.57
NasNetMobile	MobileNetv2	ResNet-101	96.16	96.17
MobileNetv2	ResNet-50	DenseNet-169	97.21	97.24
MobileNetv2	ResNet-50	DenseNet-201	95.8	95.89
InceptionV3	ResNet-152	DenseNet-201	94.78	94.78
NasNetMobile	EfficientNetV2L	DenseNet-169	95.21	95.2
EfficientNetB3	ResNet-152	DenseNet-201	92.46	92.88
NasNetMobile	ResNet-50	DenseNet-169	96.14	96.46
NasNetMobile	EfficientNetB7	DenseNet-201	92.97	92.89
GoogLeNet	ResNet-101	DenseNet-121	97.21	97.15
EfficientNetV2M	ResNet-152	DenseNet-201	95.02	95.06
GoogLeNet	ResNet-50	DenseNet-201	97.1	97.12
GoogLeNet	MobileNetv2	DenseNet-121	97.16	97.19
GoogLeNet	ResNet-18	MobileNetv2	97.14	97.54
EfficientNetV2L	MobileNetv2	EfficientNetV2M	97.12	97.02
EfficientNetB7	**ResNet-152**	**DenseNet-169**	**97.23**	**98.07**

suggested ensemble framework in this study, and the probability scores are utilized to determine the ensemble's creation. **Table 4.9** tabulates the results of McNemar's test on the LIDC-IDRI CT scan dataset used in this study. The McNemar's test's p-value must be less than 0.05 (5%) in order to find statistical evidence against the null hypothesis; according to **Table 4.9** the p-value is less than 0.05 for the used dataset. Thus, the null hypothesis is rejected by the results of this statistical test which establishes that the developed ensemble model captures complementary information from the base models and that its predictions are superior. As a result, it makes sure that none of the contributing base models are statistically similar to the ensemble classifier.

4.4.6 Error Analysis

In **Fig 4.20**, two sample images from the LIDC-IDRI dataset are displayed. The two base classifiers make inaccurate predictions with a low confidence value, whereas the third base model makes an accurate prediction with a high confidence value. However, our developed ensemble framework predicts the sample correctly. **Fig 4.20**(a) shows a sample where ResNet-152 predicts "Malignant" with a confidence score of 63.1%,

Table 4.6: Performance comparison of the popular ensemble schemes in the literature with the developed ensemble method for the LIDC-IDRI dataset. All ensembles utilise the same base models: ResNet152, EfficientNetB7, DenseNet169.

Ensemble Method	**Acc(%)**	**Pre(%)**	**Rec(%)**	**F1(%)**	**AUC(%)**
Weighted Average (Acc only)	97.18	97.02	97.29	98.08	93.89
Average Probability	97.15	97.11	97.19	97.15	94.51
Maximum Probability	97.07	97.04	97.09	97.06	93.11
Majority Voting	96.89	96.11	96.6	96.35	94.62
developed **Ensemble**	**97.23**	**98.58**	**98.07**	**98.33**	**94.68**

Table 4.7: Performance comparison of the developed ensemble model with state-of-the-art methods for the LIDC-IDRI dataset.

Work Ref.	**Method**	**Acc (%)**
[264]	Forward and Backward GAN and Multi-scale VGG16	95.20
[282]	Deep Residual Learning	84.00
[261]	Resolved Ambiguity Local Binary Pattern	94.90
[263]	Semi-supervised Adversarial Classification mode	92.50
[280]	Novel DCNN	94.00
[262]	Convolutional block attention module	90.70
[281]	Deep Learning-Assisted SVM-Based Model	94.00
[26]	2-Pathway Morphology-based CNN	96.10
[209]	LCD-CapsNet	94.00
Developed	**An ensemble of three CNN models**	**97.23**

DenseNet-169 predicts “Malignant” with a confidence score of 75.3%, and EfficientNetB7 predicts “Benign” with a confidence score of 79.4%. With a confidence score of 79.4%, the developed ensemble framework eventually properly predicts that the sample belongs to the ”Benign” class. Similarly, in the case of **Fig 4.20**(b), ResNet-152 predicts “Benign” with a confidence score of 88.9%, EfficientNet-B7 predicts “Ma-

Table 4.8: Comparison of several standard CNN models with the developed ensemble framework on the LIDC dataset.

Model	**Acc(%)**	**Pre(%)**	**Rec(%)**	**F1(%)**	**AUC(%)**
AlexNet	97.17	97.22	97.18	97.19	93.17
GoogLeNet	96.89	97.12	97.12	97.12	93.89
DenseNet-121	96.23	96.63	96.24	96.31	94.23
ResNet-152	97.19	98.31	97.29	98.3	94.29
EfficientNet-B7	97.11	97.12	97.09	97.14	95.09
Developed	**97.23**	**98.58**	**98.07**	**98.33**	**94.68**

Table 4.9: McNemar's statistical test results for the ensemble model and the base models on the LIDC-IDRI dataset. The p-value is less than 0.05 for the base models against which the developed model is compared, and hence the null hypothesis is rejected.

McNemar's Statistical Test	**p value**
Base Model	**Dataset (LIDC-IDRI)**
ResNet-152	0.0018
DenseNet-169	0.0027
EfficientNet-B7	0.0431

lignant" with a confidence score of 68.3%, and DenseNet-169 predicts "Malignant" with a confidence score of 71.3%. Finally, with a confidence score of 76.4 %, the suggested ensemble framework accurately predicts that the sample will be "Benign". This demonstrates how effective the ensemble architecture put forth here is.

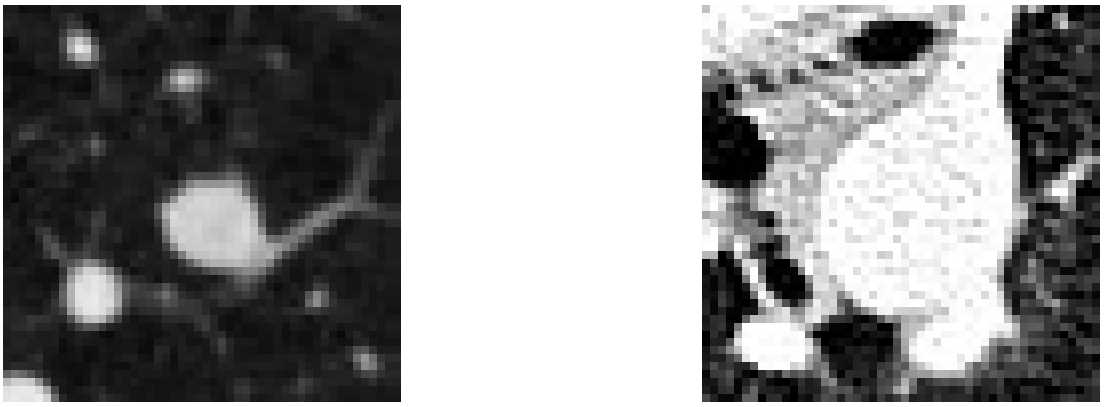


Figure 4.20: Samples taken from the LIDC-IDRI dataset where 2 of the 3 base models produce inaccurate predictions while the ensemble method produces the correct prediction. They both belong to the "Benign" category. **(a) Case-1:** DenseNet-169 predicts "Malignant" with a confidence score of 75.3%, EfficientNet-B7 predicts "Benign" with a confidence score of 79.4%, and ResNet-152 predicts "Malignant" with a confidence score of 63.1%. With a confidence rate of 74.1%, the suggested ensemble framework predicts "Benign" (right categorization) **(b) Case-2:** ResNet-152 predicts "Benign" with a confidence score of 88.9%, EfficientNet-B7 predicts "Malignant" with a confidence score of 68.3%, and DenseNet-169 predicts "Malignant" with a confidence score of 71.3%. With a confidence rate of 76.4%, the suggested ensemble framework predicts "Benign" (right categorization).

4.5 Experimental Results for Method 2

In this section, we present the detailed experimental results along with an in-depth analysis of the developed CNN-based ensemble model for lung cancer detection in CT scans. The distribution of images in the dataset is provided in Table 4.1. The implications of the obtained results are also discussed. Furthermore, a comparative evaluation is conducted to demonstrate that the developed method outperforms existing individual models as well as commonly used ensemble techniques reported in the literature.

Table 4.1: Distribution of images found in the IQ-OTHNCCD dataset into Benign, Malignant, and Normal classes.

Dataset	Class	No. of images
IQ-OTHNCCD	Benign	120
	Malignant	561
	Normal	416

4.5.1 System Configuration

The entire set of experiments has been run on a Jupyter Notebook equipped with a 12 GB NVIDIA Tesla T4 GPU that has been made available by Google's collaboratory environment. For the experiments, have implemented our developed method using Python 3, and the primary open-source modules used are Tensorflow, Keras, Matplotlib, Scikit, Numpy, and Pandas.

4.5.2 Evaluation Metrics

Evaluation metrics are important in assessing the effectiveness and strength of a predictive/learning model. In other words, these metrics help to quantify how well the model performs in terms of its ability to predict outcomes accurately. It is important to use a variety of standard evaluation metrics to get a complete picture of the model's performance and to ensure that it meets the requirements of the problem under consideration. In classification problems, these metrics are used to evaluate the performance in predicting the correct class label for a given input. Table 4.2 shows various performance measures used in our classification method. Consider there is a two-class classification problem, where one class is termed as 'positive' and another one is termed as 'negative'. Most of these measures are computed using a confusion matrix employing four fundamental components, such as true positive (T^P) rate, true negative (T^N) rate, false positive(F^P) rate, and false negative(F^N) rate.

Table 4.2: Evaluation matrix used in our classification methods

Metric	**Formula**
Accuracy(Acc)	$\frac{T^P+T^N}{T^P+T^N+F^P+F^N}$
Precision(Pre)	$\frac{T^P}{T^P+F^P}$
Recall(Re)	$\frac{T^P}{T^P+F^N}$
F1-Score(F1)	$\frac{T^P}{T^P+\frac{1}{2}(F^P+F^N)}$

Initially, perform extensive experimentation with different combinations of CNN models to determine the best combination of base learners for our developed ensemble technique. The hyperparameters selected for this experiment are mentioned in Table 4.3. The results of the experimentation are shown in Table 4.4.

As evident from the results, get the best accuracy with a combination of Xception, InceptionResnetV2, and MobileNetV2 for forming the ensemble approach. Thus, select these three models as our base learners. The results given by these models along with some additional information are shown in Table 4.5.

Table 4.3: Hyperparameters of the base classifiers

Hyperparameter	**Value/Name**
Optimizer	Adam
Loss function	Sparse Categorical Cross Entropy
Learning rate	0.001
No. of epochs	60

Table 4.4: Results of experiments implemented to determine the base classifiers for forming the ensemble in this study.

Model 1	**Model 2**	**Model 3**	**Accuracy (%)**
NasNetMobile	MobileNet	ResNet-152	95.98
NasNetMobile	MobileNetV2	ResNet-50	94.87
NasNetMobile	MobileNetV2	DenseNet-169	95.21
NasNetMobile	MobileNetV2	InceptionResNetV2	92.56
MobileNetV2	ResNet-152	DenseNet-169	96.41
MobileNetV2	ResNet-50	DenseNet-169	92.99
MobileNetV2	ResNet-50	InceptionResNetV2	97.17
MobileNetV2	ResNet-152	InceptionResNetV2	96.67
NasNetMobile	ResNet-152	DenseNet-169	97.00
NasNetMobile	ResNet-152	InceptionResNetV2	96.41
NasNetMobile	ResNet-50	DenseNet-169	96.06
NasNetMobile	ResNet-50	InceptionResNetV2	96.92
Xception	ResNet-152	DenseNet-121	97.77
Xception	ResNet-152	InceptionResNetV2	95.04
Xception	ResNet-18	InceptionResNetV2	98.20
Xception	MobileNet	DenseNet-121	98.29
Xception	ResNet-18	MobileNetV2	98.54
Xception	MobileNet	NasNetMobile	98.12
Xception	**InceptionResNetV2**	**MobileNetV2**	99.45

Table 4.5: Performance measure of each model along with their total number of parameters

Model	**No. of parameters**	**Acc (%)**	**Pre (%)**	**Re (%)**
Xception	20,867,627	99.02	99.38	99.59
InceptionResNetV2	54,341,347	97.26	88.33	93.80
MobileNetV2	2,261,827	99.45	99.59	99.79
Developed Ensemble	**77,470,801**	**99.54**	**99.62**	**98.61**

These three models, i.e., Xception, InceptionResNetV2, and MobileNetV2 give us an accuracy score of 99.02%, 97.26%, and 99.45%, respectively. The results show that these models are very reliable for being chosen for this fuzzy ensemble. In our research, the Xception model has 20.87M parameters, the InceptionResNetV2 has 54.34M parameters, and the MobileNetV2 has only 2.26M parameters.

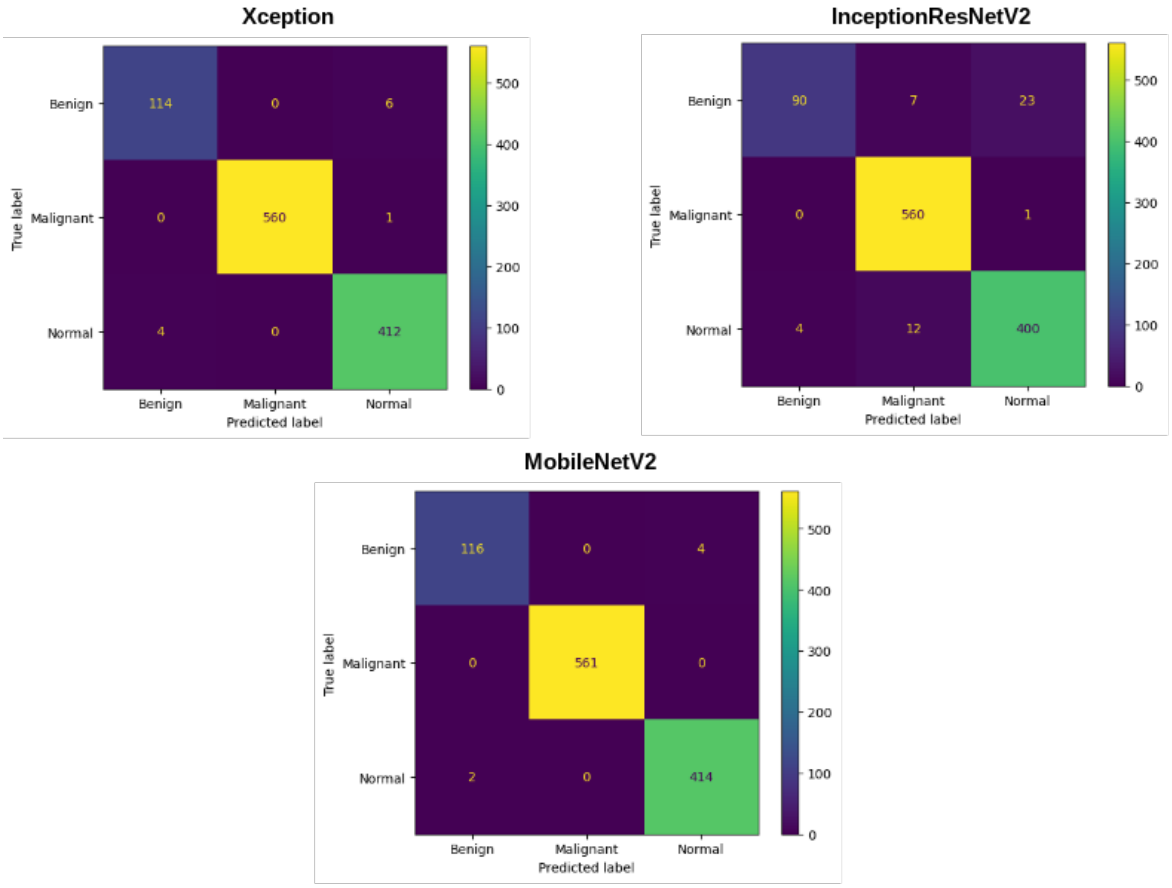


Figure 4.21: Confusion matrix obtained on the IQ-OTHNCCD dataset by Xception, InceptionResNetV2, and MobileNetV2 respectively.

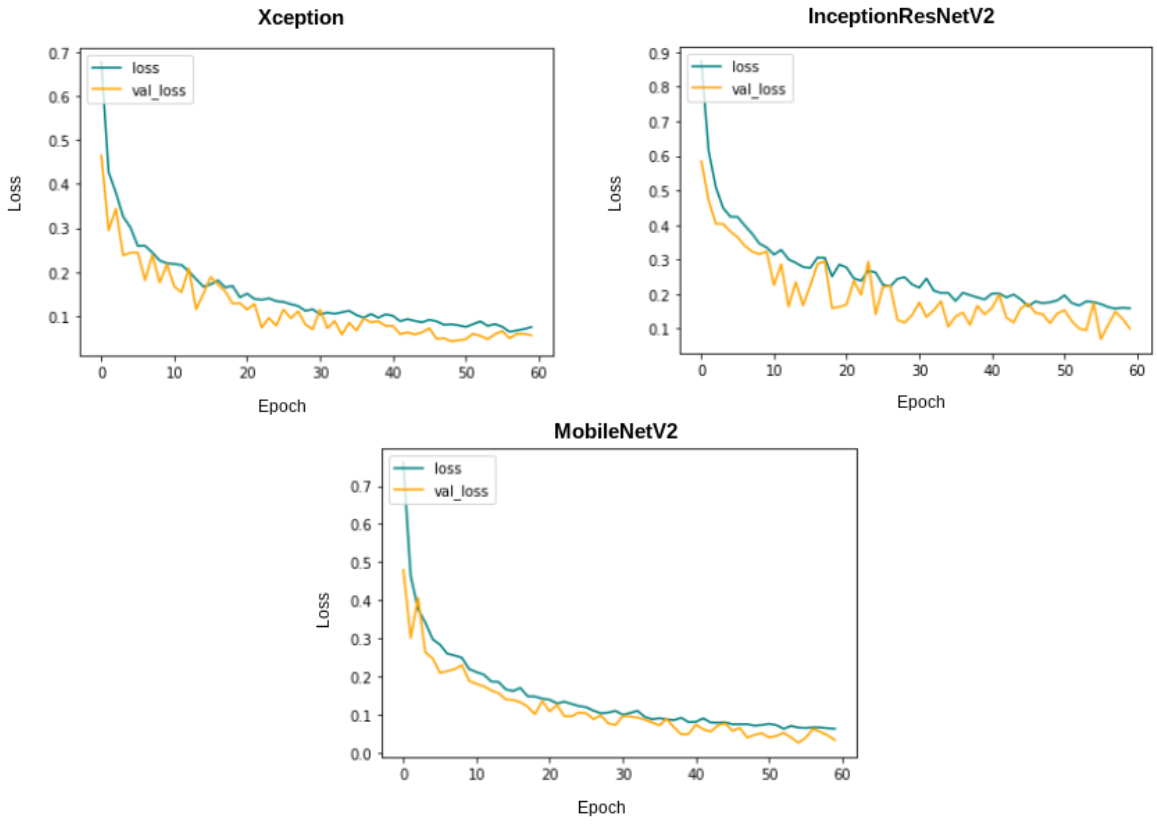


Figure 4.22: Training and validation loss plots for 60 epochs on the IQ-OTHNCCD dataset

