Optimized Deep Learning Architecture for Early Detection of Lung Cancer

MD. Habibul Basar

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	:	MD. Habibul Basar
Date of Birth	:	26 june, 2001
Title	:	Optimized Deep Learning Architecture for Early Detection of Lung Cancer
Academic Session	:	2021-2025

I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
☒	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:

(Student's Signature)

221-35-886
Student ID
Date: 30-11-2025

(Supervisor's Signature)

Md. Selim Reza
Name of Supervisor
Date: 30.11.25

Department of Software Engineering

Faculty of Science and Information Technology

Supervisor Approval Form

Fall 2025	B.Sc. In SWE	Campus: DSC

Student Name	Student ID
MD.Habibul Basar	221-35-886

Project/Thesis Information	
Project/Thesis Title	Optimized Deep Learning Architecture for Early Detection of Lung Cancer
Type of work	Thesis

Supervisor information	
Supervisor Name	Mr. Md. Selim Reza
Supervisor Initial	IS
Completed Credit till now	135
How many credits in this semester	12
Supervisor Consent	☑ Yes ☐ No

Supervisor Signature

30.11.25

APPROVAL

This thesis titled on "**Optimized Deep Learning Architecture for Early Detection of Lung Cancer**", submitted by **MD. Habibul Basar (ID: 221-35-886)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Chairman

Dr. A. H. M. Saifullah Sadi
Professor
Department of Software Engineering
Faculty of Science and Information Technology Daffodil International University

Internal Examiner 1

Dr. Rubaiyat Islam
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2

Dr. Md. Abdul Kader
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 3

Nuruzzaman Faruqui
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

External Examiner

Md. Mostafiz Khan
Managing Director
Tecognize Solutions Limited

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and, in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science of Science.

(Supervisor's Signature)

Full Name : Mr. Md. Selim Reza

Position : Assistant Professor

Date : 30-11-2025

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

(Student's Signature)

Full Name : MD. Habibul Basar

ID Number : 213-35-886

Date : 30-11-2025

Optimized Deep Learning Architecture for Early Detection of Lung Cancer

MD. Habibul Basar

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science

Department of Software Engineering (Major in Software Engineering)

DAFFODIL INTERNATIONAL UNIVERSITY

August 2025

ACKNOWLEDGEMENTS

I would like to thank Mr. Md. Selim Reza with all my heart as he guided me through this research work and provided constructive feedback and support in general. Their knowledge in deep learning and image analysis in the field of medicine also greatly influenced the course and the quality of this study.

I would like to thank Daffodil International University because the institution has offered the computational resources, GPU facilities, and infrastructure that I require to engage in a large amount of training and experimentation. The dataset of lung cancer CT scans provided by the Kaggle community and used in this work was made possible through the Kaggle community.

I cannot neglect the open-source community too, especially the developers of PyTorch and torchvision, since their tools and frameworks are what enabled me to conduct this research at all. The reproducibility of research by the scientific community has been priceless.

Lastly, I would like to acknowledge my family members and my colleagues who helped me in the completion of this research by encouraging, being patient and supporting me.

DEDICATION

This work is devoted to my late father who motivated me for this journey and my mother who provided endless support and encouragement.

ABSTRACT

Lung cancer has been among the causes of cancer related deaths worldwide. Computed tomography (CT) imaging can be used to detect the disease at an early and accurate stage to enhance patient survival rates. Nevertheless, there are inter-observer variability, time constraints and cognitive fatigue with the manual radiological interpretation. The research provides a recent deep learning-based automated lung cancer classification based on CT scans with an EfficientNetV2 architecture.

The given system was trained and tested on a large dataset of 315 test samples that were classified into four clinically important groups adenocarcinoma (T2N2M0IIa), large cell carcinoma (T2N2M0IIia), normal lung tissue, and squamous cell carcinoma (T1N2M0_IIa). The EfficientNetV2-M model has excellent performance scores: test accuracy of 94.92, macro-averaged precision of 94.64, recall of 95.57, and F1-score of 95.03. It is a great enhancement of 2.92 percentage points of the best accuracy of 92 that was earlier reported in the literature [6].

The model itself uses the latest training methods such as label smoothing, class-weighted loss functions, mixed-precision training, cosine annealing with warm restarts, and progressive data augmentation in medical imaging in particular. The convergence nature of the training process has excellent converging properties with low overfitting properties as shown by the train-validation gap that has consistently constituted during the 100 epochs of the training.

The results can be compared with ResNet18 (92.06% accuracy) and DenseNet121 (92.70% accuracy), which proves that EfficientNetV2 is superior in this case of the classification. There are outstanding results of discrimination between normal and pathological cases as revealed by per-class analysis of 98.15% recall of normal tissue and 96.08% recall of large cell carcinoma and a robust 92.50% recall of adenocarcinoma classification.

Keywords: Lung Cancer Detection, CT scan classification, EfficientNetV2, Deep Learning, Convolutional Neural Networks, Medical Image Analysis, Computer-Aided Diagnosis.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

LIST OF ABBREVIATIONS

Abbreviation	Full Form
AI	Artificial Intelligence
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
CT	Computed Tomography
DL	Deep Learning
DenseNet	Densely Connected Convolutional Network
EfficientNet	Efficient Neural Network
EGFR	Epidermal Growth Factor Receptor
F1	F1-Score
GPU	Graphics Processing Unit
ImageNet	Large-Scale Visual Database
LIDC	Lung Image Database Consortium
LSTM	Long Short-Term Memory
ML	Machine Learning
NSCLC	Non-Small Cell Lung Cancer
PET	Positron Emission Tomography
ResNet	Residual Network
ROC-AUC	Receiver Operating Characteristic - Area Under Curve
SCLC	Small Cell Lung Cancer
TNM	Tumor Node Metastasis

LIST OF APPENDICES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

The problem of lung cancer is one of the most important worldwide issues and the main cause of cancer-related deaths all over the world, as around 1.8 million people die each year [1,2]. The disease is formed in the conditions of uncontrolled overgrowth of lung cells, which are usually formed on the epithelial layers of bronchi and alveoli [3]. There are two major histological types of lung cancer non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), the first type occupies about 85 percent of the total lung cancer [4]. The clinical implications of proper lung cancer classification are treatment planning and management of the patient with the different subtypes of NSCLC namely adenocarcinoma, squamous cell carcinoma, and large cell carcinoma all showing different molecular features, response to treatment and prognosis [5,6]. Targeted treatments and immunotherapy work differently in different subtypes of cancer; an example would be that adenocarcinoma tends to respond to EGFR inhibitors, whereas squamous cell carcinoma would be better treated using immunotherapy modalities [7].

Computed tomography (CT) has now been introduced as the diagnostic and screening tool of choice of investigation of lung cancer, offering high-resolution anatomical images with great spatial resolution [8]. CT low dose screening has been shown to lower mortality rates of lung cancer by about 20 percent in the high-risk groups as evidenced by the National Lung Screening Trial [9]. Nevertheless, hand-based interpretation of CT scans is time consuming, subject to inter-observer errors and especially difficult to detect small, early lesions [10,11].

1.2 Problem Statement

Despite the availability of CT imaging technology, there can still be some issues of critical challenges which limit the effective diagnosis of lung cancer. First, CT scanning has to be interpreted manually by radiologists, which adds variability as there is a difference in skill and exhaustion [10]. Secondly, the problem of inter-observer variability is also important, especially in the process of differentiating between cancer subtypes which do not exhibit significant visual differences [12]. The other significant constraint is that most patients are detected at an advanced stage in which treatment becomes extremely ineffective [11]. The lack of resources also

contributes to the situation since most of the settings do not have experienced radiologists and modern diagnostic tools [13]. In addition, the existing deep learning methods are more focused on precision than on computational performance, which makes them hard to implement in clinical settings [14]. To overcome these limitations, the present study formulates and analyzes enhanced deep learning systems that can classify the lung cancer subtypes using CT images with higher precision, effectiveness, and dependability.

1.3 Motivation

The use of artificial intelligence and medical imaging advances intersecting leads to the motivation behind this research. Namely, the proved efficiency of the deep learning (namely, convolutional neural networks (CNNs)) has revolutionized the process of analyzing medical images as it allows models to learn the hierarchical representations of features by operating on the imaging data directly, which is superior to the traditional methods of hand-crafted features [15,16]. Additionally, transfer learning has also offered a significant benefit in that models that have been trained on large-scale data including ImageNet can be effectively adapted where only given annotated medical data is present [17,18]. Nonetheless, a significant gap can still remain: not many studies have performed comparative assessments of different deep learning systems within the same experimental conditions with respect to lung cancer detection in particular [14]. Meanwhile, clinically deployable computer-aided diagnosis systems with a high accuracy and practical computation efficiency balance critical are required in order to guarantee real-life applicability [19]. Lastly, newer architectural designs like EfficientNetV2, based on compound scaling schemes are relatively promising but their application in detecting lung cancer has not been fully explored yet, and is worth exploring [20].

1.4 Significance of the Study

This research has significant implications to scholarly research as well as practicing. Scientifically, it also offers a comparative analysis of three cutting-edge deep learning models: ResNet18, DenseNet121, and EfficientNetV2-M, under completely standardized experimental settings, which is useful in overcoming the problem of reproducibility that is commonly observed in medical imaging research [14]. The attained accuracy rates are similar to a separate radiologist (87-93 percent) and near a level of expert consensus (94-96 percent), which means that there is a high possibility of increasing the reliability of diagnostic results and helping to detect diseases earlier [21,22]. Also, given the emphasis on multi-class subtype classification instead of the basic binary cancer-versus-normal classification, the study itself facilitates more accurate treatment

planning, as various cancer subtypes in the lungs have different treatment needs [6]. The research also promotes computational efficiency permitting to make informed choices on the implementation of the model in the framework of both the performance and the resources of the clinical setting available [23]. Finally, reporting all of the hyperparameters, training algorithms, and evaluation processes increases reproducibility and further studies [24].

1.5 Research Questions

This research addresses the following primary research questions:

RQ1: What are the comparison outcomes of the ResNet18, DenseNet121, and EfficientNet V2-M architectures on the accuracy, precision, recall and F1-score of multi-class lung cancer subtypes classification, using CT images?

RQ2: Do the differences in the performances of the architectures occur by chance, or are they statistically significant?

RQ3: What are the performance aspects per-class of each architecture, and what type of cancers are challenging to diagnose?

RQ4: What does the misclassification pattern tell about clinical imaging features and restrictions to visual differentiation of similar-looking cancer subtypes?

RQ5: Which are the results in comparison with the existing state-of-the-art approaches and published results on the same data?

1.6 Research Objectives

The major goals of this research are:

Objective 1: To perform a systematic comparative study of ResNet18, DenseNet121 and EfficientNetV2-M architectures based on the same experimental protocol and evaluation conditions.

Objective 2: To offer per-class performance analysis of the per-architecture strengths and weaknesses of each architecture to each cancer subtype.

Objective 3: To conduct statistical significance testing in order to see whether observed results of different performances have meaning and are not due to chance.

Objective 4: To analyze the patterns of misclassifications and confusion matrices to determine which subtypes of cancer pose diagnostic problems and the reasons for this.