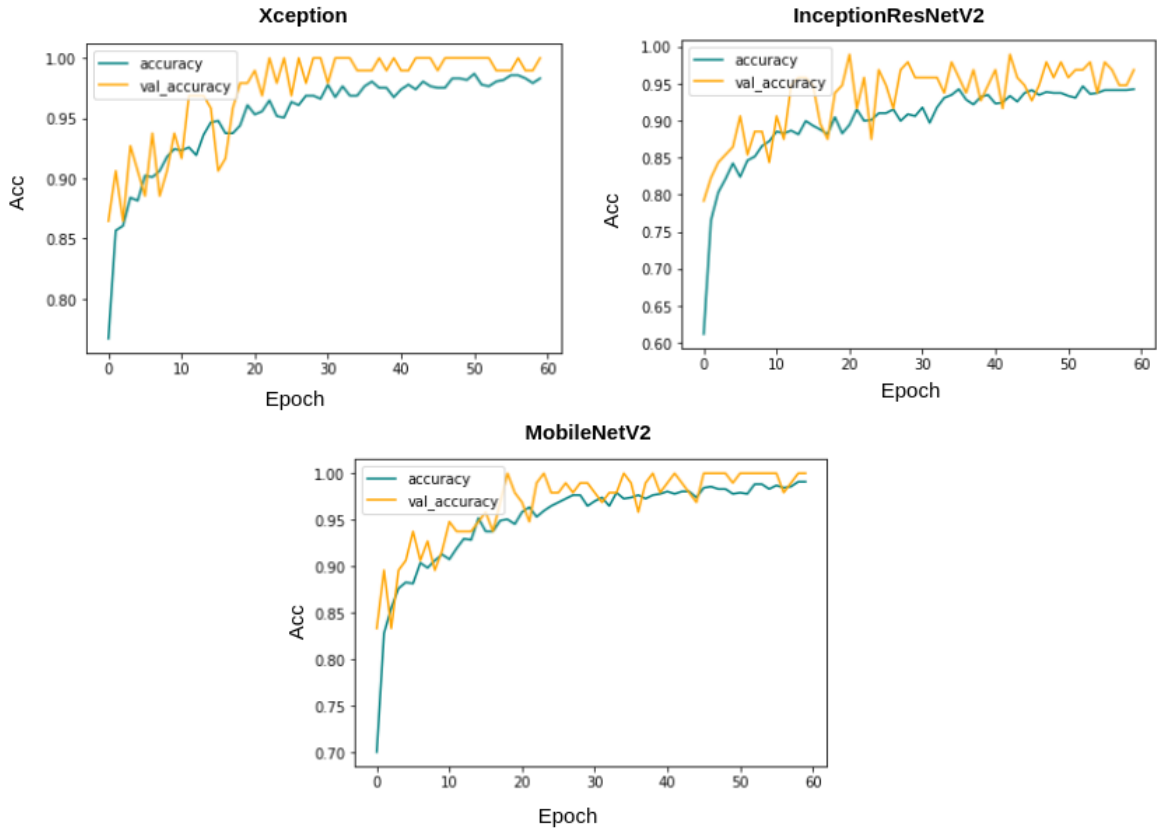


Figure 4.23: Training and validation accuracy plots for 60 epochs on the IQ-OTHNCCD dataset

Table 4.6: Class-wise and overall results obtained by the developed ensemble model.

Class	Pre	Re	F1	Support	Acc	Overall Acc	Overall Re	Overall F1
Benign	1.0000	0.9583	0.9787	120	0.95833333			
Malignant	0.9982	1.0000	0.9991	561	1.	0.9954	0.9861	0.9910
Normal	0.9905	1.0000	0.9952	416	1.			

Each of these base models generates the confidence scores for each class for every single image in the dataset. These confidence scores of all the base models for each image are generated and stored for each individual classifier. For the results obtained in Table 4.5, the three transfer-learned models have been trained for 60 epochs with the Adam optimizer individually. Figure 4.21 shows their respective confusion matrices on the dataset used. Although there are some misclassifications, the ratio of misclassifications to that of the correct classifications is very low, indicating that these models are reliable for the work. The loss curves obtained by each of the base learners(or classifiers) are shown in Figure 4.22. As our models are all transfer learning-based models that have been pre-trained on the ImageNet dataset, just had to fine-tune the models on our IQ-OTHNCCD lung cancer dataset. From the loss curves of the three models, can see that there is hardly any problem with overfitting in the three models. The accuracy curves as shown in Figure 4.23 display the corresponding accuracy scores as shown in Table 4.5. After fusing the confidence scores of the three models using the developed Mitscherlich function-based ensemble model, get the results.

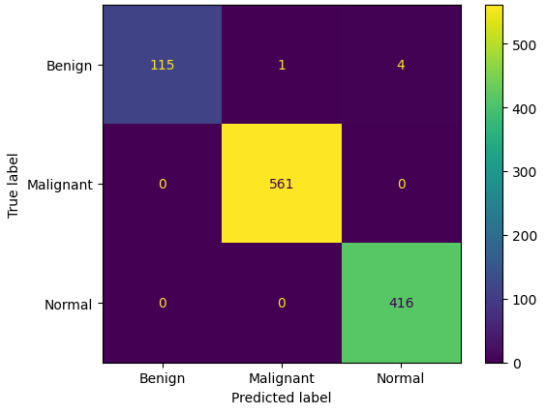


Figure 4.24: Confusion matrix of developed ensemble using Mitscherlich function (MENet)

From Figure 4.24, Figure 4.25 and Table 4.6, can observe that the overall accuracy achieved is 99.54% after combining the results of the three base classifiers. The class-wise results have also been given in the Table 4.6. All these high classification accuracies indicate that the model is highly reliable. The final ROC curve obtained on this dataset shows that the False Positive rate is significantly low. This indicates that

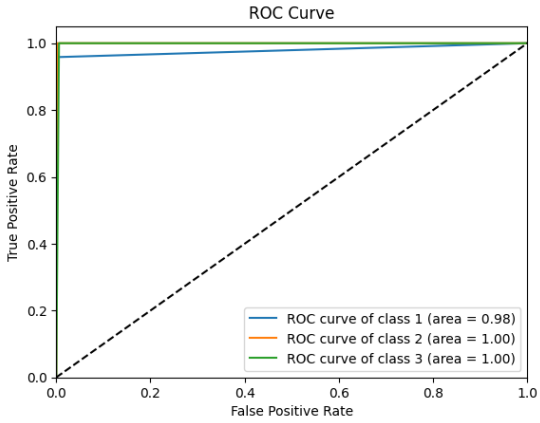


Figure 4.25: ROC-AUC curve obtained by the developed ensemble model (MENet)

this developed fuzzy ensemble method is very efficient for the problem under consideration. The Confusion Matrix of this developed fuzzy ensemble method, as shown in Figure 4.24 also shows very few mispredictions than all the three base classifiers. The area under the curve (AUC) is shown in Figure 4.25. Here can see for class 1, i.e., Benign: 98%, for class 2, i.e., Malignant: 100%, and for class 3, i.e., Normal: 100%. As the graph in Figure 4.25 is shifted to the top left corner, it indicates the highly efficient ability of the final model to classify test images into one of the three classes.

4.5.3 Comparison with state-of-the-art methods

The Mitscherlich function-based fuzzy ranking-based ensemble approach is evaluated on a new dataset. Fuzzy ranks are generated for each of the base classifiers, and then these ranks are combined using the ensemble method. The developed method is compared with the methods found in the literature, and it has been observed that the developed method achieves the highest accuracy score among others. In comparison, the dataset have used and compared with others is the same dataset (i.e., the IQ-OTHNCCD lung cancer dataset). Table 4.7 provides an illustration of the comparative study of our suggested model with others.

In our experiments, the developed ensemble approach achieves the best performance accuracy of 99.54% as compared to the other methods or individual classifiers. Out of these, the methods developed by [283,284], and [285] are based on the deep CNN methods and achieve accuracy of 95.71%, 97.00%, and 99.45%, respectively. [286] have attained 98.83% accuracy by using three transfer learning models, VGG16, VGG19, and Xception, with CNN for the classification of lung cancer on the same dataset. [287] and [288] achieve accuracy of 98.58% and 97.38%.

In this section, experimental results are compiled in Table 4.8 to demonstrate

Table 4.7: Performance comparison of the developed ensemble model with state-of-the-art methods on the IQ-OTH/NCCD dataset. Results are in %

Method	Approach	Acc
[283]	AlexNet	95.71
[286]	VGG16,VGG19 and Xception	98.83
[287]	Alexnet and ResNet50	98.58
[288]	Transfer Learning with GoogleNet	94.38
[284]	SMOTE based on 2D-CNN	97.00
[285]	CNN	99.45
MENet	**An ensemble of CNN models**	**99.54**

the developed ensemble scheme's superiority to well-known traditional ensemble techniques. The ensembles used the identical three base CNN learners, Xception, InceptionResNetV2, and MobileNetV2. The suggested ensemble method performed better than several popularly used ensemble schemes. It is clear from the results that the performance of the weighted average ensemble, which only takes the accuracy metric into account when determining the weights, came the closest to matching the performance of the developed ensemble technique. In the majority voting-based ensemble, the class that received the highest votes from the base learners is predicted to be the class of the sample.

Table 4.8: Performance comparison of the developed ensemble method with some commonly used ensemble methods evaluated on the IQ-OTHNCCD dataset. Scores are in % .

Ensemble method	Acc	Pre	Re	F1-score	AUC
Maximum Probability	98.11	98.43	97.79	97.76	97.79
Average Probability	98.82	97.81	97.79	97.78	97.81
Sum Rule	96.55	96.11	96.79	97.47	96.81
Majority Voting	98.11	98.13	98.12	98.10	97.85
Weighted Average	98.92	98.74	98.52	98.20	98.13
MENet	**99.54**	**99.62**	**98.61**	**99.10**	**98.88**

4.5.4 Data Visualization

In this study, GradCAM, a technique that creates a gradient-weighted class activation map as developed by [289], to produce visual explanations of the model predictions, is utilized. These visualizations illustrate how the neural networks make their classification decisions. As shown in Figure 4.26, Figure 4.27, and Figure 4.28, GradCAM was employed to generate class activation maps for Malignant, Benign, and Normal lung CT images from the IQ-OTHNCCD dataset using the three base models incorporated

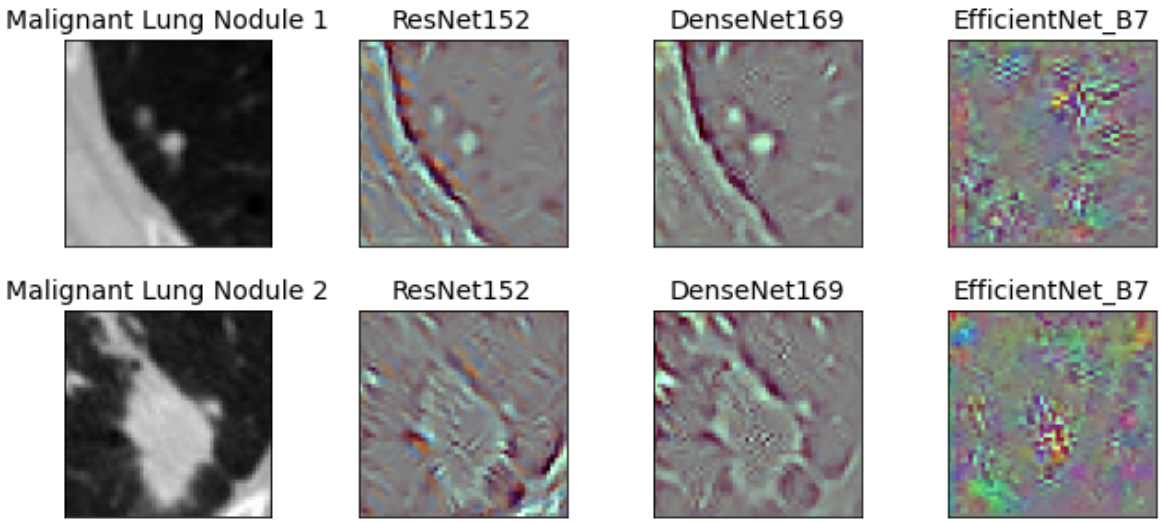


Figure 4.26: GradCAM visualization of the Malignant lung images taken from the IQ-OTHNCCD dataset with the three base models used to form the ensemble

in the developed ensemble. It can be observed that each model focuses on different regions of the lung scans, indicating that the individual learners capture diverse and complementary features essential for ensemble learning. For example, in Figure 4.26, the Xception model emphasizes the right lung region, the InceptionResNetV2 model highlights the lung nodules, and the MobileNetV2 model concentrates on the peripheral areas of the lungs. This diversity in learned representations contributes to the improved effectiveness of the ensemble approach used in this study.

t-SNE plots

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a widely used nonlinear dimensionality reduction technique designed to visualize high-dimensional data in a low-dimensional space, as introduced by [290]. The algorithm begins by converting the pairwise Euclidean distances between data points in the high-dimensional space into conditional probability scores that reflect their similarities. This is accomplished using Stochastic Neighbor Embedding (SNE). The conditional probability $P_{j|i}$, which represents the similarity of data point x_j to data point x_i, is defined in Equation 4.12.

$$P_{i|j} = \frac{exp(\frac{-\|x_i - x_j\|^2}{2\sigma_i^2})}{\sum_{k \neq i} exp(\frac{-\|x_i - x_j\|^2}{2\sigma_i^2})} \tag{4.12}$$

It is observed that the images corresponding to the different stages of lung cancer are clearly separated into different clusters of points. The t-SNE plot visualizations

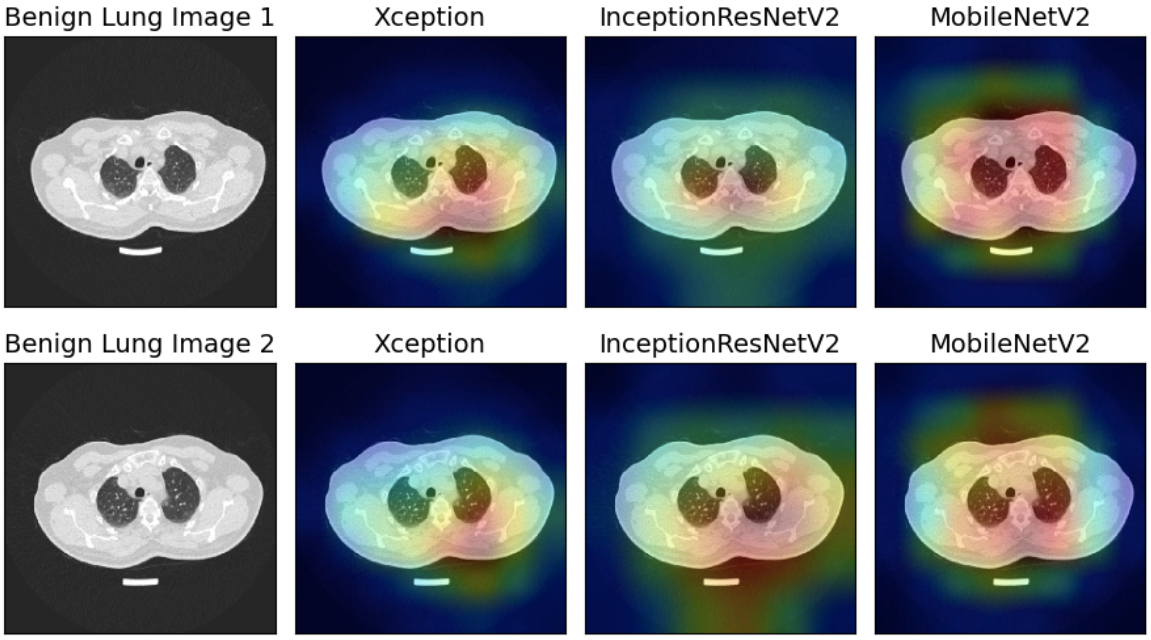


Figure 4.27: GradCAM visualization of the Benign lung images taken from the IQ-OTHNCCD dataset with the three base models used to form the ensemble

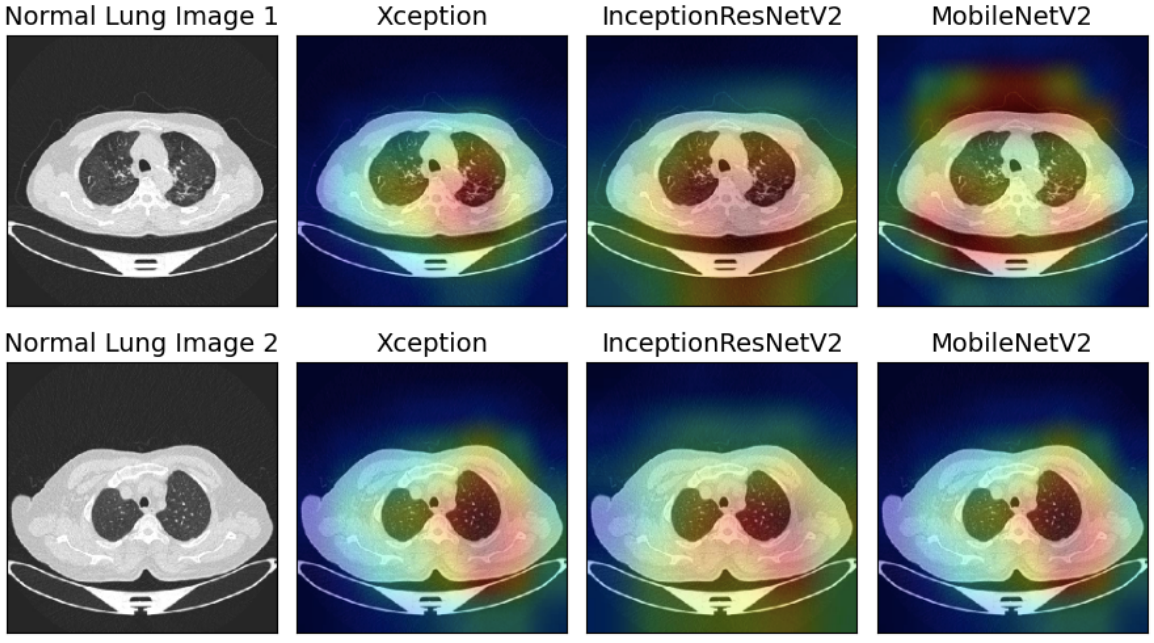


Figure 4.28: GradCAM visualization of the Normal lung images taken from the IQ-OTHNCCD dataset with the three base models used to form the ensemble

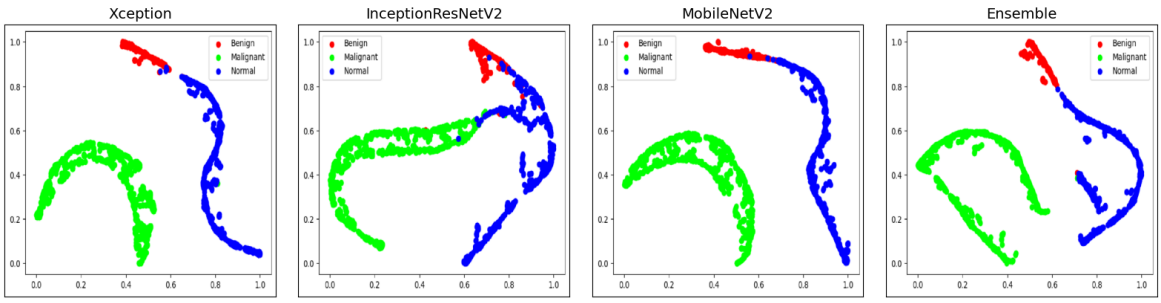


Figure 4.29: Graphical representation of the t-SNE plots of samples of the IQ-OTHNCCD dataset with the three base models and the final ensemble model.

for the Benign, Malignant, and Normal classes of lung nodule CT images from the IQ-OTHNCCD dataset are presented in the first three images of Figure 4.29. These visualizations have been generated using the three selected base models that form the ensemble. The t-SNE plot of the ensemble model is presented in the fourth image of Figure 4.29.