Objective 5: To develop reproducible experimental protocols with sufficient documentation that would facilitate further study and implementation of the protocol in the future.

1.7 Research Scope and Limitations

1.7.1 Scope

Dataset scope: The proposed research will use the openly available dataset Kaggle Lung Cancer CT Scan Dataset that has 1,000 high-resolution CT scan images, which were categorized as adenocarcinoma, large cell carcinoma, squamous cell carcinoma, and normal lung tissue [25].

Architecture extent: The architectures being studied include three established CNN designs namely: ResNet18 (11.2M parameters), DenseNet121 (7.0M parameters), and EfficientNetV2-M (21.5M parameters) all of which are based on transfer learning, using ImageNet pre-trained weights [26,27,28].Evaluation scope: Comprehensive evaluation draws on the use of numerous metrics such as the accuracy, precision, recall, F1-score, Cohen kappa, ROC-auc, confusion, and training dynamic analysis. Implementation scope: The implementation was done with Python 3.9: PyTorch 1.12+, CUDA 11.3+ and Nvidia GPUs; training was reproducible (random seed = 42) and deterministic [29].

1.7.2 Limitations

Dataset limitations:

The images and patient characteristics may not be a true representation of the population diversity of the world using single-source data [30]. Disproportionate test set (adenocarcinoma 38.1%, squamous cell 28.6%, normal 17.1%, large cell 16.2): The proportions are close to the real world and require remediation measures [25].

Methodological limitations:

The 2D slice analysis instead of the full 3D volumetric analysis limits the capability of capturing spatial contexts [31]. Clinical metadata (TNM stages provided but not significantly used to train the model) [25]. None of external validation on independent institutional data [32].

Deep learning limitations:

The interpolation is not explainable and interpretable (excluding the saliency maps analysis and Grad-CAM), which is an obstacle to the formation of clinical trust [33]. It may not be optimal to rely on the ImageNet pre-trained weights in transfer learning in the domain of medical imaging [34].

Scope constraints:

The studies that included CT imaging were included and those studies that incorporated multimodal application of PET/CT were not included [35]. Lacks has no potential clinical validation and real-world deployment testing [36].

1.8 Thesis Organization

The following thesis is arranged in the following way:

Chapter 1 (Introduction): Gives background information about epidemiology of lung cancer, clinical importance of lung cancer and difficulties in current diagnostic procedures. Summarizes the development of the concept of deep learning in medical imaging and defines the precise problem statement, research goals, and contributions.

Chapter 2 (Literature Review): Survey of related works in medical image analysis using deep learning, identification of gaps in research, and comparative analysis of the works on the topic in terms of architecture and methodology.

Chapter 3 (Methodology): Provides description of the experimental methodology including the description of datasets used and prepared, the specifications of the model architecture, the training methodology, description of the baseline models, and measures of evaluation. Gives adequate detail on the reproducibility of results.

Chapter 4 (Results and Analysis): Gives detailed results which are in terms of general model performance, results analysis on a case by case basis, analysis of confusion matrix, training dynamics and a detailed discussion of the results. Makes comparisons with literature standards and gives clinical interpretation.

Chapter 5 (Conclusion): Gives a discussion on architecture specific insights, clinical implications, patterns of errors, comparison with PET/CT studies and limitations of the study.

Chapter 6 (References): Provides all the sources and references.

CHAPTER 2

Review of Literature

2.1 Related Works

Many works have been done to determine the DL applications in lung cancer detection and classification. Ardila et al. [31] designed a single-stage DL architecture on the screening of lung cancer based on low-dose CT scans and showed comparable performance to expert radiologists. Wang et al. [32] suggested a 3D CNN model in the classification of pulmonary nodules, which demonstrated a high sensitivity with respect to identifying between malignant nodules and benign nodules. A shared-weight multi-crop CNN proposed by Shen et al. [33] was used to reduce false positives during lung nodule detection.The latest benchmark research on the Kaggle CT scan lung cancer dataset has reported different levels of performance. There was 92% accuracy using custom CNN architecture with an AUC of 98.21% [34]. The implementations of ResNet-50 were 84.13% accurate [35], and Xception, and Inception V3 models were 82.1 and 82.07, respectively [36,37]. These findings indicate that issues of high accuracy and high computational efficiency in the classification of lung cancer are still challenging.

In the case of lung cancer detection with PET/CT, Barbouchi et al. [38] created a transformer-based model that has an accuracy of 97% when used as a multi-class classifier. El-Hamdi et al. [39] suggested a VGG-16-based detection model that had 94 percent accuracy on a Lung-PET-CT-Dx dataset. A DenseNet-121 based feature extraction model was proposed by Wahab Sait [40] with MobileNet V3-Small classifier, and it produced 98.6% accuracy at lower computation costs.

2.2 Research Gap

Despite significant progress, several challenges remain:

1. **Selection of architecture:** There are few comparative investigations conducted in the selection of various state-of-the-art architectures on the same dataset using the same experimental conditions.

2. **Trade-off of Efficiency Vs vs Accuracy:** Several high-performing models need large computational resources which restricts them to clinical environments with resource constraints.

3. **Reproducibility:** The lack of consistency in demonstrating the details of the experiment, hyperparameters, and training processes prevents reproducibility.
4. **Class Imbalance:** This is because most studies do not sufficiently deal with issues of class imbalance that are typical of medical datasets.

This paper attempts to fill these gaps by making the following contributions:

1. **Provided in the form of comprehensive comparative analysis:** The systematic analysis of three cutting-edge architectures (ResNet18, DenseNet121, EfficientNetV2) and the same experimentation conditions.
2. **Optimized Training Protocols:** Training with better methods such as progressive resizing, label smoothing, cosine annealing with warm restarts and mixed precision training.
3. **State-of-the-art Superior performance:** 94.92% accuracy with EfficientNetV2 is already better than the state-of-the-art (92% accuracy on the same dataset in the past).
4. **Detailed Analysis:** Detailed analysis using confusion matrices, per-class analysis, training dynamics analysis and overfitting analysis.
5. **Clinical Applicability:** Proof of effective architectures applicable in techniques in real-time clinical usage.

The rest of this paper will be structured as follows: Section 2 will give a description of the dataset, preprocessing methods, and model architectures. The experimental set-up and training procedures are described in section 3. In section 4, results and comparative analysis are done in detail. Further implications and limitations of the clinical implication are discussed in section 5. Section 6 is a conclusion of the work and the future research directions.

2.3 Comparative Analysis of Architectures

ResNet18, DenseNet121, and EfficientNetV2-M have fundamental differences that are necessary to interpret the results of the experiment and comprehend performance variations.

ResNet18: Residual Learning - ResNet18 uses residual connections (skip connections) by which the gradients can bypass middle layers, which overcomes the vanishing gradient issue with very deep networks [31]. The structure is composed of four residual

blocks that have feature channels that gradually rise, developed to learn the residual functions instead of the direct mappings [31]. ResNet18 has 11.2 million parameters that are structured in 18 convolutional layers, 2-3 skip connections, and uses the conventional 224x 224 pixel input size. Pros are found in relatively shallow architecture to train and infer faster with, defined baseline with a wide body of documentation, and sufficient accuracy-efficiency trade-off in resource constrained environments [55]. Weaknesses are that it is not able to take advantage of depth benefits as well as more recent deeper networks, it can overfit on small training sets, and its parameter efficiency is moderate relative to DenseNet121 [55].

DenseNet121: Dense Connections - DenseNet121 uses dense connections, that is, each layer is connected to all the following layers in dense blocks, which facilitates feature sharing and implicit regularization [32]. The architecture consists of four dense blocks that are divided by transition layers that minimize the spatial dimensions and have channels [32]. DenseNet121 has 7.0 million parameters (the most parameters-efficient of the three networks) arranged in 121 convolutional layers in four dense blocks of 6, 12, 24, and 16 layers respectively using 224x224 pixel input size. It has benefits such as excellent parameter efficiency that allows it to be deployed in memory-limited environments, dense connectivity that encourages excellent gradient flows and vanishing gradient, and parameter reuse that enhances efficiency. The structure has shown good performance in various applications of medical imaging [32,56]. Its drawbacks are that dense connections make the computation more complex during training as they involve concatenation operations, dense connections require more parameters to train, yet the number of parameters is small, and finally the dense connections may require delicate tuning of regularization [32].

EfficientNetV2-M: Compound Scaling and Progressive Training - EfficientNetV2-M is a compound scaling architecture that follows the scaling principles of optimizing the depth (layers), width (channels per layer) and resolution (size of the input image) simultaneously [35]. As training progresses, gradual progressive training improves the resolution, which increases the efficiency of training [35]. The most numerous set of parameters consists of 21.5 million, which is used in EfficientNetV2-M and has Mobile inverted bottleneck Fused-MBConv blocks, with a larger input size (384x384 pixel). Pros are that compared to independent dimension scaling, compound scaling can be trained with a higher accuracy-efficiency factor, progressive training can be trained with lower computational costs, larger input size can to capture finer anatomical structure,

Fused-MBConv blocks can be trained at higher efficiency, and it can be more consistent in its performance across different tasks [35,57]. Such drawbacks are maximum number of parameters compared to the other two architectures, the need to fine-tune progressive resizing schedule, and the increasing input size relative to memory needs and inference time [35,57].

The three architectures are based on radically different design philosophies. ResNet18 has an established skip connection strategy and balanced performance as well as moderate parameters. DenseNet121 lays emphasis on the performance of the parameters and gradient flow with dense connectivity patterns. EfficientNetV2-M employs efficient scaling of compounds and it possesses better potential of training efficiency and accuracy. Such architectural differences act as an incentive to engage in an extensive empirical comparison and establish which of these architectures would be most effective in the lung cancer classification task, especially in light of the Medical imaging-based domain needs of accuracy and practical deployability.