Empirical significance of the ensemble model

In this section, some cases that have arisen during the testing and analysis of our model is reported. Figure 4.30 and Figure 4.31 highlight the cases when for a benign image and a malignant image, respectively, two of our base models have given correct answers, and one has not, but our developed ensemble model is still able to give the correct result. Figure 4.32 highlights the condition where for a normal image, only one of our base models is able to give the correct result, while the other two could not. In this case, our developed ensemble method can also give the correct results. These empirical results ensure the effectiveness of the developed ensemble model in varied scenarios.

Statistical Analysis

The McNemar's statistical test, a non-parametric test, is used to determine the statistical significance of the results obtained by the developed ensemble method. The ensemble result is compared against three base models' results, namely Xception, InceptionResNetV2, and MobileNetV2, on the IQ-OTHNCCD dataset used in this study. The probability scores are used to design the ensemble model. In this case, the null hypothesis is that there is no significant difference in the performance of the three base CNN models in predicting lung cancer outcomes. The results of McNemar's test are presented in Table 4.9, and the p-value for the dataset used in this study is found to be less than 0.05 (5%), indicating statistical evidence against the null hypothesis. The

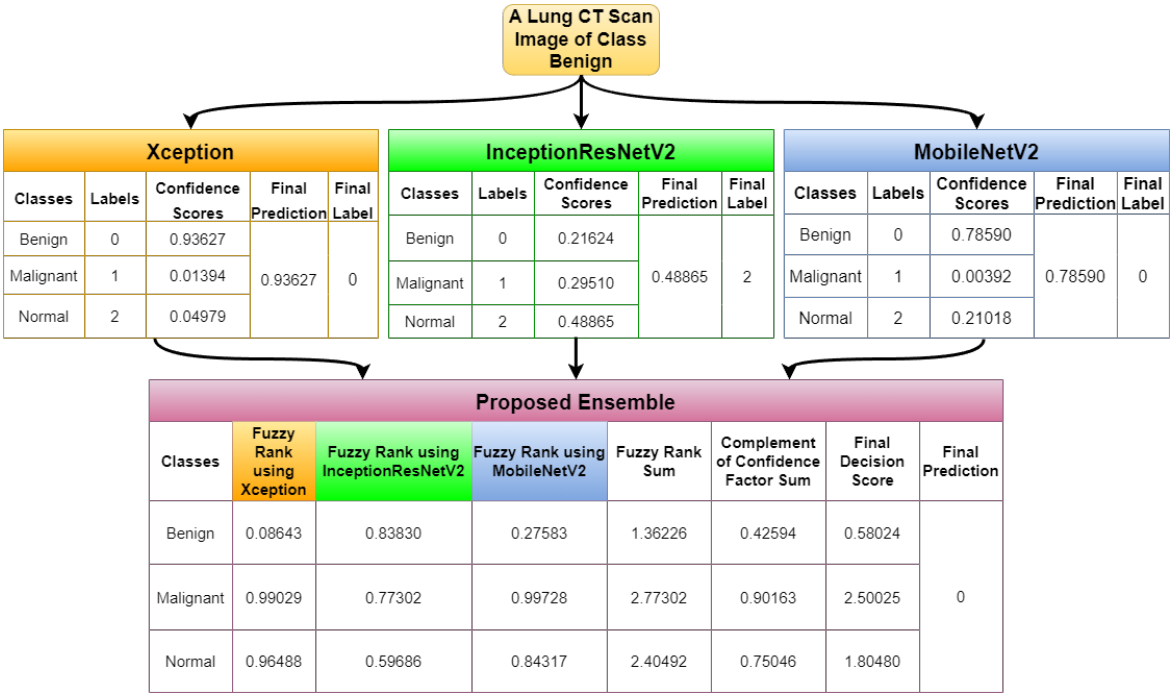


Figure 4.30: Demonstration for the case when two base models give the correct results, but the third one gives an erroneous result for a benign image.

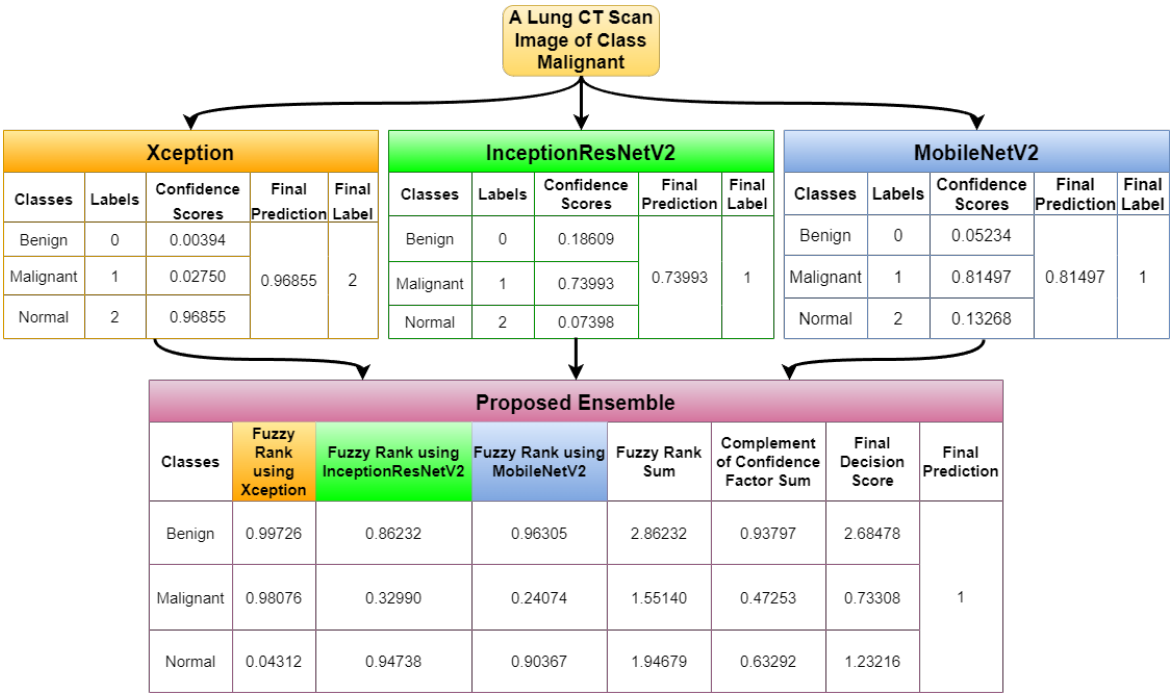


Figure 4.31: Demonstration for the case when two base models give the correct results, but the third one gives an erroneous result for a malignant image.

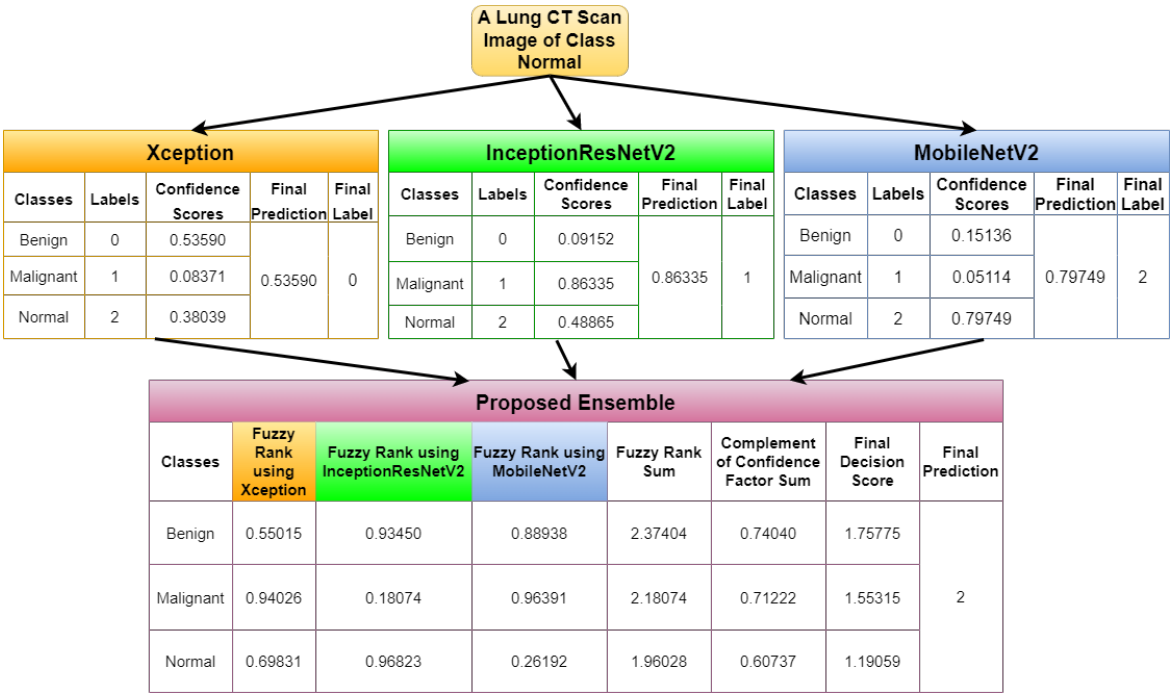


Figure 4.32: Demonstration for the case when one base model gives the correct results, but the remaining two give erroneous results for a normal image.

statistical test confirms that the developed ensemble model captures complementary information from the base models and performs better than individual base models. Consequently, none of the contributing base models are statistically equivalent to the ensemble classifier.

Table 4.9: Results of McNemar's statistical test show that the null hypothesis is rejected for the base models on the IQ-OTHNCCD dataset when compared to the developed ensemble model, as the p-value is found to be less than 0.05.

Base model	p-value
Xception	0.0014
InceptionResNetV2	0.0317
MobileNetV2	0.0005

In recent times, it has been observed that CAD systems can help diagnostic precision and decrease the likelihood of human errors by automating the diagnosis process. In the present work, have designed an ensemble model, called MENet, using three transfer learning-based CNN models, namely Xception, InceptionResNetV2, and MobileNetV2, to enhance the accuracy of a lung cancer prediction model. In doing so, have used a fuzzy ranking-based approach, which considers the uncertainty of each classifier's predictions and assigns different levels of importance to each classifier. The fuzzy ranking system is designed based on the Mitscherlich function, which combines the outputs of the base classifiers to form a final prediction model that is more ac-

curate than each classifier's individual prediction ability. The developed method is evaluated on the open-access lung CT scan dataset, IQ-OTHNCCD, and the obtained results are better than many recently developed methods. However, there are some false positives and false negatives, which is a significant challenge in the medical field as it would directly affect the treatment of patients. Hence, in the future, need to reduce such errors. may apply some attention mechanisms to the base CNN models that might help to generate better feature maps, which, in turn, might produce a better prediction model. In the future, may also implement some lightweight CNN models to make the system more applicable in practical cases.

Dry Run of the entire developed method

In this section, a dry-run demonstration of the developed fuzzy-ranking-based ensemble method is shown, applied for lung cancer detection using CT scans. This dry-run would help the common readers to understand better. This dry run's steps include generating confidence scores, merging these scores using the developed Mitscherlich function-based ensemble procedure, and obtaining the final prediction on test data.

The confidence or probability scores of the training and testing images are generated by individual base classifiers and stored in their individual CSV files.
Table 4.5 in the main paper shows the results of the selected base models.

Table A1: Sample contents of the CSV file for the three base classifiers showing the confidence scores of each class: Xception (Table A1a), InceptionResNetV2 (Table A1b) and MobileNetV2 (Table A1c)

Benign	Malignant	Normal
0.8627	0.0002	0.1371
0.5998	0.2816	0.1185
0.8915	0.0071	0.1013
0.7174	0.0036	0.2789
0.9304	0.0012	0.0684
0.9363	0.0139	0.0498

(a) Xception

Benign	Malignant	Normal
0.8456	3.81E-07	0.1544
0.7019	0.00021	0.2978
0.5832985	2.85E-06	0.4167
0.7554	0.1430	0.1016
0.6337	6.50E-06	0.3663
0.6286	0.2517	0.1197

(b) InceptionResNetV2

Benign	Malignant	Normal
0.9112	1.45E-06	0.0888
0.9123	2.41E-06	0.0876
0.9257	1.32E-06	0.0743
0.7389	0.0793	0.1818
0.9659	3.36E-07	0.0341
0.7188	0.0197	0.2616

(c) MobileNetV2

In Table A1, columns represent classes (Benign, Malignant, Normal), and rows represent images in the dataset. Here the tables show results for six randomly chosen samples.

Then using the Mitscherlich function in Equation 2 (in the main paper), generate the fuzzy ranks for all the samples that belong to various classes.

$$A_1 = \begin{pmatrix} \begin{bmatrix} 0.08036 & 0.99997 & 0.95817 \end{bmatrix} \\ \begin{bmatrix} 0.18151 & 0.99986 & 0.90033 \end{bmatrix} \\ \begin{bmatrix} 0.09625 & 0.99996 & 0.94948 \end{bmatrix} \\ \begin{bmatrix} 0.96089 & 0.99964 & 0.07595 \end{bmatrix} \\ \begin{bmatrix} 0.88624 & 0.99995 & 0.20436 \end{bmatrix} \\ \begin{bmatrix} 0.97705 & 0.99958 & 0.04569 \end{bmatrix} \end{pmatrix} \tag{4.13}$$

$$A_2 = \begin{pmatrix} \begin{bmatrix} 0.20297 & 0.99999 & 0.88705 \end{bmatrix} \\ \begin{bmatrix} 0.37327 & 0.99985 & 0.77073 \end{bmatrix} \\ \begin{bmatrix} 0.50173 & 0.99999 & 0.66513 \end{bmatrix} \\ \begin{bmatrix} 0.87021 & 0.99563 & 0.23746 \end{bmatrix} \\ \begin{bmatrix} 0.94939 & 0.99989 & 0.09654 \end{bmatrix} \\ \begin{bmatrix} 0.88377 & 0.99950 & 0.20914 \end{bmatrix} \end{pmatrix} \tag{4.14}$$

$$A_3 = \begin{pmatrix} \begin{bmatrix} 0.11936 & 0.99999 & 0.93653 \end{bmatrix} \\ \begin{bmatrix} 0.11789 & 0.99999 & 0.93736 \end{bmatrix} \\ \begin{bmatrix} 0.10038 & 0.99999 & 0.94716 \end{bmatrix} \\ \begin{bmatrix} 0.99048 & 0.99923 & 0.02038 \end{bmatrix} \\ \begin{bmatrix} 0.99239 & 0.99996 & 0.01520 \end{bmatrix} \\ \begin{bmatrix} 0.91669 & 0.99823 & 0.15706 \end{bmatrix} \end{pmatrix} \tag{4.15}$$

The matrices A_1, A_2 *and* A_3 in Equation 4.13, Equation 4.14 and Equation 4.15 show sample fuzzy ranks generated using Equation 2 (in the main paper) for Xception, InceptionResNetV2 and MobileNetV2 base classifier, respectively. Each column represents a different class (Benign, Malignant, and Normal), whereas each row represents an image from the dataset.

Now with the help of Equation 4.8 (in the main paper) and Equation 4.9 (in the main paper), generate the fuzzy rank sum and the complement of confidence factor sum as shown in Equation 4.16 by matrix B and Equation 4.17 by matrix C.

$$B = \begin{pmatrix} \begin{bmatrix} 0.40269 & 3. & 2.78176 \end{bmatrix} \\ \begin{bmatrix} 0.67267 & 3. & 2.60843 \end{bmatrix} \\ \begin{bmatrix} 0.69836 & 3. & 2.56177 \end{bmatrix} \\ \begin{bmatrix} 2.82158 & 3. & 0.33379 \end{bmatrix} \\ \begin{bmatrix} 2.82802 & 3. & 0.31609 \end{bmatrix} \\ \begin{bmatrix} 2.77752 & 3. & 0.41189 \end{bmatrix} \end{pmatrix} \tag{4.16}$$

$$C = \begin{pmatrix} \begin{bmatrix} 0.10078 & 1. & 0.89924 \end{bmatrix} \\ \begin{bmatrix} 0.17431 & 1. & 0.82583 \end{bmatrix} \\ \begin{bmatrix} 0.18738 & 1. & 0.81264 \end{bmatrix} \\ \begin{bmatrix} 0.91831 & 1. & 0.08432 \end{bmatrix} \\ \begin{bmatrix} 0.92080 & 1. & 0.07929 \end{bmatrix} \\ \begin{bmatrix} 0.89773 & 1. & 0.10356 \end{bmatrix} \end{pmatrix} \tag{4.17}$$

Both in matrix B in Equation 4.16 and matrix C in Equation 4.17, columns, as usual, represent classes in the order Benign, Malignant, and Normal, whereas rows represent each CT scan image in the dataset.

To obtain the final decision score, perform element-wise multiplication between the fuzzy rank sum matrix B and the complement of confidence factor sum matrix C that will be used to generate the final predictions of our ensemble approach. use the Equation 4.10 to get the result represented by matrix D in Equation 4.18.

$$D = \begin{pmatrix} \begin{bmatrix} 0.04058 & 3. & 2.50147 \end{bmatrix} \\ \begin{bmatrix} 0.11726 & 3. & 2.15411 \end{bmatrix} \\ \begin{bmatrix} 0.13086 & 3. & 2.08179 \end{bmatrix} \\ \begin{bmatrix} 2.59109 & 3. & 0.02815 \end{bmatrix} \\ \begin{bmatrix} 2.60404 & 3. & 0.02506 \end{bmatrix} \\ \begin{bmatrix} 2.49347 & 3. & 0.04266 \end{bmatrix} \end{pmatrix} \tag{4.18}$$

In the end, to get the final prediction of the developed MENet model, use Equation 4.11. Hence, based on it, the final prediction for:

$$image_1 = \min(0.04058,\ 3.,\ 2.50147) = 0.04058 = Benign$$

$$image_2 = \min(0.11726,\ 3.,\ 2.15411) = 0.11726 = Benign$$

$$image_3 = \min(0.13086,\ 3.,\ 2.08179) = 0.13086 = Benign$$

$$image_4 = \min(2.5911,\ 3.,\ 0.0282) = 0.0282 = Normal$$

$$image_5 = \min(2.6040,\ 3.,\ 0.0251) = 0.0251 = Normal$$

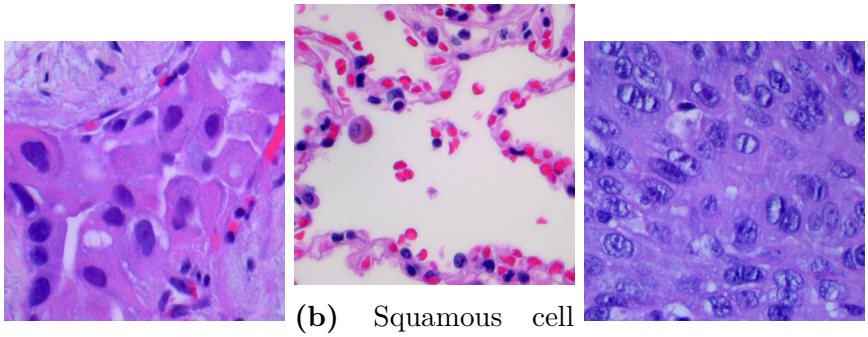
$$image_6 = \min(2.4935,\ 3.,\ 0.0427) = 0.0427 = Normal$$

This is how our developed ensemble model, called MENet, classifies images into Benign, Malignant, and Normal categories.

4.6 Experimental Results for Method 3

4.6.1 Dataset Description

The dataset used in this study, contains nearly 15,000 histopathological images with 3 classes of lung cancer consisting of benign tissue, squamous cell carcinoma and adenocarcinoma. It has been taken from the LC25000 dataset, which is available as a public data repository for lung and colon cancer images. This dataset comprises approximately 15000 sample images of lung cancer and 10000 sample images of colon cancer. All images are 768 x 768 pixels in size and are in jpeg file format. For our work, have only taken the lung adenocarcinoma, lung squamous cell carcinoma, and benign lung tissue images each numbering 5000 i.e., have a total of 15000 images. The details have been tabulated in Table A2. Three sample images, one from each class, have been shown in Figure 4.33.



(a) Benign tissue **(b)** Squamous cell carcinoma **(c)** Adenocarcinoma

Figure 4.33: Samples from the three different classes of lung cancer histopathological images taken from the dataset LC25000

In order to identify the optimal feature extractor, tested several CNN models based on transfer learning (pretrained on ImageNet) on the dataset. The models compared include VGG16, DenseNet121, and Inception V3. Among these, DenseNet121 yielded the highest accuracy. Table A3 presents a comparison of the accuracies of these models.

Table A2: Details of the lung histopathological dataset

Split	ACA	N	SCC	Total
Train	3500	3500	3500	10500
Test	1000	1000	1000	3000
Val	500	500	500	1500
Total	5000	5000	5000	15000

Additionally, evaluated VGG19, ResNet50, Xception, and Inception ResNet V2. While VGG19 demonstrated high training accuracy (98.23%), its test accuracy was lower (95.42%). ResNet50, Xception, and Inception ResNet V2 showed poorer performance, with ResNet50 having the lowest test accuracy (78.55%).

Table A4 illustrates the precision, recall, and F1-score comparisons of the developed model. Table A5 provides a 5-fold cross-validation accuracy comparison among the tested models. DenseNet121 achieved the highest mean accuracy (97.6%), surpassing VGG16 and Inception V3, both of which attained a mean accuracy of 96.2%. Standard deviations were 0.4% for VGG16 and Inception V3, and 0.3% for DenseNet121, indicating consistent performance. The additional models did not outperform DenseNet121. Table A6 presents a performance comparison of the standard DenseNet model, DenseNet with an additional MLP layer, and DenseNet enhanced with Channel Attention, highlighting their respective accuracies and feature counts. have also tested performance of using feature selection with the KNN classifier and our developed methodology. The results of the comparison are shown in Table A7.

Table A3: Performance comparison of different CNN models tested on lung histopathological images with 25 epochs and cross-categorical entropy as the loss function. Accuracy scores are in %.

Sl No.	**Model**	**Train Acc**	**Test Acc**
1	VGG16	97.04	96.01
2	DenseNet121	97.99	97.20
3	Inception V3	96.47	96.10
4	VGG19	98.23	95.42
5	ResNet50	83.74	78.55
6	Xception	90.05	83.36
7	Inception ResNet V2	94.13	92.87

Table A4: Performance comparison of developed model on various classes of lung cancer images in terms of Precision, F1-score, and the Recall

Class	**Precision**	**F1-Score**	**Recall**
Lung adenocarcinoma	99.75	99.59	99.45
Benign lung tissue	100.0	100.0	100.0
Squamous cell carcinoma	99.45	99.66	99.80
Average	**99.73**	**99.75**	**99.75**

Table A5: 5-fold Cross-validation Results for CNN Models

Model	**Fold Accuracy (%)**				
	1	**2**	**3**	**4**	**5**
VGG16	96.5	95.8	96.2	95.9	96.7
DenseNet121	97.2	97.6	97.9	97.4	97.8
Inception V3	96.1	96.3	95.8	96.6	96.1
Mean Accuracy (%)					
VGG16	96.2 ± 0.4				
DenseNet121	97.6 ± 0.3				
Inception V3	96.2 ± 0.4				

Table A6: Performance comparison of DenseNet, DenseNet with one MLP layer, and DenseNet with Channel Attention.

Method	**Accuracy (%)**
DenseNet	97.20
DenseNet with one MLP layer	97.31
DenseNet with Channel Attention	98.95

Table A7: Performance comparison between KNN classifier trained on original DenseNet features and the developed GA-based KNN classifier.

Method	**Accuracy (%)**	**Precision (%)**	**Recall (%)**
KNN on original DenseNet features	97.45	96.80	97.10
developed GA-based FS with KNN classifier	99.75	99.60	99.65

Table A8: Performance comparison of model with only DenseNet 121 with Channel Attention, without FS, and with FS.

Method	Accuracy (%)	No. of features
Without FS	98.95	128
With FS	99.75	22

4.6.2 Parameter Tuning of GA

To thoroughly ascertain the performance of the developed model, the accuracy of the model without FS and with FS is evaluated. Upon comparison, observe that the accuracy of the channel attention enabled DenseNet121 based model is 98.95%, while that with FS, it is 99.75% along with the fact that the number of features have reduced drastically from 128 to 22 as described in Table A8. This clearly shows that FS improves the accuracy of the overall model which can be supported further by comparing the Figure 4.34 where have the confusion matrix obtained by the Channel attention enabled DenseNet121 model **without FS** with the Figure 4.35 showing the confusion matrix for the model with FS.

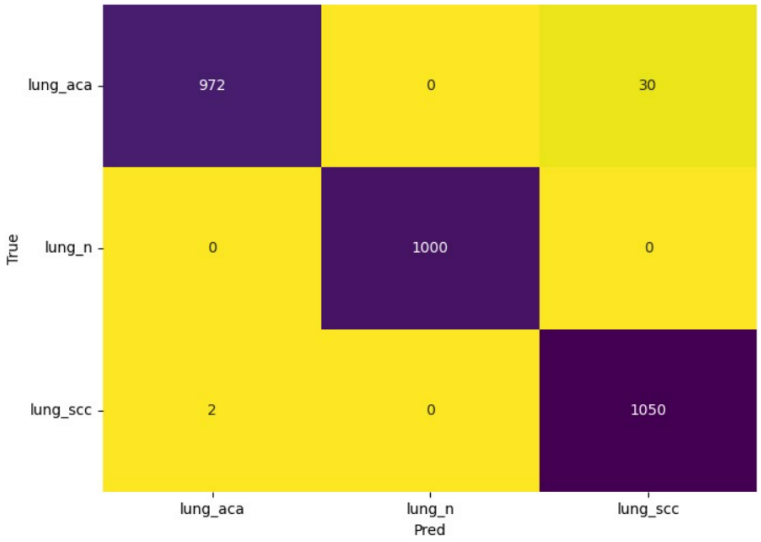


Figure 4.34: Confusion matrix obtained by DenseNet121 with channel attention on the LC25000 dataset i.e., **without FS**

The performance of the classification model is strongly influenced by the FS hyper-parameters. In this section, analyze the impact of crucial FS hyper-parameters, namely population size, crossover probability, and the number of iterations, on the overall accuracy of the model.

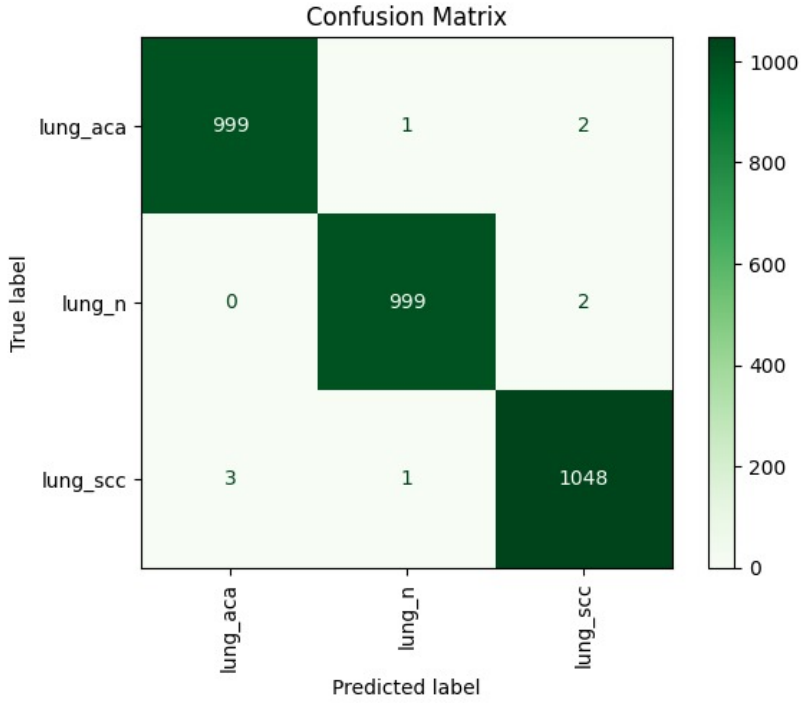


Figure 4.35: Confusion matrix obtained by the developed method on the LC25000 dataset **after applying FS**

4.6.3 Population Size

The population size plays a crucial role in the search for the best solution within the given search space. A larger population size enhances the chances of finding an optimal solution. By increasing the population size, the exploration of the search space is expanded, allowing for a more comprehensive search and potentially leading to better results. In this paper, experiment with different population sizes starting from 10 and increasing to 115 with a fixed interval of 5. The population vs accuracy graph is shown in Figure 4.36. It is observed that the accuracy follows a zig-zag pattern with that maximum accuracy being 99.71% at a population size of 20. Hence take 20 as the default population size for our developed method.

4.6.4 Number of Iterations

Figure 4.37 depicts the variation of the accuracy with the number of iterations. start from 5 iterations and go up to 95 iterations. From the graph observe that the accuracy ranges from 99.68 to 99.75 till 60 iterations and then registers a sharp fall. The maximum accuracy obtained is 99.75% at 50 iterations. Hence take 50 iterations as the default number of iterations for our developed method. Figure 4.38 shows the variation of the number of iterations with the number of selected features. The

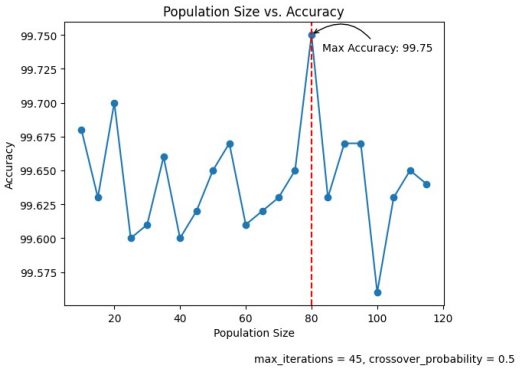


Figure 4.36: Population vs Accuracy curve showing maximum accuracy of 99.75%.

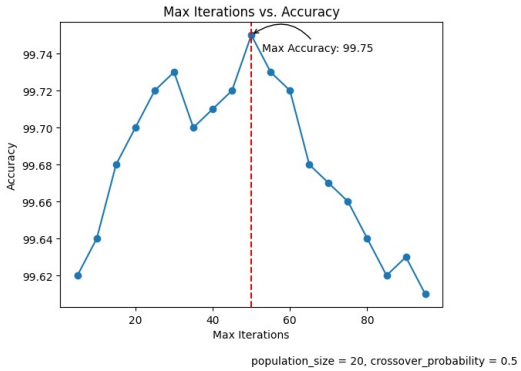


Figure 4.37: Number of iterations vs Accuracy showing maximum accuracy of 99.75%.

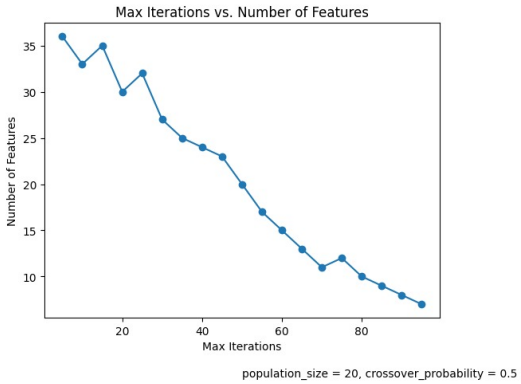


Figure 4.38: Number of iterations vs Number of features in GA.

details of the hyperparameters selected at various stages of developed method has been highlighted in Table A9.

Crossover Probability

Crossover serves as a genetic operator that randomly generates fresh potential solutions by combining existing individuals within a population. The probability of crossover determines the chance of a mating event leading to a crossover. In this case, have varied the crossover probability starting from 0.1 up to 0.95.

4.6.5 Minimum Mutation Probability

Mutation on the newly created offspring generated from crossover is performed. For each offspring, the mutation is applied to randomly selected features. The probability of mutation is determined by mutation-prob, which gradually increases from min mutation probability to max mutation probability over the generations. In this case, have varied the min mutation probability starting from 0.00 up to 0.10. The maximum accuracy achieved is 99.71% at a min mutation probability of 0.01. Hence take 0.01 as the default min mutation probability for our work.

Table A9: Hyperparameter details of the various stages of our developed method.

Stage	Hyperparameter	Value
Feature Extraction	Optimizer	Adam
	Learning rate	0.01
	No. of epochs	25
	Batch size	32
	Loss function	categorical_crossentropy
Feature Selection	Population size	20
	Crossover probability	0.8
	No. of iteration	60
	Minimum mutation probability	0.01
	Maximum mutation probability	0.04
	Alpha	0.99
Classification	K value for KNN	5

Comparison with state-of-the-art methods

Table A10 shows the performances of the past methods that had been applied on the LC25000 dataset. From the table, it can be concluded that our model shows better accuracy than the current state-of-the-art model while maintaining a low resource overhead.

Table A10: Comparison of the developed method with past methods

Model	Year	Accuracy (in %)
Mehmood et al. [10]	2022	89.90
Masud et al. [34]	2021	96.33
Mangal et al. [31]	2020	97.89
Hatuwal et al. [24]	2020	97.20
Nishio et al. [38]	2021	99.43
Nandita et al.	**2024**	**99.75**

t-SNE plots

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a popular dimensionality reduction technique that can be used to visualize high-dimensional data in a lower-dimensional space. At first, the algorithm transforms the high-dimensional Euclidean distances among the data points into conditional probability scores that indicate similarities. This is achieved using SNE (Stochastic Neighbor Embedding) on the data points.

It is observed that the images corresponding to the three different classes of image from lung histopathological dataset are clearly separated into different clusters of samples. The t-SNE plot visualizations for the lung-n, lung-aca, and lung-scc classes, before feature reduction, are presented in Figure 4.39. These visualizations have been generated using the DenseNet121 model. The t-SNE plot of the DenseNet121 model after feature reduction with GA is presented in Figure 4.40.

Statistical Test

To test the hypothesis that the new algorithm produces outcomes similar to those of established algorithms, have utilized the Wilcoxon Rank-Sum test [291]. have selected some well-known FS algorithms and performed 15 separate executions for each algorithm on the employed dataset, documenting the classification accuracy of each run. Our set significance level is a p-value of less than 0.05. If the calculated p-value falls below this threshold, have rejected the hypothesis due to insufficient evidence supporting the claim of comparability. Table A11 lists the p-values obtained for different FS algorithms. This statistically significant result suggests that the developed algorithm consistently performs better than existing algorithms across different splits of the data, highlighting its stability and superior performance. Thus, conclude that compared to the existing FS algorithms, our method demonstrates statistical superiority.

Table A11: p-values for different FS algorithms with our developed method on LC25000 lung cancer images

Algorithm	p-value
Grey Wolf Optimisation (GWO)	0.000197
Mayfly Optimization Algorithm (MA)	0.008389
Binary Bat Algorithm (BBA)	0.000750
Particle Swarm Optimization (PSO)	0.000068
Equilibrium Optimizer (EO)	0.001199
Red Deer Algorithm (RDA)	0.000069
Whale Optimization Algorithm (WOA)	0.001592

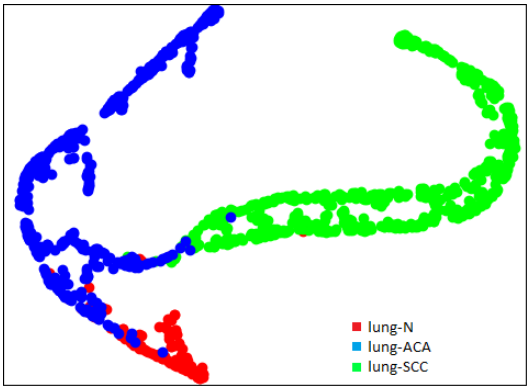


Figure 4.39: t-SNE plot of lung histopathological image samples obtained using the DenseNet121 model.

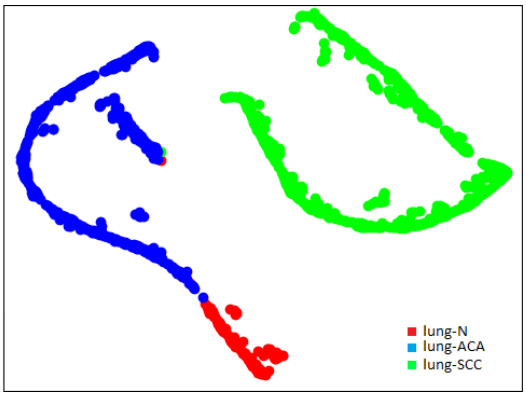


Figure 4.40: t-SNE plot of lung histopathological image samples using a combination of the DenseNet121 and GA.