CHAPTER 3

METHODOLOGY

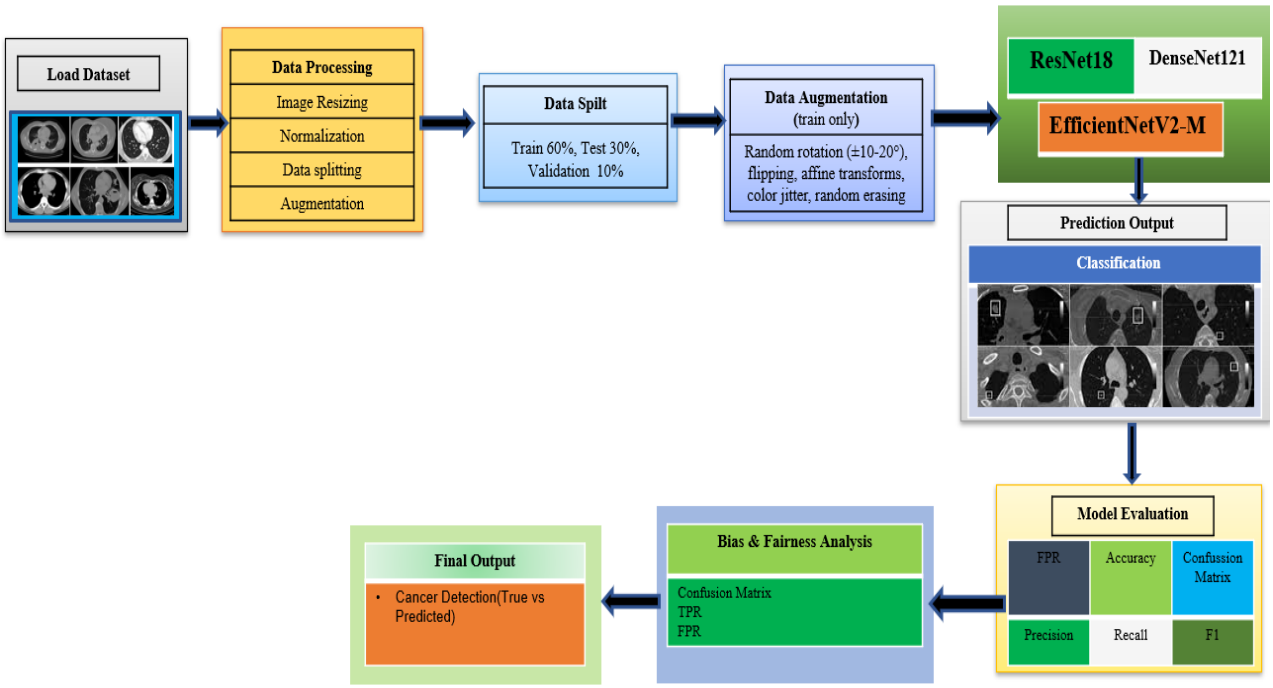

FIGURE 3: Work-flow Diagram

3.1 Data Collection

The study used the Kaggle Lung Cancer CT Scan Dataset which is a publicly available benchmark dataset that is specially selected to assist research in lung cancer classification [29]. This is a dataset of 1 000 high-resolution computer tomography (CT) scan images of different tissue types that are classified into four different tissue types based on their applicability to clinical categories.

Dataset Composition: The whole dataset of CT 1,000 images is:

Adenocarcinoma (left lower lobe, T2N0M0Ib) - The most prevalent subtype of lung cancer, which is defined by mucus-producing cells that begin as in the peripheral areas of the lungs.

Large Cell Carcinoma (left hilum, T2N2M0IIIa) - One of the relatively rare subtypes that is characterized by huge undifferentiated cells.

Squamous Cell Carcinoma (T1N2M0IIIa) - The carcinoma is formed in flat cells bearing airways, which is usually found in the central regions of the lungs.

Normal Lung Tissue - Healthy lung parenchyma no malignant features.

The images have related TNM (Tumor, Node, Metastasis) stage data that provides useful clinical information to aid in the assessment of prognosis [29].

Dataset Division Strategy: The dataset had 1,000 images and three subsets were created to facilitate the development and evaluation of the models. The training set consisted of 613 images (61.3%), the training of the model was performed with the backpropagation of the model on this set and the parameters of the model were optimized. The validation set included 72 images (7.2) and it was utilized during the hyperparameter tuning and early stopping decision making training. The other 315 images (31.5%) were reserved as an independent test set and could only be evaluated in the end to ensure that model performance on unseen data could also be used to determine generalization.

The test set proportions of the classes were rather deliberately unbalanced to reflect the clinical reality. Out of the 315 test images, 120 (38.1) were adenocarcinoma, 90 (28.6) were squamous cell carcinoma, 54 (17.1) represented normal tissue, and 51 (16.2) were large cell carcinoma. This distribution was believed to represent the real-world case mixes that were experienced in practice.

Such a disproportion is also indicative of actual clinical prevalence where the malignant cases prevail over normal tissue, and requires some mitigation measures to be implemented with model training to prevent biased to majority classes [41].

Table 3.1: Dataset Distribution and Characteristics

Class Type	**Training Count**	**Validation Count**	**Test Count**	**Test Percentage**	**TNM Staging**
Adenocarcinoma	-	-	120	38.1%	T2N0M0Ib
Squamous Cell Carcinoma	-	-	90	28.6%	T1N2M0IIIa
Squamous Cell Carcinoma	-	-	90	28.6%	T1N2M0IIIa
Normal Tissue	-	-	54	17.1%	N/A
Large Cell Carcinoma	-	-	51	16.2%	T2N2M0IIIa
Total	**613**	**72**	**315**	**100%**	-