Results on the LIDC-IDRI Dataset

In order to test the efficacy of the developed method, have also tested our method on another dataset, namely the LIDC-IDRI [292] dataset. The dataset description can be found in Table A12, which shows the two classes - benign and malignant and also shows the number of images in each class. The accuracy achieved by the model before feature selection and after feature selection has been provided in Table A13. It clearly shows the improvement achieved by feature selection after feature extraction by DenseNet121 model, as opposed to only using DenseNet121. Further the efficacy and robustness of the developed method can be verified by comparing its performance against state-of-the-art models based on the LIDC-IDRI dataset. From Table A14, it can be observed that the developed solution performs comparably with the state-of-the-art models. Hence from these evidences, the robustness and efficacy of the developed solution can be ascertained. Additionally, have also tested the current developed method using small cell lung cancer (SCLC) dataset, results of which are shown in Table A15. This table presents the results with SCLC histopathological images as training & validation

dataset.

Table A12: LIDC-IDRI dataset description

Image Type	**Benign**	**Malignant**	**Total**
Train	4342	845	5187
Test	1340	282	1622
Validation	1073	224	1297
Total	6755	1351	8106

Table A13: Performance comparison of developed method on LIDC-IDRI without FS and with FS.

Method	**Accuracy (%)**
Without FS	80.77
With FS	95.38

Table A14: Performance comparison of the developed model with some existing methods for the LIDC-IDRI dataset

Work Ref.	Method	Accuracy(%)
Rey et al. [293]	Fuzzy clustering, SVM and ANN	94.00
Jiang et al. [294]	Convolutional block attention module	91.00
Shaffie et al. [295]	Resolved Ambiguity Local Binary Pattern	92.50
Xie et al. [296]	Semi-supervised Adversarial Classification mode	92.50
Zhang et al. [297]	Basic 3D CNN, multi-output network, 3D DenseNet	90.40
Nandita et al.	**Deep FS algorithm**	**95.38**

Table A15: Performance comparison of developed model on various lung cancer datasets

Dataset	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%**
NSCLC	99.75	97.66	99.12	98.38
SCLC	91.54	90.56	92.10	91.34
Combined (NSCLC + SCLC)	97.89	96.45	97.66	98.56

Chapter 5

Conclusion and Future Directions

5.1 Future Scope

In order to improve the effectiveness of lung nodule segmentation using DL and aligned approaches, the potential future prospects presented in this part may be taken into consideration by the upcoming researchers. Based on our review of a few studies, can say that DL-based approaches have demonstrated good performance in applications linked to lung nodule segmentation, but there are still many opportunities as well as obstacles in the field of medical image segmentation research including lung cancer. Here, list a number of difficulties and potential research directions related to lung nodule segmentation.

- **Dataset Availability:** Public access to larger hospital datasets, with ethical considerations, would enhance model accuracy and robustness, providing more comprehensive data for training.
- **Advanced DL Models:** More sophisticated DL models are needed to better utilize large datasets and handle challenges posed by limited labeled data.
- **Semi-supervised Learning:** Leveraging both labeled and unlabeled data can significantly improve segmentation performance, especially in data-limited situations [298] [299].
- **Explainable AI:** Explainable AI techniques like GANs or VAEs can enhance trust and interpretability in lung segmentation models [300].
- **Ensemble Learning:** Combining diverse base models allows for better segmentation accuracy by capturing complex patterns and perspectives.
- **Lightweight Models:** Lightweight models are crucial for deployment in IoT environments, ensuring efficiency in resource-constrained settings.
- **Metaheuristic Algorithms:** Metaheuristic algorithms, such as PSO and GA, can optimize segmentation methods by adapting to the unique properties of lung image data.

- **Multimodal Data:** Integrating CT scans, X-rays, and histopathology allows for more accurate diagnoses and personalized treatments [301].

- **Medical Collaboration:** Collaborating with clinicians ensures AI models remain clinically relevant and improve patient care.

- **Universal Models in Lung Nodule Segmentation:** Universal models such as Segment Anything Model (SAM), MedSAM, and SAMed2D represent significant advancements in medical image segmentation, including applications in lung nodule detection. In their survey on the SAM for medical image segmentation, the authors provide a comprehensive overview of its applications and challenges in the healthcare domain [302].

  Segment Anything Model: SAM is designed for versatile and robust image segmentation across various domains, including medical imaging. Its ability to handle diverse imaging data makes it suitable for segmenting lung nodules from CT scans with high precision. Khalili et al. [303] developed an automatic lung segmentation method in chest X-ray images by integrating SAM with YOLO prompts. SAM's adaptability to different contexts enhances its utility in detecting and delineating lung nodules, potentially improving diagnostic accuracy and aiding in early lung cancer detection. Mazurowski et al. conducted an experimental study applying the SAM to medical image analysis. The study explores SAM's performance across different medical imaging tasks, providing valuable insights into its potential applications in healthcare [304].

  MedSAM: Tailored specifically for medical imaging, MedSAM leverages advanced deep learning techniques to perform detailed and accurate segmentation of medical structures. Sridhar and Kamaleswaran proposed the Lung Segment Anything Model (LuSAM), a novel prompt-integrated framework designed for automated lung segmentation on ICU chest X-ray images. Their approach leverages the flexibility of SAM to enhance segmentation accuracy in critical care settings. In lung cancer diagnosis, MedSAM excels at identifying and segmenting lung nodules with high sensitivity and specificity. Its model architecture is optimized for medical images, allowing for precise nodule delineation, which is critical for assessing cancerous growths and planning treatment.

  SAMed2D: An extension of SAM focused on 2D medical imaging, SAMed2D is particularly effective in segmenting nodules in chest X-rays and other 2D imaging modalities. By integrating cutting-edge segmentation algorithms, SAMed2D enhances the detection of lung nodules, contributing to improved diagnostic workflows and better patient outcomes. In their work, Bui et al. introduced

SAM3D, extending the SAM to handle volumetric medical images. The study demonstrates how SAM3D improves 3D medical segmentation tasks by leveraging foundation models and transfer learning techniques [305].

5.2 Limitations and Recommendations

Though have achieved state-of-the-art accuracy, there are a few false positive and false negative cases. As this research work is related to cancer, such errors are not acceptable in real-life scenarios. Hence, in the future, will try to address these wrong classification cases. may apply channel and spatial attention mechanisms to extract better feature maps. Dimensionality reduction can also be done to reduce the number of features and retain as much information as possible. As this is an ensemble-based approach, in the future, may plan to use some lightweight CNN models as the base learner. Another future plan is to apply the proposed method to some other medical image datasets to ensure the robustness of the method.

5.3 Conclusion

In conclusion, this thesis demonstrates the strong potential of deep learning in automating lung cancer detection, classification, and segmentation using medical imaging data. By integrating advanced architectures with ensemble learning strategies, the research achieved significant improvements in segmentation precision, classification accuracy, and overall model sensitivity, while effectively minimizing false negatives.

The study also explored adaptive feature selection techniques to enhance model interpretability and reduce redundant information, leading to more efficient learning and reliable outcomes. Collectively, these efforts contribute toward building intelligent, data-driven diagnostic systems that can support radiologists in identifying lung abnormalities more accurately and consistently. The findings reaffirm that well-designed deep learning frameworks can play a crucial role in improving early diagnosis, optimizing clinical workflows, and ultimately enhancing patient care in lung cancer management.

References

[1] WHO. World health organisation. *https://www.who.int/news-room/fact-sheets/detail/cancer*, 2020.

[2] Arash Heidari, Danial Javaheri, Shiva Toumaj, Nima Jafari Navimipour, Mahsa Rezaei, and Mehmet Unal. A new lung cancer detection method based on the chest ct images using federated learning and blockchain systems. *Artificial Intelligence in Medicine*, 141:102572, 2023.

[3] Samuel Greenbank and David Howey. Automated feature extraction and selection for data-driven models of rapid battery capacity fade and end of life. *IEEE Transactions on Industrial Informatics*, 18(5):2965–2973, 2022.

[4] Lareib Fatima Talib, Javaria Amin, Muhammad Sharif, and Mudassar Raza. Transformer-based semantic segmentation and cnn network for detection of histopathological lung cancer. *Biomedical Signal Processing and Control*, 92:106106, 2024.

[5] Anum Masood, Bin Sheng, Ping Li, Xuhong Hou, Xiaoer Wei, Jing Qin, and Dagan Feng. Computer-assisted decision support system in pulmonary cancer detection and stage classification on ct images. *Journal of Biomedical Informatics*, 79:117–128, 2018.

[6] Amir S. Moosavi, Ali Mahboobi, Farshid Arabzadeh, Nasrin Ramezani, Hossein S. Moosavi, and Gholamreza Mehrpoor. Segmentation and classification of lungs ct-scan for detecting covid-19 abnormalities by deep learning technique: U-net model. *Journal of Family Medicine and Primary Care*, 13(2):691–698, Feb 2024. Epub 2024 Mar 6.

[7] Rukhmini Roy, Tapabrata Chakraborti, and Ananda S Chowdhury. A deep learning-shape driven level set synergism for pulmonary nodule segmentation. *Pattern Recognition Letters*, 123:31–38, 2019.

[8] Hyung Min Kim, Taehoon Ko, In Young Choi, and Jun-Pyo Myong. Asbestosis diagnosis algorithm combining the lung segmentation method and deep learning model in computed tomography image. *International Journal of Medical Informatics*, 158:104667, 2022.

[9] Qixin Gao, ShengJun Wang, Dazhe Zhao, and Jiren Liu. Accurate lung segmentation for x-ray ct images. In *Third International Conference on Natural Computation (ICNC 2007)*, volume 2, pages 275–279, 2007.

[10] Xu Y Zhao C, Tang J He Z, Han J Zhang Y, and Shi Y. Lung segmentation and automatic detection of covid-19 using radiomic features from chest ct images. *ScienceDirect*, 2021.

[11] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.

[12] Tao Zhou, YaLi Dong, HuiLing Lu, XiaoMin Zheng, Shi Qiu, and SenBao Hou. Apu-net: An attention mechanism parallel u-net for lung tumor segmentation. *BioMed Research International*, 2022, 2022.

[13] Hexuan Hu, Qingqiu Li, Yunfeng Zhao, and Ye Zhang. Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors. *IEEE Transactions on Industrial Informatics*, 17(4):2880–2889, 2020.

[14] Worku J Sori, Jiang Feng, Arero W Godana, Shaohui Liu, and Demissie J Gelmecha. Dfd-net: lung cancer detection from denoised ct scan image using deep learning. *Frontiers of Computer Science*, 15(2):1–13, 2021.

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[16] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

[17] Anita Khanna, Narendra D Londhe, S Gupta, and Ashish Semwal. A deep residual u-net convolutional neural network for automated lung segmentation in computed tomography images. *Biocybernetics and Biomedical Engineering*, 40(3):1314–1327, 2020.

[18] Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M Taha, and Vijayan K Asari. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.

[19] Qing-qing LIU, Zhi-yong ZHOU, Guo-hua FAN, Xu-sheng QIAN, Ji-su HU, Guang-qiang CHEN, and Ya-kang DAI. Semi-supervised learning segmentation method of liver ct images based on 3d scse-unet. *Journal of ZheJiang University (Engineering Science)*, 55(11):2033–2044, 2021.

[20] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[21] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[22] Siddique Nahian, Sidike Paheding, Elkin Colin, and Devabhaktuni Vijay. U-net and its variants for medical image segmentation: Theory and applications. *Preprint at arxiv*, 1118:v1, 2011.

[23] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. *CoRR*, abs/1807.10165, 2018.

[24] Chang Yao, Jingyu Tang, Menghan Hu, Yue Wu, Wenyi Guo, Qingli Li, and Xiao-Ping Zhang. Claw u-net: A unet-based network with deep feature concatenation for scleral blood vessel segmentation. *Conference: 2022 5th International Conference on Information Communication and Signal Processing (ICICSP)*, 2020.

[25] Xiao Xiao, Shen Lian, Zhiming Luo, and Shaozi Li. Weighted res-unet for high-quality retina vessel segmentation. *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, pages 327–331, 2018.

[26] Ozan Oktay, Jo Schlemper, Loïc Le Folgoc, Matthew C. H. Lee, Mattias P. Heinrich, Kazunari Misawa, Kensaku Mori, Steven G. McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention u-net: Learning where to look for the pancreas. *CoRR*, abs/1804.03999, 2018.

[27] Yao Chang, Menghan Hu, Guangtao Zhai, and Xiao-Ping Zhang. Transclaw u-net: Claw u-net with transformers for medical image segmentation. *CoRR*, abs/2107.05188, 2021.

[28] Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng-Ann Heng. H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes. *IEEE Transactions on Medical Imaging*, 37(12):2663–2674, 2018.

[29] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.

[30] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. *CoRR*, abs/1709.01507, 2017.

[31] Sainbayar Sukhbaatar, Edouard Grave, Piotr Bojanowski, and Armand Joulin. Adaptive attention span in transformers. *CoRR*, abs/1905.07799, 2019.

[32] Shaoqing Xu, Dingfu Zhou, Jin Fang, Junbo Yin, Zhou Bin, and Liangjun Zhang. Fusionpainting: Multimodal fusion with adaptive attention for 3d object detection. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 3047–3054, 2021.

[33] Jiasen Lu, Caiming Xiong, Devi Parikh, and Richard Socher. Knowing when to look: Adaptive attention via a visual sentinel for image captioning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[34] Hanguang Xiao, Li Li, Qiyuan Liu, Xiuhong Zhu, and Qihang Zhang. Transformers in medical image segmentation: A review. *Biomedical Signal Processing and Control*, 84:104791, 2023.

[35] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou. Transunet: Transformers make strong encoders for medical image segmentation. *CoRR*, abs/2102.04306, 2021.

[36] S Niranjan Kumar, P Malin Bruntha, S Isaac Daniel, J Ajay Kirubakar, R Elaine Kiruba, Siril Sam, and S Immanuel Alex Pandian. Lung nodule segmentation using unet. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1:420–424, 2021.

[37] Andrew A. Borkowski, Marilyn M. Bui, L. Brannon Thomas, Catherine P. Wilson, Lauren A. DeLand, and Stephen M. Mastorides. Lung and colon cancer histopathological image dataset (lc25000). *arxiv access*, 2019.

[38] Mizuho Nishio, Mari Nishio, Naoe Jimbo, and Kazuaki Nakane. Homology-based image processing for automatic classification of histopathological images of lung tissue. *Cancers*, 13, 2021.

[39] Qinhua Hu, Luis Fabricio de F Souza, Gabriel Bandeira Holanda, Shara SA Alves, Francisco Hercules dos S Silva, Tao Han, and Pedro P Reboucas Filho. An effective approach for ct lung segmentation using mask region-based convolutional neural networks. *Artificial intelligence in medicine*, 103:101792, 2020.

[40] Tao Peng, Caishan Wang, You Zhang, and Jing Wang. H-segnet: hybrid segmentation network for lung segmentation in chest radiographs using mask region-based convolutional neural network and adaptive closed polyline searching method. *Physics in Medicine & Biology*, 67(7):075006, 2022.

[41] Shuangfeng Dai, Ke Lu, Jiyang Dong, Yifei Zhang, and Yong Chen. A novel approach of lung segmentation on chest ct images using graph cuts. *Neurocomputing*, 168:799–807, 2015.

[42] Ayesha Fatima, Anam Tariq, Mahmood Akhtar, and Hira Zahid. Segmentation of chest radiographs for tuberculosis screening using kernel mapping and graph cuts. *Applications of Intelligent Technologies in Healthcare*, pages 25–34, 2019.

[43] Wei Ju, Deihui Xiang, Bin Zhang, Lirong Wang, Ivica Kopriva, and Xinjian Chen. Random walk and graph cut for co-segmentation of lung tumor on pet-ct images. *IEEE Transactions on Image Processing*, 24(12):5854–5867, 2015.

[44] Mohd Hanafi Ahmad Hijazi, Stefanus Kieu Tao Hwa, Abdullah Bade, Razali Yaakob, and Mohammad Saffree Jeffree. Ensemble deep learning for tuberculosis detection using chest x-ray and canny edge detected images. *IAES International Journal of Artificial Intelligence*, 8(4):429, 2019.

[45] Benjamin Hell, Marc Kassubeck, Pablo Bauszat, Martin Eisemann, and Marcus Magnor. An approach toward fast gradient-based image segmentation. *IEEE Transactions on Image Processing*, 24(9):2633–2645, 2015.

[46] GT Shrivakshan and Chandramouli Chandrasekar. A comparison of various edge detection techniques used in image processing. *International Journal of Computer Science Issues (IJCSI)*, 9(5):269, 2012.

[47] Pune Vit. Comparison of various edge detection technique. *Int. J. Signal Process. Image Process. Pattern Recognit*, 9:143–158, 2016.

[48] Ashish Anand, Sanjaya Shankar Tripathy, and R Sukesh Kumar. An improved edge detection using morphological laplacian of gaussian operator. In *2015 2nd International conference on signal processing and integrated networks (SPIN)*, pages 532–536. IEEE, 2015.

[49] Sheng Chen, Liping Yao, and Bao Chen. A parameterized logarithmic image processing method with laplacian of gaussian filtering for lung nodule enhancement in chest radiographs. *Medical & biological engineering & computing*, 54:1793–1806, 2016.

[50] Qi Mao, Shuguang Zhao, Dongbing Tong, Shengchao Su, Zhiwei Li, and Xiang Cheng. Hessian-mrlog: Hessian information and multi-scale reverse log filter for pulmonary nodule detection. *Computers in Biology and Medicine*, 131:104272, 2021.

[51] Soudeh Saien, Abdol Hamid Pilevar, and Hamid Abrishami Moghaddam. Refinement of lung nodule candidates based on local geometric shape analysis and laplacian of gaussian kernels. *Computers in biology and medicine*, 54:188–198, 2014.

[52] Sergei V Fotin, David F Yankelevitz, Claudia I Henschke, and Anthony P Reeves. A multiscale laplacian of gaussian (log) filtering approach to pulmonary nodule detection from whole-lung ct scans. *arXiv preprint arXiv:1907.08328*, 2019.

[53] K Bhargavi and S Jyothi. A survey on threshold based segmentation technique in image processing. *International Journal of Innovative Research and Development*, 3(12):234–239, 2014.

[54] Azrin Khan, Rachael Garner, Marianna La Rocca, Sana Salehi, and Dominique Duncan. A novel threshold-based segmentation method for quantification of covid-19 lung abnormalities. *Signal, Image and Video Processing*, pages 1–8, 2022.

[55] Jaryd R Christie, Omar Daher, Hannah van Dongen, Rory Gilliland, Mohamed Abdelrazek, and Sarah A Mattonen. A semi-automatic threshold-based segmentation algorithm for lung cancer delineation. In *Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 12036, pages 495–500. SPIE, 2022.

[56] Hailing Zhou, Dmitry B Goldgof, Samuel Hawkins, Lei Wei, Ying Liu, Doug Creighton, Robert J Gillies, Lawrence O Hall, and Saeid Nahavandi. A robust

approach for automated lung segmentation in thoracic ct. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 2267–2272. IEEE, 2015.

[57] Jing Gong, Ji-yu Liu, Li-jia Wang, Bin Zheng, and Sheng-dong Nie. Computer-aided detection of pulmonary nodules using dynamic self-adaptive template matching and a flda classifier. *Physica Medica*, 32(12):1502–1509, 2016.