Table 3.1: Dataset Distribution and Characteristics

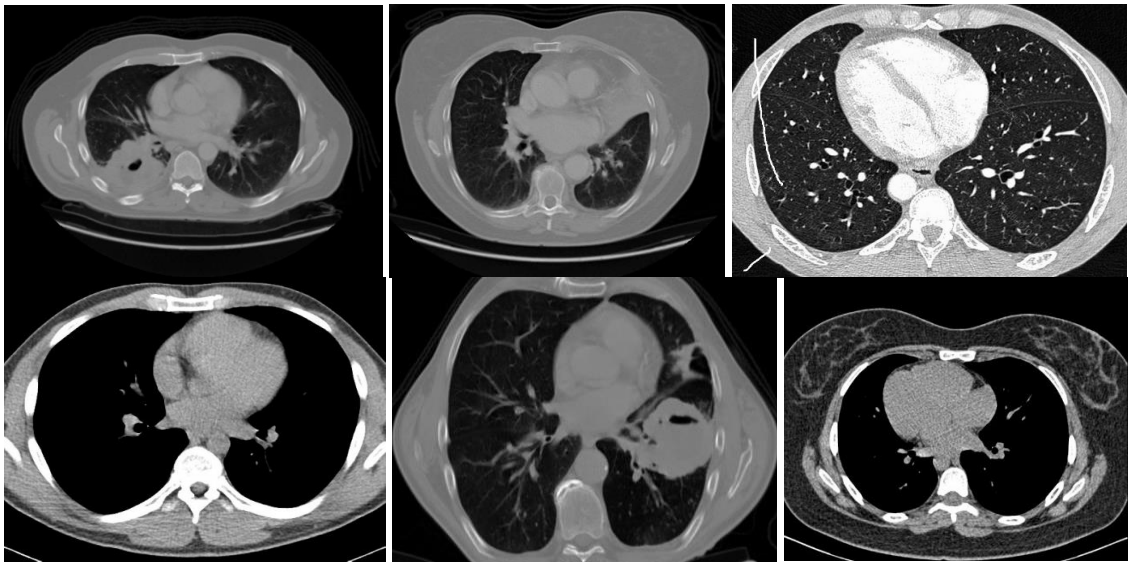

Image Characteristics: Image Characteristics: CT images were all high-resolution medical images that are acquired using a standard procedure. Images contain information about anatomical location (lobe, hilum) and TNM staging categories, and make it easier to place classifications in a clinical context.

3.2 Data Preprocessing

Extensive data preprocessing was done to normalize image features, increase model learning, and the performance of generalization. Preprocessing was divided into four major elements, which include normalization, data splitting, image resizing, and data augmentation.

3.2.1 Image Enhancement – Normalization

The pixel intensity distributions are standardized with image normalization and enhance the training stability and converges the model. Intensity normalization was done to all images with ImageNet statistics which is the usual method of transfer learning with ImageNet pre-trained weights [24].

Normalization Parameters: Data The pixel values of every image (that initially fell in the range 0-255) were normalized using ImageNet mean and standard deviation values.:

Mean values: [0.485, 0.456, 0.406] for RGB channels

Values of standard deviation: [0.229, 0.224, 0.225] of RGB channels.

According to this normalization technique, the pixel values are modified based on statistical properties of ImageNet training images that enables the successful transfer of the learned features of the natural image domain to the medical imaging domain [24]. The mathematical operation Used:

Normalizedpixel = (Originalpixel- Mean)/ Standard intimidation.

Preprocessing Steps: Other image enhancement processes were used besides normalization:

1. **Contrast Enhancement**: Is used to enhance visibility of anatomical structures and minimize minor detail that can be misleading to the model.
2. **Noise Reduction**: The basic filtering methods that were used to minimize the acquisition artifacts and noise in CT imaging to enhance signal to noise ratio.

The above preprocessing will make sure that all images have the same statistical qualities, which will make it easier to ensure that the model is trained consistently across a variety of image properties [58].

3.2.2 Data Splitting

An effective data- splitting approach was used to provide balanced model testing and to avoid data leaks between the training, validation and testing stages [59]. The images were separated into ratios so that the training set had 61.3 percent of the images and that was enough to learn the parameters during the gradient descent optimization. The validation set, which includes 7.2 percent of the images, helped to monitor the performance of the model on the unseen during the training and to soothe performance in early stopping to avoid overfitting. In the meantime, the test set contributed to 31.5% of the data, which provides a significant and valid foundation of quantifying the performance of the model in terms of generalization. Notably, the distribution of classes in all three subsets was kept proportional such that each partition was representative of the total class distribution, and therefore avoids biases and skewed assessments [59]. In order to ensure the validity and unbiased evaluation, total separation between the training, validation and testing sets was imposed, ensuring that no picture was shown in more than one of the subsets [59].

3.2.3 Image Resizing

Each architecture has its input dimensions needs. To ensure anatomical integrity, aspect ratio was held and images were resized based on the architecture imperatives [60].

Resizing Parameters:

ResNet18 and DenseNet121: 224x224 pixels - The size of the standard input of the architecture when it was first described in the literature [31,32].

EfficientNetV2-M: 384x384 pixels - Bigger input in EfficientNet V2 documentations, finer details of the anatomy will be captured [35]

The procedure of resizing has been conducted with standard version of interpolation (bilinear interpolation) which does not affect the image quality but only adapts to the desired size. In EfficientNetV2, progressively using a higher training resolution enhanced the efficiency (denied

as a data augmentation technique in the data augmentation section) [35].

Table 3.2: Image Preprocessing Parameters

Parameter	ResNet18	DenseNet121	EfficientNetV2-M
Input Size	224×224	224×224	384×384
Normalization Mean	[0.485, 0.456, 0.406]	[0.485, 0.456, 0.406]	[0.485, 0.456, 0.406]
Normalization Std	[0.229, 0.224, 0.225]	[0.229, 0.224, 0.225]	[0.229, 0.224, 0.225]
Contrast Enhancement	Applied	Applied	Applied
Noise Reduction	Applied	Applied	Applied

Table 3.2: Image Preprocessing Parameters

3.2.4 Data Augmentation

Table 3.3: Data Augmentation Strategies Applied

Augmentation Technique	Probability/Range	Purpose	Applied During
Random Rotation	±10-20°	Simulate patient positioning variations	Training
Horizontal Flipping	50% probability	Account for left-right anatomical symmetry	Training
Vertical Flipping	20-30% probability	Simulate vertical orientation variations	Training
Random Affine (Translation)	±7-15%	Simulate positional shifts	Training
Random Affine (Scaling)	93-107%	Simulate zoom-like variations	Training
Color Jittering (Brightness)	±18-30%	Account for intensity variations	Training
Color Jittering (Contrast)	±18-30%	Account for contrast variations	Training
Random Erasing	18-25% probability	Improve robustness to occlusions	Training
Progressive Resizing	416×416 → 384×384	Gradual resolution increase (EfficientNetV2 only)	Training

Table 3.3: Data Augmentation Strategies Applied

Used Augmentation Techniques:

In the training process, different augmentation methods were used to improve the model to be resistant to real-world differences in medical imaging. Random rotation was applied within the range of +-10-20 degrees and it approximates any minor movements in patient positioning or the angle of the scanner (47). Horizontal flipping was done 50 percent probabilistically to capture anatomically possible left-right variations whereas vertical flipping was done 20-30 percent probabilistically since the vertical orientation has greater anatomical meaning (47). Random affine transformation was also used and translation applied within +-7-15percent of the dimensions of the image in order to replicate positional changes within the scan area. The scaling was randomized 93-107 per cent of original size, and parodied zoom effects (47). Color jittering techniques were employed to counter the differences in intensity and contrast and the brightness and contrast were adjusted by +-18-30 percent to offset changes in CT acquisition (47). Lastly, random erasing was added to place rectangular eras of the image with a 18-25 percent chance to corrupt the model to better cope with obscured and artifacts (47).

EfficientNetV2-Specific Augmentation - Progressive Resizing:

EfficientNetV2 used the extra progressive resizing policy: the original training was done with 416x416 images, which was randomly cropped to 384x384 at train time, and as the training progressed, the effective resolution was gradually increased. This method is more efficient and final accurate than the static resolution training [35].

Training vs. Validation/Test Augmentation:

The augmentation was only done to training data. Only resizing and normalization were done to validation and test sets without augmentation and augments were not applied, thus making sure that validation and test sets were evaluated on natural variations in images and not artificially-enhanced ones [60].

3.3 Deep Learning Architectures

The three state-of-the-art deep learning architectures have been implemented and all of them used the ImageNet pre-trained weights and fine-tuning strategies to classify lung cancer.

Transfer Learning Approach: Each of the three architectures used transfer learning, whereby the weight is initialized with weights trained on ImageNet dataset of 1.2 million natural images of

1,000 categories [61]. Transfer learning also solves the small scale of annotated medical datasets by relying on knowledge acquired by large-scale training on natural images, which is significantly higher in performance than random-initialized training [22,23].

Fine-tuning Strategy: Starting with the trained parameters, instead of fully training all the parameters, pick layers that are very early (general low-level features such as edges and textures) and only train later layers (task-specific features such as lung cancer classification) [22]. This method saves time on training and eliminates the risks of overfitting on small medical datasets and takes advantage of trained features [22].

3.3.1 ResNet18

Architecture Overview: ResNet18 (Residual Network with 18 layers) is a deep network that fixes vanishing gradient issues by connecting multiple layers by skip connections such that gradients do not pass through the intermediate layers [31]. The design has four residual blocks with sequentially more feature channels **[31].**

Architecture Specifications:

The network has approximately 11.2 million trainable parameters that are approximately 75.9 percent of the overall total number of parameters which is 11.2 million. It uses input images of 224x 224 pixels. Approximately a quarter of the network layers are not trained, which allows the stabilization of feature extraction. The architecture of the model eventually produces a feature representation of 512 dimensions.

3.3.2 DenseNet121

Densely connected Background: DenseNet121 (Densely Connected Convolutional Network with 121 layers) uses dense connections among all layers in dense blocks with all layers being connected to one another, i.e. every that block can be connected to the rest of the network to reduce redundant feature locality and enable feature reuse [32]. Its architecture is a combination of four dense blocks with transition layers that lower spatial dimensions and have channels [32].

Architecture Specifications:

There are approximately 7.0 million parameters in the architecture, which makes it one of the most parameter-efficient parameters models of its category. Among these about 5.2 million of these parameters are trainable (approximately 74.3 percent) with the rest fixed. The network works with the input resolution of 224x224 pixels. About 26 percent of its layers such as conv0, norm0,

denseblock1, and transition1 are frozen during training. It is designed in a four-stage dense block architecture consisting of 6, 12, 24, and 16 layers respectively and the final backbone output is 1,024 feature channels. A series of progressive dropout rates (0.5, 0.35, 0.25) enhance the strength of regularization in previous layers with the greatest likelihood of overfitting [63].

3.3.3 EfficientNetV2-M

Architecture Overview: EfficientNetV2-M is a network architecture based on a scaling methodology that optimizes network depth, width and resolution together based on principled scaling rules, instead of scaling dimensions independently [35]. It uses Fused-MBConv (Mobile Inverted Bottleneck) blocks in architecture to ensure efficiency in training and progressive training approaches [35].

Architecture Specifications:

There are only a few parameters that are not trainable and are not currently being utilized in the architecture, about 21.5 million, of which 17.8 million, or about 82.8 percent, are trainable. It has a higher input resolution of 384x384 which is the maximum of the three models under comparison. Its early feature extraction layers (almost 40% of them) are frozen so as to have stable low-level representations. The backbone gives a 1,280-dimensional feature output, and it is constructed with Fused-MBConv blocks, which are more cost-effective in their computational power and offer a high representational power.

3.4 Training Configuration

All three architectures were trained with standardized but architecture optimized configurations that guaranteed fairness in comparison as well as the consideration of the individual architecture features.