[58] Narsimha Raj Kasu and Chandran Saravanan. Segmentation on chest radiographs using otsu's and k-means clustering methods. In *2018 International conference on inventive research in computing applications (ICIRCA)*, pages 210–213. IEEE, 2018.

[59] Ji-kui Liu, Hong-yang Jiang, Meng-di Gao, Chen-guang He, Yu Wang, Pu Wang, He Ma, and Ye Li. An assisted diagnosis system for detection of early pulmonary nodule in computed tomography images. *Journal of medical systems*, 41:1–9, 2017.

[60] Zhibin Wang, Kaiyi Wang, Feng Yang, Shouhui Pan, and Yanyun Han. Image segmentation of overlapping leaves based on chan–vese model and sobel operator. *Information processing in agriculture*, 5(1):1–10, 2018.

[61] Nur Najihah Sofia Mohd Marzuki, Iza Sazanita Isa, Noor Khairiah A Karim, Ibrahim Lutfi Shuaib, Zainal Hisham Che Soh, and Siti Noraini Sulaiman. Demarcation of lung lobes in ct scan images for lung cancer detection using watershed segmentation. In *Proceedings of the 2020 12th International Conference on Computer and Automation Engineering*, pages 70–74, 2020.

[62] S Avinash, K Manjunath, and S Senthil Kumar. An improved image processing analysis for the detection of lung cancer using gabor filters and watershed segmentation technique. In *2016 International Conference on Inventive Computation Technologies (ICICT)*, volume 3, pages 1–6. IEEE, 2016.

[63] Tamanna Tajrin, Mamun Ahmed, and Sabina Zaman. Detection of lung cancer stages on computed tomography image using laplacian filter and marker controlled watershed segmentation technique. *Periodica Polytechnica Electrical Engineering and Computer Science*, 66(2):105–115, 2022.

[64] CMAK Zeelan Basha, B Lakshmi Pravallika, D Vineela, and S Lakshmi Prathyusha. An effective and robust cancer detection in the lungs with bpnn

and watershed segmentation. In *2020 International Conference for Emerging Technology (INCET)*, pages 1–6. IEEE, 2020.

[65] Yanyan Wu and Qian Li. The algorithm of watershed color image segmentation based on morphological gradient. *Sensors*, 22(21):8202, 2022.

[66] Aldísio G Medeiros, Matheus T Guimarães, Solon A Peixoto, Lucas de O Santos, Antônio C da Silva Barros, Elizângela de S Rebouças, Victor Hugo C de Albuquerque, and Pedro P Rebouças Filho. A new fast morphological geodesic active contour method for lung ct image segmentation. *Measurement*, 148:106687, 2019.

[67] K Sathish, YV Narayana, Mahammad Shareef Mekala, Patan Rizwan, and Suresh Kallam. Efficient tumor volume measurement and segmentation approach for ct image based on twin support vector machines. *Neural Computing and Applications*, 34(9):7199–7207, 2022.

[68] Shiwen Shen, Alex AT Bui, Jason Cong, and William Hsu. An automated lung segmentation approach using bidirectional chain codes to improve nodule detection accuracy. *Computers in biology and medicine*, 57:139–149, 2015.

[69] Gregory Z Ferl, Kai H Barck, Jasmine Patil, Skander Jemaa, Evelyn J Malamut, Anthony Lima, Jason E Long, Jason H Cheng, Melissa R Junttila, and Richard AD Carano. Automated segmentation of lungs and lung tumors in mouse micro-ct scans. *Iscience*, 25(12):105712, 2022.

[70] Heba M Afify, Ashraf Darwish, Kamel K Mohammed, and Aboul Ella Hassanien. An automated cad system of ct chest images for covid-19 based on genetic algorithm and k-nearest neighbor classifier. *Ingénierie des Systèmes d Inf.*, 25(5):589–594, 2020.

[71] Tao Peng, Thomas Canhao Xu, Yihuai Wang, and Fanzhang Li. Deep belief network and closed polygonal line for lung segmentation in chest radiographs. *The Computer Journal*, 65(5):1107–1128, 2022.

[72] Malathi Murugesan, Kalaiselvi Kaliannan, Shankarlal Balraj, Kokila Singaram, Thenmalar Kaliannan, and Johny Renoald Albert. A hybrid deep learning model for effective segmentation and classification of lung nodules from ct images. *Journal of Intelligent & Fuzzy Systems*, 42(3):2667–2679, 2022.

[73] Srishti Sharma, Prasenjeet Fulzele, and Indu Sreedevi. Hybrid model for lung nodule segmentation based on support vector machine and k-nearest neighbor.

In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, pages 170–175. IEEE, 2020.

[74] Sotiris B Kotsiantis. Decision trees: a recent overview. *Artificial Intelligence Review*, 39:261–283, 2013.

[75] Wei Guo, Uday K. Rage, and Seishi Ninomiya. Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Computers and Electronics in Agriculture*, 96:58–66, 2013.

[76] Seung Hoon Yoo, Hui Geng, Tin Lok Chiu, Siu Ki Yu, Dae Chul Cho, Jin Heo, Min Sung Choi, Il Hyun Choi, Cong Cung Van, Nguen Viet Nhung, et al. Deep learning-based decision-tree classifier for covid-19 diagnosis from chest x-ray imaging. *Frontiers in medicine*, 7:427, 2020.

[77] Beatrice Berthon, Christopher Marshall, Mererid Evans, and Emiliano Spezi. Atlaas: an automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography. *Physics in Medicine & Biology*, 61(13):4855, 2016.

[78] Daniel Markel, Curtis Caldwell, Hamideh Alasti, Hany Soliman, Yee Ung, Justin Lee, and Alexander Sun. Automatic segmentation of lung carcinoma using 3d texture features in 18-fdg pet/ct. *International journal of molecular imaging*, 2013, 2013.

[79] Caixia Liu, Ruibin Zhao, and Mingyong Pang. Lung segmentation based on random forest and multi-scale edge detection. *IET Image Processing*, 13(10):1745–1754, 2019.

[80] Caixia Liu, Ruibin Zhao, and Mingyong Pang. A fully automatic segmentation algorithm for ct lung images based on random forest. *Medical Physics*, 47(2):518–529, 2020.

[81] May Phu Paing, Kazuhiko Hamamoto, Supan Tungjitkusolmun, Sarinporn Visitsattapongse, and Chuchart Pintavirooj. Automatic detection of pulmonary nodules using three-dimensional chain coding and optimized random forest. *Applied Sciences*, 10(7):2346, 2020.

[82] Qiang Li, Lei Chen, Xiangju Li, Xiaofeng Lv, Shuyue Xia, and Yan Kang. Prf-rw: a progressive random forest-based random walk approach for interactive semi-automated pulmonary lobes segmentation. *International Journal of Machine Learning and Cybernetics*, 11(10):2221–2235, 2020.

[83] Guotai Wang, Maria A Zuluaga, Rosalind Pratt, Michael Aertsen, Tom Doel, Maria Klusmann, Anna L David, Jan Deprest, Tom Vercauteren, and Sébastien Ourselin. Dynamically balanced online random forests for interactive scribble-based segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 352–360. Springer, 2016.

[84] K. Lan, J. Zhou, X. Jiang, J. Wang, S. Huang, J. Yang, Q. Song, R. Tang, X. Gong, K. Liu, Y. Wu, and T. Li. Group theoretic particle swarm optimization for multi-level threshold lung cancer image segmentation. *Quantitative Imaging in Medicine and Surgery*, 13(3):1312–1322, 3 2023. Epub 2022 Oct 8.

[85] Shabana R. Ziyad, V. Radha, and Thavavel Vayyapuri. A novel lung extraction approach for ldct images using discrete wavelet transform with adaptive thresholding and fuzzy c-means clustering enhanced by genetic algorithm. *Research on Biomedical Engineering*, 38(2):581–598, 2022.

[86] Supiksha Jain, Sanjeev Indora, and Dinesh Kumar Atal. Lung nodule segmentation using salp shuffled shepherd optimization algorithm-based generative adversarial network. *Computers in Biology and Medicine*, 137:104811, 2021.

[87] Yixin Wang, Jinshun Ding, Weiqing Fang, and Jian Cao. Segmentation-assisted diagnosis of pulmonary nodule recognition based on adaptive particle swarm image algorithm. In Huansheng Ning, editor, *Cyberspace Data and Intelligence, and Cyber-Living, Syndrome, and Health*, pages 504–512, Singapore, 2019. Springer Singapore.

[88] S. Shaik Parveen and C. Kavitha. Segmentation of ct lung nodules using fcm with firefly search algorithm. In *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pages 1–6, 2015.

[89] Juan–juan Zhao, Guo–hua Ji, Yong Xia, and Xiao–long Zhang. Cavitary nodule segmentation in computed tomography images based on self–generating neural networks and particle swarm optimisation. *International Journal of Bio-Inspired Computation*, 7(1):62–67, 2015.

[90] P. Badura and E. Pietka. Soft computing approach to 3d lung nodule segmentation in ct. *Computers in Biology and Medicine*, 53:230–243, 2014.

[91] Tianyi Zhao, Dashan Gao, Jiao Wang, and Zhaozheng Yin. Lung segmentation in ct images using a fully convolutional neural network with multi-instance and conditional adversary loss. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 505–509. IEEE, 2018.

[92] Ying Chen, Yerong Wang, Fei Hu, and Ding Wang. A lung dense deep convolution neural network for robust lung parenchyma segmentation. *IEEE Access*, 8:93527–93547, 2020.

[93] Haichao Cao, Hong Liu, Enmin Song, Chih-Cheng Hung, Guangzhi Ma, Xiangyang Xu, Renchao Jin, and Jianguo Lu. Dual-branch residual network for lung nodule segmentation. *Applied Soft Computing*, 86:105934, 2020.

[94] Sangheum Hwang and Sunggyun Park. Accurate lung segmentation via network-wise training of convolutional networks. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pages 92–99. Springer, 2017.

[95] Xia Huang, Wenqing Sun, Tzu-Liang Bill Tseng, Chunqiang Li, and Wei Qian. Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic ct scans using deep convolutional neural networks. *Computerized Medical Imaging and Graphics*, 74:25–36, 2019.

[96] Zhaohui Bu, Xuejun Zhang, Jianxiang Lu, Huan Lao, Chan Liang, Xianfu Xu, Yini Wei, and Hongjie Zeng. Lung nodule detection based on yolov3 deep learning with limited datasets. *Mol. Cell. Biomech*, 19(1):17–28, 2022.

[97] Hao Tang, Chupeng Zhang, and Xiaohui Xie. Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22*, pages 266–274. Springer, 2019.

[98] Shuo Wang, Mu Zhou, Zaiyi Liu, Zhenyu Liu, Dongsheng Gu, Yali Zang, Di Dong, Olivier Gevaert, and Jie Tian. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Medical image analysis*, 40:172–183, 2017.

[99] Shuo Wang, Mu Zhou, Olivier Gevaert, Zhenchao Tang, Di Dong, Zhenyu Liu, and Tian Jie. A multi-view deep convolutional neural networks for lung nodule segmentation. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1752–1755. IEEE, 2017.

[100] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):640–651, 2017.

[101] Zhaojin Fu, Jinjiang Li, and Zhen Hua. Msa-net: Multiscale spatial attention network for medical image segmentation. *Alexandria Engineering Journal*, 70:453–473, 2023.

[102] Mehmet Akif Cifci et al. Segchanet: A novel model for lung cancer segmentation in ct scans. *Applied Bionics and Biomechanics*, 2022, 2022.

[103] Haojie Song, Yuefei Wang, Shijie Zeng, Xiaoyan Guo, and Zheheng Li. Oau-net: Outlined attention u-net for biomedical image segmentation. *Biomedical Signal Processing and Control*, 79:104038, 2023.

[104] Ran Gu, Guotai Wang, Tao Song, Rui Huang, Michael Aertsen, Jan Deprest, Sébastien Ourselin, Tom Vercauteren, and Shaoting Zhang. Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. *IEEE transactions on medical imaging*, 40(2):699–711, 2020.

[105] Xiaocong Chen, Lina Yao, and Yu Zhang. Residual attention u-net for automated multi-class segmentation of covid-19 chest ct images. *arXiv preprint arXiv:2004.05645*, 2020.

[106] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. *IEEE Transactions on Medical Imaging*, 39(6):1856–1867, 2020.

[107] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 801–818. Springer, 2018.

[108] Nandita Gautam, Abhishek Basu, Dmitry Kaplun, and Ram Sarkar. An ensemble of unet frameworks for lung nodule segmentation. In Anatoly Alikhanov,

Pavel Lyakhov, and Irina Samoylenko, editors, *Current Problems in Applied Mathematics and Computer Science and Systems*, pages 450–461, Cham, 2023. Springer Nature Switzerland.

[109] Yang Xu, Shike Hou, Xiangyu Wang, Duo Li, and Lu Lu. A medical image segmentation method based on improved unet 3+ network. *Diagnostics*, 13(3):576, 2023.

[110] Zezhi Wu, Xiaoshu Li, and Jianhui Zuo. Rad-unet: Research on an improved lung nodule semantic segmentation algorithm based on deep learning. *Frontiers in Oncology*, 13:1084096, 2023.

[111] Tong Wang, Fubin Wu, Haoran Lu, and Shengzhou Xu. Ca-unet: Convolution and attention fusion for lung nodule segmentation. *International Journal of Imaging Systems and Technology*, n/a(n/a), 2023.

[112] Baihua Zhang, Shouliang Qi, Yanan Wu, Xiaohuan Pan, Yudong Yao, Wei Qian, and Yubao Guan. Multi-scale segmentation squeeze-and-excitation unet with conditional random field for segmenting lung tumor from ct images. *Computer Methods and Programs in Biomedicine*, 222:106946, 2022.

[113] Guobin Zhang, Zhiyong Yang, and Shan Jiang. Automatic lung tumor segmentation from ct images using improved 3d densely connected unet. *Medical & Biological Engineering & Computing*, 60(11):3311–3323, 2022.

[114] Lei Yang, Yuge Gu, Benyan Huo, Yanhong Liu, and Guibin Bian. A shape-guided deep residual network for automated ct lung segmentation. *Knowledge-Based Systems*, 250:108981, 2022.

[115] Wessam M Salama and Moustafa H Aly. Lung ct image segmentation: A generalized framework based on u-net architecture and preprocessing models. In *2021 31st International Conference on Computer Theory and Applications (ICCTA)*, pages 141–146. IEEE, 2021.

[116] S Niranjan Kumar, P. Malin Bruntha, S Isaac Daniel, J. Ajay Kirubakar, R Elaine Kiruba, Siril Sam, and S. Immanuel Alex Pandian. Lung nodule segmentation using unet. In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 420–424, 2021.

[117] Wei Chen, Fengchang Yang, Xianru Zhang, Xin Xu, and Xu Qiao. Mau-net: Multiple attention 3d u-net for lung cancer segmentation on ct images. *Procedia Computer Science*, 192:543–552, 2021. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 25th International Conference KES2021.

[118] Kuan bing Chen, Ying Xuan, Ai jun Lin, and Shao hua Guo. Lung computed tomography image segmentation based on u-net network fused with dilated convolution. *Computer Methods and Programs in Biomedicine*, 207:106170, 2021.

[119] Zhitao Xiao, Bowen Liu, Lei Geng, Fang Zhang, and Yanbei Liu. Segmentation of lung nodules using improved 3d-unet neural network. *Symmetry*, 12(11), 2020.

[120] Yeganeh Jalali, Mansoor Fateh, Mohsen Rezvani, Vahid Abolghasemi, and Mohammad Hossein Anisi. Resbcdu-net: A deep learning framework for lung ct image segmentation. *Sensors*, 21(1), 2021.

[121] Joana Rocha, António Cunha, and Ana Maria Mendonça. Conventional filtering versus u-net based models for pulmonary nodule segmentation in ct images. *Journal of Medical Systems*, 44(4):81, 2020.

[122] Muhammad Usman, Byoung-Dai Lee, Shi-Sub Byon, Sung-Hyun Kim, Byung-il Lee, and Yeong-Gil Shin. Volumetric lung nodule segmentation using adaptive roi with multi-view residual learning. *Scientific Reports*, 10(1):12839, 2020.

[123] Guofeng Tong, Yong Li, Huairong Chen, Qingchun Zhang, and Huiying Jiang. Improved u-net network for pulmonary nodules segmentation. *Optik*, 174:460–469, 2018.

[124] Zimeng Tan, Jianjiang Feng, and Jie Zhou. Sgnet: Structure-aware graph-based network for airway semantic segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 153–163. Springer, 2021.

[125] Nicolás Gaggion, Lucas Mansilla, Diego H Milone, and Enzo Ferrante. Hybrid graph convolutional neural networks for landmark-based anatomical segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 600–610. Springer, 2021.

[126] Zhiwei Zhai, Marius Staring, Xuhui Zhou, Qiuxia Xie, Xiaojuan Xiao, M Els Bakker, Lucia J Kroft, Boudewijn PF Lelieveldt, Gudula JAM Boon,

Frederikus A Klok, et al. Linking convolutional neural networks with graph convolutional networks: application in pulmonary artery-vein separation. In *International Workshop on Graph Learning in Medical Imaging*, pages 36–43. Springer, 2019.

[127] Jie Lian, Yonghao Long, Fan Huang, Kei Shing Ng, Faith MY Lee, David CL Lam, Benjamin XL Fang, Qi Dou, and Varut Vardhanabhuti. Imaging-based deep graph neural networks for survival analysis in early stage lung cancer using ct: A multicenter study. *Frontiers in Oncology*, 12, 2022.

[128] Haozhe Jia, Haoteng Tang, Guixiang Ma, Weidong Cai, Heng Huang, Liang Zhan, and Yong Xia. Psgr: Pixel-wise sparse graph reasoning for covid-19 pneumonia segmentation in ct images. *arXiv preprint arXiv:2108.03809*, 2021.

[129] Marc'Aurelio Ranzato, Christopher Poultney, Sumit Chopra, and Yann Cun. Efficient learning of sparse representations with an energy-based model. *Advances in neural information processing systems*, 19, 2006.

[130] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103, 2008.

[131] Rifai Salah, P Vincent, X Muller, et al. Contractive auto-encoders: Explicit invariance during feature extraction. In *Proc. of the 28th International Conference on Machine Learning*, pages 833–840, 2011.

[132] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *International conference on artificial neural networks*, pages 52–59. Springer, 2011.

[133] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[134] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhudinov. Unsupervised learning of video representations using lstms. In *International conference on machine learning*, pages 843–852. PMLR, 2015.

[135] Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey. Adversarial autoencoders. *arXiv preprint arXiv:1511.05644*, 2015.

[136] Ihsan Ullah, Farman Ali, Babar Shah, Shaker El-Sappagh, Tamer Abuhmed, and Sang Hyun Park. A deep learning based dual encoder–decoder framework for anatomical structure segmentation in chest x-ray images. *Scientific Reports*, 13(1):1–14, 2023.

[137] Shichao Luo, Jina Zhang, Ning Xiao, Yan Qiang, Keqin Li, Juanjuan Zhao, Liang Meng, and Ping Song. Das-net: A lung nodule segmentation method based on adaptive dual-branch attention and shadow mapping. *Applied Intelligence*, pages 1–15, 2022.

[138] Abbas Khan, Hyongsuk Kim, and Leon Chua. Pmed-net: Pyramid based multi-scale encoder-decoder network for medical image segmentation. *IEEE Access*, 9:55988–55998, 2021.

[139] Samar Ibrahim, Kareem Elgohary, Mahmoud Higazy, Thanaa Mohannad, Sahar Selim, and Mustafa Elattar. Lung segmentation using resunet++ powered by variational auto encoder-based enhancement in chest x-ray images. In *Annual Conference on Medical Image Understanding and Analysis*, pages 339–356. Springer, 2022.

[140] Amitava Halder and Debangshu Dey. Atrous convolution aided integrated framework for lung nodule segmentation and classification. *Biomedical Signal Processing and Control*, 82:104527, 2023.

[141] Zhengyang Wang and Shuiwang Ji. Smoothed dilated convolutions for improved dense prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2486–2495, 2018.

[142] Shahruk Hossain, Suhail Najeeb, Asif Shahriyar, Zaowad R Abdullah, and M Ariful Haque. A pipeline for lung tumor detection and segmentation from ct scans using dilated convolutional neural networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1348–1352. IEEE, 2019.

[143] Haiying Xia, Weifan Sun, Shuxiang Song, and Xiangwei Mou. Md-net: multi-scale dilated convolution network for ct images segmentation. *Neural Processing Letters*, 51:2915–2927, 2020.

[144] Lei Geng, Siqi Zhang, Jun Tong, and Zhitao Xiao. Lung segmentation method with dilated convolution based on vgg-16 network. *Computer Assisted Surgery*, 24(sup2):27–33, 2019.

[145] Hanif Rasyidi and Salman Khan. Historical document text binarization using atrous convolution and multi-scale feature decoder. In *2019 Digital Image Computing: Techniques and Applications (DICTA)*, pages 1–8. IEEE, 2019.

[146] Mourad Gridach. Pydinet: Pyramid dilated network for medical image segmentation. *Neural networks*, 140:274–281, 2021.

[147] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[148] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.

[149] Anushikha Singh, Brejesh Lall, Bijaya K Panigrahi, Anjali Agrawal, Anurag Agrawal, Balamugesh Thangakunam, and DJ Christopher. Deep lf-net: Semantic lung segmentation from indian chest radiographs including severely unhealthy images. *Biomedical Signal Processing and Control*, 68:102666, 2021.

[150] Ching-Sheng Chang, Jin-Fa Lin, Ming-Ching Lee, and Christoph Palm. Semantic lung segmentation using convolutional neural networks. In *Bildverarbeitung für die Medizin 2020: Algorithmen–Systeme–Anwendungen. Proceedings des Workshops vom 15. bis 17. März 2020 in Berlin*, pages 75–80. Springer, 2020.

[151] N Venugopal. Automatic semantic segmentation with deeplab dilated learning network for change detection in remote sensing images. *Neural Processing Letters*, 51:2355–2377, 2020.

[152] Xiangyu Zhao, Peng Zhang, Fan Song, Guangda Fan, Yangyang Sun, Yujia Wang, Zheyuan Tian, Luqi Zhang, and Guanglei Zhang. D2a u-net: Automatic segmentation of covid-19 ct slices based on dual attention and hybrid dilated convolution. *Computers in biology and medicine*, 135:104526, 2021.

[153] Mian Muhammad Naeem Abid, Tehseen Zia, Mubeen Ghafoor, and David Windridge. Multi-view convolutional recurrent neural networks for lung cancer nodule identification. *Neurocomputing*, 453:299–311, 2021.

[154] S Akila Agnes, J Anitha, and A Arun Solomon. Two-stage lung nodule detection framework using enhanced unet and convolutional lstm networks in ct images. *Computers in Biology and Medicine*, 149:106059, 2022.

[155] Ping Xuan, Bin Jiang, Hui Cui, Qiangguo Jin, Peng Cheng, Toshiya Nakaguchi, Tiangang Zhang, Changyang Li, Zhiyu Ning, Menghan Guo, et al. Convolutional bi-directional learning and spatial enhanced attentions for lung tumor segmentation. *Computer Methods and Programs in Biomedicine*, 226:107147, 2022.

[156] Yeganeh Jalali, Mansoor Fateh, Mohsen Rezvani, Vahid Abolghasemi, and Mohammad Hossein Anisi. Resbcdu-net: a deep learning framework for lung ct image segmentation. *Sensors*, 21(1):268, 2021.

[157] Maryam Asadi-Aghbolaghi, Reza Azad, Mahmood Fathy, and Sergio Escalera. Multi-level context gating of embedded collective knowledge for medical image segmentation. *arXiv preprint arXiv:2003.05056*, 2020.

[158] Vinit Kumar Gunjan, Ninni Singh, Fahimudin Shaik, and Sudipta Roy. Detection of lung cancer in ct scans using grey wolf optimization algorithm and recurrent neural network. *Health and Technology*, 12(6):1197–1210, 2022.

[159] M Kanipriya, C Hemalatha, N Sridevi, SR SriVidhya, and SL Jany Shabu. An improved capuchin search algorithm optimized hybrid cnn-lstm architecture for malignant lung nodule detection. *Biomedical Signal Processing and Control*, 78:103973, 2022.

[160] Shuaijing Xu, Junqi Guo, Guangzhi Zhang, and Rongfang Bie. Automated detection of multiple lesions on chest x-ray images: classification using a neural network technique with association-specific contexts. *Applied Sciences*, 10(5):1742, 2020.

[161] Riqiang Gao, Yucheng Tang, Kaiwen Xu, Yuankai Huo, Shunxing Bao, Sanja L Antic, Emily S Epstein, Steve Deppen, Alexis B Paulson, Kim L Sandler, et al. Time-distanced gates in long short-term memory networks. *Medical image analysis*, 65:101785, 2020.

[162] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

[163] Jiaxin Cai and Hongfeng Zhu. Lung image segmentation by generative adversarial networks. In *2019 International Conference on Image and Video Processing, and Artificial Intelligence*, volume 11321, pages 175–180. SPIE, 2019.

[164] Jiaxing Tan, Longlong Jing, Yumei Huo, Lihong Li, Oguz Akin, and Yingli Tian. Lgan: Lung segmentation in ct scans using generative adversarial network. *Computerized Medical Imaging and Graphics*, 87:101817, 2021.

[165] Faizan Munawar, Shoaib Azmat, Talha Iqbal, Christer Grönlund, and Hazrat Ali. Segmentation of lungs in chest x-ray image using generative adversarial networks. *Ieee Access*, 8:153535–153545, 2020.

[166] Swati P Pawar and Sanjay N Talbar. Lungseg-net: Lung field segmentation using generative adversarial network. *Biomedical Signal Processing and Control*, 64:102296, 2021.

[167] Shweta Tyagi and Sanjay N Talbar. Cse-gan: A 3d conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation. *Computers in Biology and Medicine*, 147:105781, 2022.

[168] Mizuho Nishio, Koji Fujimoto, Hidetoshi Matsuo, Chisako Muramatsu, Ryo Sakamoto, and Hiroshi Fujita. Lung cancer segmentation with transfer learning: usefulness of a pretrained model constructed from an artificial dataset generated using a generative adversarial network. *Frontiers in Artificial Intelligence*, 4:694815, 2021.

[169] Supiksha Jain, Sanjeev Indora, and Dinesh Kumar Atal. Lung nodule segmentation using salp shuffled shepherd optimization algorithm-based generative adversarial network. *Computers in Biology and Medicine*, 137:104811, 2021.

[170] Lei Cui, Jun Feng, and Lin Yang. Semi-supervised deep linear discriminant analysis for histopathology image classification. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 2333–2337, 2018.

[171] Zeyu Gao, Bangyang Hong, Yang Li, Xianli Zhang, Jialun Wu, Chunbao Wang, Xiangrong Zhang, Tieliang Gong, Yefeng Zheng, Deyu Meng, and Chen Li. A semi-supervised multi-task learning framework for cancer classification with weak annotation in whole-slide images. *Medical Image Analysis*, 83:102652, 2023.

[172] Yu Fu, Peng Xue, Taohui Xiao, Zhili Zhang, Youren Zhang, and Enqing Dong. Semi-supervised adversarial learning for improving the diagnosis of pulmonary nodules. *IEEE Journal of Biomedical and Health Informatics*, 27(1):109–120, 2023.

[173] Vemund Fredriksen, Svein Ole M Sevle, André Pedersen, Thomas Langø, Gabriel Kiss, and Frank Lindseth. Teacher-student approach for lung tumor segmentation from mixed-supervised datasets. *Plos one*, 17(4):e0266147, 2022.

[174] Feng Shi, Bojiang Chen, Qiqi Cao, Ying Wei, Qing Zhou, Rui Zhang, Yaojie Zhou, Wenjie Yang, Xiang Wang, Rongrong Fan, Fan Yang, Yanbo Chen, Weimin Li, Yaozong Gao, and Dinggang Shen. Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest ct images. *IEEE Transactions on Medical Imaging*, 41(4):771–781, 2022.

[175] Naji Khosravan and Ulas Bagci. Semi-supervised multi-task learning for lung cancer diagnosis. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 710–713, 2018.

[176] Bruno Rodrigues Froz, Antonio Oseas de Carvalho Filho, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Rodolfo Acatauassú Nunes, and Marcelo Gattass. Lung nodule classification using artificial crawlers, directional texture and support vector machine. *Expert Systems with Applications*, 69:176–188, 2017.

[177] Antonio Oseas de Carvalho Filho, Wener Borges de Sampaio, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Rodolfo Acatauassú Nunes, and Marcelo Gattass. Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index. *Artificial Intelligence in Medicine*, 60(3):165–177, 2014.

[178] Dawid Poap, Marcin Wozniak, Robertas Damaševičius, and Wei Wei. Chest radiographs segmentation by the use of nature-inspired algorithm for lung disease detection. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 2298–2303, 2018.

[179] Stelmo Magalhães Barros Netto, João Otávio Bandeira Diniz, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Rofolfo Acatauassú Nunes, and Marcelo Gattass. Modified quality threshold clustering for temporal analysis and classification of lung lesions. *IEEE Transactions on Image Processing*, 28(4):1813–1823, 2019.

[180] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.

[181] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[182] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[183] Aya Hage Chehade, Nassib Abdallah, Jean-Marie Marion, Mohamad Oueidat, and Pierre Chauvet. Lung and colon cancer classification using medical imaging: a feature engineering approach. *Physical and engineering sciences in medicine*, page 729–746, 2022.

[184] Deepak Rawat, Meenakshi, Lokesh Pawar, Gaurav Bathla, and Ravi Kant. Optimized deep learning model for lung cancer prediction using ann algorithm. In *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pages 889–894, 2022.

[185] Ali Safiyari and Reza Javidan. Predicting lung cancer survivability using ensemble learning methods. In *2017 Intelligent Systems Conference (IntelliSys)*, pages 684–688, 2017.

[186] Zhiwei Cai, Dazhen Xu, Qiang Zhang, Jiyang Zhang, Sai-Ming Ngai, and Jie Shao. Classification of lung cancer using ensemble-based feature selection and machine learning methods. *Molecular BioSystems*, 11(3):791–800, Mar 2015.

[187] Rehan Raza, Fatima Zulfiqar, Muhammad Owais Khan, Muhammad Arif, Atif Alvi, Muhammad Aksam Iftikhar, and Tanvir Alam. Lung-effnet: Lung cancer classification using efficientnet from ct-scan images. *Engineering Applications of Artificial Intelligence*, 126:106902, 2023.

[188] Mamoona Humayun, R. Sujatha, Saleh Naif Almuayqil, and N. Z. Jhanjhi. A transfer learning approach with a convolutional neural network for the classification of lung carcinoma. *Healthcare*, 10(6), 2022.

[189] Walker H. Land, Dan Mckee, Tatyana Zhukov, Dansheng Song, and Wei Qian. An end-to-end process for cancer identification from images of lung tissue. In *2007 IEEE 7th International Symposium on BioInformatics and BioEngineering*, pages 233–239, 2007.

[190] Muntasir Mamun, Afia Farjana, Miraz Al Mamun, and Md Salim Ahammed. Lung cancer prediction model using ensemble learning techniques and a systematic review analysis. In *2022 IEEE World AI IoT Congress (AIIoT)*, pages 187–193, 2022.

[191] Snehal Dabade, Shubhangi Chaudhari, Sneha Jadhav, and Arjun Nichal. A review paper on computer aided system for lung cancer detection. In *2017 International Conference on Big Data, IoT and Data Science (BID)*, pages 97–102, 2017.

[192] Subrato Bharati, Prajoy Podder, and M. Rubaiyat Hossain Mondal. Hybrid deep learning for detecting lung diseases from x-ray images. *Informatics in Medicine Unlocked*, 20:100391, 2020.

[193] H. Abdel-Nabi, M. Ali, and A. Awajan. A comprehensive review of the deep learning-based tumor analysis approaches in histopathological images: segmentation, classification and multi-learning tasks. *Cluster Comput*, 2023.

[194] Shweta Tyagi and Sanjay N. Talbar. Lcscnet: A multi-level approach for lung cancer stage classification using 3d dense convolutional neural networks with concurrent squeeze-and-excitation module. *Biomedical Signal Processing and Control*, 80:104391, 2023.

[195] Basra Jehangir, Soumya Ranjan Nayak, and Sourav Shandilya. Lung cancer detection using ensemble of machine learning models. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 411–415, 2022.

[196] Jayakumar S. Kasinathan G. Cloud-based lung tumor detection and stage classification using deep learning techniques. *Biomed Res Int.*, 2022.

[197] Gilu K Abraham, Preethi Bhaskaran, and V. S Jayanthi. Lung nodule classification in ct images using convolutional neural network. In *2019 9th International Conference on Advances in Computing and Communication (ICACC)*, pages 199–203, 2019.

[198] Wei Shen, Mu Zhou, Feng Yang, Caiyun Yang, and Jie Tian. Multi-scale convolutional neural networks for lung nodule classification. In Sebastien Ourselin, Daniel C. Alexander, Carl-Fredrik Westin, and M. Jorge Cardoso, editors, *Information Processing in Medical Imaging*, pages 588–599, Cham, 2015. Springer International Publishing.

[199] Suren Makaju, P.W.C. Prasad, Abeer Alsadoon, A.K. Singh, and A. Elchouemi. Lung cancer detection using ct scan images. *Procedia Computer Science*, 125:107–114, 2018. The 6th International Conference on Smart Computing and Communications.

[200] Sutong Wang, Jiacheng Zhu, Yunqiang Yin, Dujuan Wang, T.C. Edwin Cheng, and Yanzhang Wang. Interpretable multi-modal stacking-based ensemble learning method for real estate appraisal. *IEEE Transactions on Multimedia*, 25:315–328, 2023.

[201] Lakshmanaprabu S.K., Sachi Nandan Mohanty, Shankar K., Arunkumar N., and Gustavo Ramirez. Optimal deep learning model for classification of lung cancer on ct images. *Future Generation Computer Systems*, 92:374–382, 2019.

[202] Shanchen Pang, Yaqin Zhang, Mao Ding, Xun Wang, and Xianjin Xie. A deep model for lung cancer type identification by densely connected convolutional networks and adaptive boosting. *IEEE Access*, 8:4799–4805, 2020.

[203] Anum Masood, Po Yang, Bin Sheng, Huating Li, Ping Li, Jing Qin, Vitaveska Lanfranchi, Jinman Kim, and David Dagan Feng. Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest ct. *IEEE Journal of Translational Engineering in Health and Medicine*, 8:1–13, 2020.

[204] Wang, Shuo, Zhou, Mu, Gevaert, Olivier, Tang, Zhenchao, Dong, Di, Liu, Zhenyu, Jie, and Tian. A multi-view deep convolutional neural networks for lung nodule segmentation. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1752–1755, 2017.

[205] Aman Agarwal, Kritik Patni, and Rajeswari D. Lung cancer detection and classification based on alexnet cnn. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, pages 1390–1397, 2021.

[206] Chun-Hui Lin, Cheng-Jian Lin, Yu-Chi Li, and Shyh-Hau Wang. Using generative adversarial networks and parameter optimization of convolutional neural networks for lung tumor classification. *Applied Sciences*, 11(2), 2021.

[207] Muhammad Asif Khan, V Rajinikanth, Suresh Chandra Satapathy, David Taniar, Jyoti Ranjan Mohanty, Usman Tariq, and Robertas Damaševičius. Vgg19

network assisted joint segmentation and classification of lung nodules in ct images. *Diagnostics (Basel)*, 11(12):2208, Nov 2021.

[208] Haiying Yuan, Yanrui Wu, Junpeng Cheng, Zhongwei Fan, and Zhiyong Zeng. Pulmonary nodule detection using 3-d residual u-net oriented context-guided attention and multi-branch classification network. *IEEE Access*, 10:82–98, 2022.

[209] A.R. Bushara, R.S. Vinod Kumar, and S.S. Kumar. Lcd-capsule network for the detection and classification of lung cancer on computed tomography images. *Multimedia Tools and Applications*, March 23 2023.

[210] Uma R. Salunkhe and Suresh N. Mali. Classifier ensemble design for imbalanced data classification: A hybrid approach. *Procedia Computer Science*, 85:725–732, 2016. International Conference on Computational Modelling and Security (CMS 2016).

[211] Stelmo Magalhães Barros Netto, João Otávio Bandeira Diniz, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Rofolfo Acatauassú Nunes, and Marcelo Gattass. Modified quality threshold clustering for temporal analysis and classification of lung lesions. *IEEE Transactions on Image Processing*, 28(4):1813–1823, 2019.

[212] Yutong Xie, Jianpeng Zhang, Yong Xia, Michael Fulham, and Yanning Zhang. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct. *Information Fusion*, 42:102–110, 2018.

[213] Abhijit Guha Roy, Nassir Navab, and Christian Wachinger. Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, pages 421–429. Springer, 2018.

[214] Mamtha V. Shetty, D. Jayadevappa, and G. N. Veena. Water cycle bat algorithm and dictionary-based deformable model for lung tumor segmentation. *International Journal of Biomedical Imaging*, 2021:Article ID 3492099, 12 pages, 2021.

[215] W. Liu, X. Liu, H. Li, M. Li, X. Zhao, and Z. Zhu. Integrating lung parenchyma segmentation and nodule detection with deep multi-task learning. *IEEE Journal of Biomedical and Health Informatics*, 25(8):3073–3081, 2021.

[216] G. Aresta et al. Automatic lung nodule detection combined with gaze information improves radiologists' screening performance. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2894–2901, 2020.

[217] J. Rocha, A. Cunha, and A.M. Mendonça. Conventional filtering versus u-net based models for pulmonary nodule segmentation in ct images. *Journal of Medical Systems*, 44(4):81, 2020.

[218] S. Zheng, J. Guo, X. Cui, R. N. J. Veldhuis, M. Oudkerk, and P. M. A. van Ooijen. Automatic pulmonary nodule detection in ct scans using cnns based on maximum intensity projection. *IEEE Transactions on Medical Imaging*, 39(3):797–805, 2020.

[219] Muhammad Usman, Byoung-Dai Lee, Shi-Sub Byon, Sung-Hyun Kim, Byung-il Lee, and Yeong-Gil Shin. Volumetric lung nodule segmentation using adaptive roi with multi-view residual learning. *Scientific Reports*, 10(1):1–15, 2020.

[220] P.H.J. Amorim, T.F. de Moraes, J.V.L. da Silva, and H. Pedrini. Lung nodule segmentation based on cnns using multi-orientation and patchwise mechanisms. In *VipIMAGE 2019*, volume 34 of *Lecture Notes in Computational Vision and Biomechanics*, pages 319–329. Springer, Cham, 2019.

[221] Z. Shi, H. Hao, M. Zhao, et al. A deep cnn based transfer learning method for false positive reduction. *Multimedia Tools and Applications*, 78:1017–1033, 2019.

[222] Guilherme Aresta, Colin Jacobs, Teresa Araújo, António Cunha, Isabel Ramos, Bram van Ginneken, and Aurélio Campilho. iw-net: an automatic and minimalistic interactive lung nodule segmentation deep network. *Scientific reports*, 9(1):1–9, 2019.

[223] Yulei Qin, Hao Zheng, Xiaolin Huang, Jie Yang, and Yue-Min Zhu. Pulmonary nodule segmentation with ct sample synthesis using adversarial networks. *Medical physics*, 46(3):1218–1229, 2019.

[224] M. Liu, J. Dong, X. Dong, H. Yu, and L. Qi. Segmentation of lung nodule in ct images based on mask r-cnn. In *9th International Conference on Awareness Science and Technology (iCAST)*, pages 1–6, 2018.

[225] H. Jiang, H. Ma, W. Qian, M. Gao, and Y. Li. An automatic detection system of lung nodule based on multigroup patch-based deep learning network. *IEEE Journal of Biomedical and Health Informatics*, 22(4):1227–1237, 2018.

[226] Botong Wu, Zhen Zhou, Jianwei Wang, and Yizhou Wang. Joint learning for pulmonary nodule segmentation, attributes and malignancy prediction. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1109–1113. IEEE, 2018.

[227] Wenzhe Wang, Yifei Lu, Bian Wu, Tingting Chen, Danny Ziyi Chen, and Jian Wu. Deep active self-paced learning for accurate pulmonary nodule segmentation. In *MICCAI*, 2018.

[228] Guofeng Tong, Yong Li, Huairong Chen, Qingchun Zhang, and Huiying Jiang. Improved u-net network for pulmonary nodules segmentation. *Optik*, 2018.

[229] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6):1856–1867, 2020.

[230] Wei Hao Khoong. Bus-unet: An ensemble u-net framework for medical image segmentation, 2020.

[231] luna16 grand challenge. Luna16 and kaggle data science bowl 2017. *https://www.kaggle.com/datasets/kmader/finding-lungs-in-ct-data*, 2017.

[232] Yiting Xie, Jianhua Zhang, Yong Xia, Michael Fulham, and Yong Zhang. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct. *Information Fusion*, 42:102–110, 2018.

[233] Reza Azad, Maryam Asadi-Aghbolaghi, Mahmood Fathy, and Sergio Escalera. Bi-directional convlstm u-net with densley connected convolutions. In *Proceedings of the IEEE/CVF international conference on computer vision workshops*, pages 0–0, 2019.

[234] Stephen LH Lau, Edwin KP Chong, Xu Yang, and Xin Wang. Automated pavement crack segmentation using u-net-based convolutional neural network. *IEEE Access*, 8:114892–114899, 2020.

[235] Syeda Furruka Banu, Md Sarker, Mostafa Kamal, Mohamed Abdel-Nasser, Domenec Puig, and Hatem A Raswan. Aweu-net: An attention-aware weight excitation u-net for lung nodule segmentation. *Applied Sciences*, 11(21):10132, 2021.

[236] Haibo Lin, Lingzhi Xia, Yunhao Zhang, Xuefeng Chen, Minzhi Chen, Huan Wang, and Junlong Zhou. Research of lung nodule segmentation algorithm based on 3d u-net network. In *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, volume 10, pages 1795–1799. IEEE, 2022.

[237] Debnath Bhattacharyya, N Thirupathi Rao, Eali Stephen Neal Joshua, and Yu-Chen Hu. A bi-directional deep learning architecture for lung nodule semantic segmentation. *The Visual Computer*, pages 1–17, 2022.

[238] Zhixun Zhou, Fangfang Gou, Yanlin Tan, and Jia Wu. A cascaded multi-stage framework for automatic detection and segmentation of pulmonary nodules in developing countries. *IEEE Journal of Biomedical and Health Informatics*, 26(11):5619–5630, 2022.

[239] Junlong Cheng, Shengwei Tian, Long Yu, Shijia Liu, Chaoqing Wang, Yuan Ren, Hongchun Lu, and Min Zhu. Ddu-net: A dual dense u-structure network for medical image segmentation. *Applied Soft Computing*, 126:109297, 2022.

[240] P. Princy Magdaline, T.R. Ganesh Babu, R. Praveena, and R. Khowshalya. A hybrid deep learning model combining agresnet, yolo, and cnn for lung tumor segmentation and classification. *Journal of Innovative Image Processing*, 6(4):472–498, 2024.

[241] Mehr Kashyap, Xi Wang, Neil Panjwani, Mohammad Hasan, Qin Zhang, Charles Huang, Karl Bush, Alexander Chin, Lucas K. Vitzthum, Peng Dong, Sandra Zaky, Billy W. Loo, Maximilian Diehn, Lei Xing, Ruijiang Li, and Michael F. Gensheimer. Automated deep learning–based detection and segmentation of lung tumors at ct imaging. *Radiology*, 2025. Deputy Editor: Mizuki Nishino; Scientific Editor: Shannyn Wolfe.

[242] Tathagat Banerjee. Towards automated and reliable lung cancer detection in histopathological images using dy-fspan: A feature-summarized pyramidal attention network for explainable ai. *Computational Biology and Chemistry*, page 108500, 2025.

[243] S.R. Vijayakumar, S. Aarthy, D. Deepa, and P. Suresh. Sustainable framework for automated segmentation and prediction of lung cancer in ct image using capsnet with u-net segmentation. *Biomedical Signal Processing and Control*, 99:106873, 2025.

[244] Jayapradha Jayaram, Su-Cheng Haw, Naveen Palanichamy, Senthil Kumar Thillaigovindhan, and Mutaz Al-Tarawneh. Title not provided. *Journal of Informatics and Web Engineering*, 4(1):Article 11, 2025. Published on 2025-02-14.

[245] S.W. Huang, C.T. Lin, S.P. Chen, Y.Y. Wu, P.H. Hsu, and S.H. Lai. Auggan: Cross domain adaptation with gan-based data augmentation. In V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, editors, *Computer Vision – ECCV 2018*, volume 11213 of *Lecture Notes in Computer Science*, pages 743–758. Springer, Cham, 2018.

[246] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[247] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[248] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019.

[249] Mwamba Kasongo Dahouda and Inwhee Joe. Neural architecture search net-based feature extraction with modular neural network for image classification of copper/ cobalt raw minerals. *IEEE Access*, 10:72253–72262, 2022.

[250] Saleh Albahli and Waleed Albattah. Detection of coronavirus disease from x-ray images using deep learning and transfer learning algorithms. *Journal of X-ray science and technology*, 28, 08 2020.

[251] Mavra Mehmood, Nasser Alshammari, Saad Awadh Alanazi, Asma Basharat, Fahad Ahmad, Muhammad Sajjad, and Kashaf Junaid. Improved colorization and classification of intracranial tumor expanse in mri images via hybrid scheme of pix2pix-cgans and nasnet-large. *Journal of King Saud University - Computer and Information Sciences*, 34(7):4358–4374, 2022.

[252] Raul Victor Medeiros da Nóbrega, Solon Alves Peixoto, Suane Pires P. da Silva, and Pedro Pedrosa Rebouças Filho. Lung nodule classification via deep transfer learning in ct lung images. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, pages 244–249, 2018.

[253] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.

[254] C Szegedy, S Ioffe, V Vanhoucke, and A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. arxi v: 1602.07261, 2018.

[255] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.

[256] Sagnik Ghosal, Mainak Sarkar, and Ram Sarkar. Nofed-net: Nonlinear fuzzy ensemble of deep neural networks for human activity recognition. *IEEE Internet of Things Journal*, 9(18):17526–17535, 2022.

[257] Zonlehoua Coulibali, Athyna Nancy Cambouris, and Serge-Étienne Parent. Site-specific machine learning predictive fertilization models for potato crops in eastern canada. *PloS one*, 15(8):e0230888, 2020.

[258] K Harmsen. A modified mitscherlich equation for rainfed crop production in semi-arid areas: 1. theory. *NJAS: Wageningen Journal of Life Sciences*, 48(3):237–250, 2000.

[259] Kashinath R Sonar and Vipin P Babhulkar. Application of mitscherlich–bray equation for fertilizer use in wheat. *Communications in soil science and plant analysis*, 33(15-18):3241–3249, 2002.

[260] Aixia Liu Zhengjung Liu and Zheng Niu Changyao Wang. Evolving neural network using real coded genetic algorithm (ga) for multispectral image classification. *Future Generation Computer Systems*, 20:1119–1129, 2004.

[261] Ahmed Shaffie, Ahmed Soliman, Hadil Abu Khalifeh, Fatma Taher, Mohammed Ghazal, Neal Dunlap, Adel Elmaghraby, Robert Keynton, and Ayman El-Baz. A novel ct-based descriptors for precise diagnosis of pulmonary nodules. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1400–1404, 2019.

[262] Hanliang, Fuhao Jiang, Fei Shen, Weidong Gao, and Han. Learning efficient, explainable and discriminative representations for pulmonary nodules classification. *Pattern Recognition*, 113:107825, 2021.

[263] Xia Y. Xie Y, Zhang J. Semi-supervised adversarial model for benign-malignant lung nodule classification on chest ct. In *Med Image Anal. 2019 Oct*, pages 237–248, 2019.

[264] Defang Zhao, Dandan Zhu, Jianwei Lu, Ye Luo, and Guokai Zhang. Synthetic medical images using f&bgan for improved lung nodules classification by multi-scale vgg16. *Symmetry*, 10:519, 10 2018.

[265] Zhang Ge, Lan Lin, and Jingxuan Wang. Lung nodule classification in ct images using 3d densenet. *Journal of Physics: Conference Series*, 1827:012155, 03 2021.

[266] Penghua Zhai, Yaling Tao, Hao Chen, Ting Cai, and Jinpeng Li. Multi-task learning for lung nodule classification on chest ct. *IEEE Access*, 8:180317–180327, 01 2020.

[267] R. Guha, M. Ghosh, S. Kapri, et al. Deluge based genetic algorithm for feature selection. *Evolutionary Intelligence*, 14(3):357–367, 2021.

[268] Pratik Bhowal, Subhankar Sen, Juan Velasquez, and Ram Sarkar. Fuzzy ensemble of deep learning models using choquet fuzzy integral, coalition game and information theory for breast cancer histology classification. *Expert Systems with Applications*, 190:116167, 11 2021.

[269] Rishav Pramanik, Momojit Biswas, Shibaprasad Sen, Luis Júnior, João Papa, and Ram Sarkar. A fuzzy distance-based ensemble of deep models for cervical cancer detection. *Computer Methods and Programs in Biomedicine*, 219:106776, 06 2022.

[270] M. Ghosh, S. Adhikary, K.K. Ghosh, et al. Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods. *Medical & Biological Engineering & Computing*, 57(1):159–176, 2019.

[271] Shyam Marjit, Trinav Bhattacharyya, Bitanu Chatterjee, and Ram Sarkar. Simulated annealing aided genetic algorithm for gene selection from microarray data. *Computers in Biology and Medicine*, 158:106854, 2023.

[272] Rishav Pramanik, Payel Pramanik, and Ram Sarkar. Breast cancer detection in thermograms using a hybrid of ga and gwo based deep feature selection method. *Expert Systems with Applications*, 219:119643, 2023.

[273] Shile Chen and Changjun Zhou. Stock prediction based on genetic algorithm feature selection and long short-term memory neural network. *IEEE Access*, 9:9066–9072, 2021.

[274] Han Chung and Kyoungsoo Shin. Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. *Neural Computing and Applications*, 32:7897–7914, 2020.

[275] Driss Saadaoui, Mustapha Elyaqouti, Khalid Assalaou, Driss Ben Hmamou, and Souad Lidaighbi. Parameters optimization of solar pv cell/module using genetic algorithm based on non-uniform mutation. *Energy Conversion and Management: X*, 12:100129, 2021.

[276] Jyoti Bhola, Sarabjeet Soni, and Gurdarshan Singh Cheema. Genetic algorithm based optimized leach protocol for energy efficient wireless sensor networks. *Journal of Ambient Intelligence and Humanized Computing*, 11:1281–1288, 2020.

[277] Gordon Opuodho. Robert Odek. F-test and p-values: A synopsis. *Journal of Management and Science*, 2023.

[278] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[279] Aditya Chattopadhay, Anirban Sarkar, Prantik Howlader, and Vineeth N Balasubramanian. Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, mar 2018.

[280] Wang Tong, Wu Fubin, Lu Haoran, and Xu Shengzhou. Ca-unet : Convolution and attention fusion for lung nodule segmentation. *International Journal of Imaging Systems and Technology*, 33, 03 2023.

[281] Imran Shafi, Sadia Din, Asim Khan, Isabel De La Torre Díez, Ramón del Jesús Palí Casanova, Kilian Tutusaus Pifarre, and Imran Ashraf. An effective method for lung cancer diagnosis from ct scan using deep learning-based support vector network. *Cancers*, 14(21), 2022.

[282] Siddharth Bhatia, Yash Sinha, and Lavika Goel. Lung cancer detection: A deep learning approach. In Jagdish Chand Bansal, Kedar Nath Das, Atulya Nagar, Kusum Deep, and Akshay Kumar Ojha, editors, *Soft Computing for Problem Solving*, pages 699–705, Singapore, 2019. Springer Singapore.

[283] Hamdalla F Al-Yasriy, Muayed S AL-Husieny, Furat Y Mohsen, Enam A Khalil, and Zainab S Hassan. Diagnosis of lung cancer based on ct scans using cnn. In *IOP Conference Series: Materials Science and Engineering*, volume 928, page 022035. IOP Publishing, 2020.

[284] Sayyada Hajera Begum, Mirza Iliyaz Baig, Mohammed Abbas Hussain, and Mohd Abdul Muqeet. A lightweight deep learning model for automatic diagnosis of lung cancer. In *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, pages 1–5. IEEE, 2022.

[285] Saleh Abunajm, Nelly Elsayed, Zag ElSayed, and Murat Ozer. Deep learning approach for early stage lung cancer detection. *arXiv preprint arXiv:2302.02456*, 2023.

[286] Mamoona Humayun, R Sujatha, Saleh Naif Almuayqil, and NZ Jhanjhi. A transfer learning approach with a convolutional neural network for the classification of lung carcinoma. In *Healthcare*, volume 10, page 1058. MDPI, 2022.

[287] Derya NARİN and Tuğba Özge ONUR. The effect of hyper parameters on the classification of lung cancer images using deep learning methods. *Erzincan University Journal of Science and Technology*, 15:258–268, 2022.

[288] Muayed S AL-Huseiny and Ahmed Sattar Sajit. Transfer learning with googlenet for detection of lung cancer. *Indonesian Journal of Electrical Engineering and Computer Science*, 22(2):1078–1086, 2021.

[289] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017.

[290] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[291] Andrew C. Leon. 3.12 - descriptive and inferential statistics. In Alan S. Bellack and Michel Hersen, editors, *Comprehensive Clinical Psychology*, pages 243–285. Pergamon, Oxford, 1998.

[292] Samuel G. Armato III, Geoffrey McLennan, Luc Bidaut, Michael F. McNitt-Gray, Charles R. Meyer, Anthony P. Reeves, Binsheng Zhao, Denise R. Aberle,

Claudia I. Henschke, Eric A. Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): A completed reference database of lung nodules on ct scans. *Medical Physics*, 38(2):915–931, 2011.

[293] Alberto Rey, Bernardino Arcay, and Alfonso Castro. A hybrid cad system for lung nodule detection using ct studies based in soft computing. *Expert Systems with Applications*, 168:114259, 2021.

[294] Hanliang, Fuhao Jiang, Fei Shen, Weidong Gao, and Han. Learning efficient, explainable and discriminative representations for pulmonary nodules classification. *Pattern Recognition*, 113:107825, 2021.

[295] Ahmed Shaffie, Ahmed Soliman, Hadil Abu Khalifeh, Fatma Taher, Mohammed Ghazal, Neal Dunlap, Adel Elmaghraby, Robert Keynton, and Ayman El-Baz. A novel ct-based descriptors for precise diagnosis of pulmonary nodules. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1400–1404, 2019.

[296] Xia Y. Xie Y, Zhang J. Semi-supervised adversarial model for benign-malignant lung nodule classification on chest ct. In *Med Image Anal. 2019 Oct*, pages 237–248, 2019.

[297] Zhang Ge, Lan Lin, and Jingxuan Wang. Lung nodule classification in ct images using 3d densenet. *Journal of Physics: Conference Series*, 1827:012155, 03 2021.

[298] Yurou Sun, Jinglei Tang, Weijie Lei, and Dongjian He. 3d segmentation of pulmonary nodules based on multi-view and semi-supervised. *IEEE Access*, 8:26457–26467, 2020.

[299] Yutong Xie, Jianpeng Zhang, and Yong Xia. Semi-supervised adversarial model for benign–malignant lung nodule classification on chest ct. *Medical Image Analysis*, 57:237–248, 2019.

[300] Min Chen, Xiaobo Shi, Yin Zhang, Di Wu, and Mohsen Guizani. Deep feature learning for medical image analysis with convolutional autoencoder neural network. *IEEE Transactions on Big Data*, 7(4):750–758, 2021.

[301] Tongxue Zhou, Su Ruan, and Stéphane Canu. A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, 3-4:100004, 2019.

[302] Yichi Zhang and Rushi Jiao. Towards segment anything model (sam) for medical image segmentation: A survey. *Elsevier*, 2024.

[303] Ebrahim Khalili, Blanca Priego-Torres, Antonio Leon-Jimenez, and Daniel Sanchez-Morillo. Automatic lung segmentation in chest x-ray images using sam with prompts from yolo. *TechRxiv*, 2024.

[304] Maciej A. Mazurowski, Haoyu Dong, Hanxue Gu, Jichen Yang, Nicholas Konz, and Yixin Zhang. Segment anything model for medical image analysis: An experimental study. *Medical Image Analysis*, 89:102918, 2023.

[305] Nhat-Tan Bui, Dinh-Hieu Hoang, Minh-Triet Tran, Gianfranco Doretto, Donald Adjeroh, Brijesh Patel, Arabinda Choudhary, and Ngan Le. Sam3d: Segment anything model in volumetric medical images. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1–4, 2024.

Nandita Gautam
11/11/25

Ram Sarkar 11/11/25

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata-700032