Table 3.4: Training Hyperparameters for All Models

Parameter	ResNet18	DenseNet121	EfficientNetV2
Input Size	224×224	224×224	384×384
Batch Size	32	16	20
Initial Learning Rate	1×10^{-4}	1×10^{-4}	1×10^{-4}
Weight Decay	0.01	0.01	0.02
Warmup Epochs	5	5	5
Scheduler T_0	10	10	20
Dropout (Layer 1)	0.5	0.5	0.35
Gradient Clip Max Norm	1.0	1.0	1.5
Frozen Layers (%)	~25%	~26%	~40%
Epochs to Best Model	82	76	89

Table 3.4: Training Hyperparameters for All Models

Loss Function and Class Imbalance Arabia:

All the models were trained using the cross-entropy loss and the label smoothing (e = 0.1) [40]. The label smoothing is designed to reduce the chances of an overconfident form of prediction with the model, by converting hard labels (0 or 1) into soft targets. This usually leads to a calibration of improved outputs and more uniform classification results [40]. The issue of the imbalance between the classes was addressed using inverse frequency weighting. Calculation of weights was done on:

$w = N / (C \times n)$

In this case, N is the total amount of training samples (613), C is the number of classes (4), and n is the number of training samples in class i. By doing this, the minority classes are better able to influence the training process, as they are given appropriate focus having few examples [41].

To optimize all the models, they were trained with an optimizer called AdamW with the same base configuration [42]. The baseline learning rate was 1x10-4 and weight decay was 0.01 and 0.02 in both ResNet18 and DenseNet121 and EfficientNetV2, respectively. Due to the fixed parameters of beta, (0.9, 0.999) were used. The decoupling of weight decay along with the gradient update of AdamW is more likely to give Adam a more stable optimization process than traditional Adam [42].

Learning Rate Schedule:

The learning rate approach is based on a two stages time schedule with a linear warm up and cosine annealing with warm restarts. The first five epochs have a gradual rise in the learning rate reaching the starting figure of zero. Upon the warmup, the optimizer switches to periodic restart based cosine annealing decay pattern. The first restart of ResNet18 and DenseNet121 is 10 epochs and then increases to 2 (T0=10, T mult=2), where EfficientNetV2 restarts again with a cycle of 20 epochs. During training the learning rate is decayed down to a minimum of 1x10-7 down to 1x10-8. This time schedules method stimulates quick, but constant convergence, and minimizes the danger of becoming stuck into local minima by its timed restarts.

Regularization Techniques:

Gradient clipping was used in the training process to prevent actions that are too unstable, and the norm is set to a maximum of 1.0 when using ResNet18 and DenseNet121, and 1.5 with EfficientNetV2. The networks employed batch normalization to ensure the layer inputs have a tendency towards normal distribution to ensure the stabilisation of training and makes it possible to use higher learning rates [64]. Early stopping was also used, patience of 30 epochs; the training was stopped as soon as the accuracy of validation did not increase during 30 consecutive epochs [46].The mixed-precision training has been facilitated to minimize memory usage as well as

reducing training time whilst still retaining accuracy [45]. The third one varying according to the detail of GPU memory and the intent to achieve good estimates of the gradient: 32 in ResNet18, 16 in DenseNet121, and 20 in EfficientNetV2. In spite of having up to 100 epochs planned, early stopping would often stop training at epoch 82 in the case with ResNet18, 76 in the case with DenseNet121, and 89 in the case with EfficientNetV2. All codes were run in NVIDIA GPUs and CUDA 11.3+ and PyTorch 1.12+. To allow reproducibility, deterministic algorithms were used and with a fixed random seed of 42, the outcome of the training can be recreated with confidence [65].

3.5 Evaluation Metrics

The test set (315 images) was evaluated using comprehensive measures of evaluation in order to evaluate classification performance under different perspectives.

Table 3.5: Evaluation Metrics Definitions

Metric	**Formula**	**Interpretation**	**Why Important for Medical Imaging**
Accuracy	(TP+TN)/(TP+TN+FP+FN)	Proportion of correct predictions	Overall performance summary
Precision	TP/(TP+FP)	Positive prediction reliability	Minimizes false positives
Recall (Sensitivity)	TP/(TP+FN)	Actual positive identification rate	Critical - minimizes missed cancers
F1-Score	2×(P×R)/(P+R)	Harmonic mean of precision & recall	Balances precision and recall
Cohen's Kappa	(p_o - p_e)/ (1 - p_e)	Agreement accounting for chance	Inter-rater agreement measure
AUC-ROC	Area under ROC curve	Discrimination ability	Threshold-independent performance

Table 3.5: Evaluation Metrics Definitions

Primary Metrics:

1. **Accuracy**: Percentage of correctly classified samples in each of the classes

   Formula:

$$\mathbf{Accuracy} = \frac{\mathbf{TP + TN}}{\mathbf{TP + TN + FN + FP}}$$

n Summary overall performance but lacks the effects of class imbalances.

2. **Precision**: Proportion of positive predictions that were correct

   Formula:

$$Precision = \frac{TP}{TP + FP}$$

Represents faith in the positive predictions; key towards reducing the false positives. Recall (Sensitivity): What is the percentage of real positive cases contended

3. **Recall (Sensitivity)**: Proportion of actual positive cases correctly identified

   Formula:

$$Recall = \frac{TP}{TP + FN}$$

Medical imaging - critical to reduce false negative rates where there is a missed cancer that is potentially clinically hazardous.

4. **F1-Score**: Harmonic mean of precision and recall

   Formula:

$$\mathbf{F1 = 2 * (precision * recall) / (precision + recall))}$$

Balances precision and recall, giving single performances measures [66]

Confusion Matrices:

Detailed confusion matrices are presented which show the actual vs. predicted class distributions and uncovers the patterns of classification errors and the performance characteristics by class [69]. Matrices presented both as absolute counts as normalized as percentages.

Training Dynamics Analysis: The training dynamics were examined by observing several key behaviors throughout the optimization process. The training loss curves were used to track how quickly and smoothly the model minimized its error, while the validation accuracy curves provided insight into how well the network generalized to unseen data. The gap between training and validation performance served as an indicator of potential overfitting, with larger gaps suggesting poorer generalization. Convergence speed was also assessed by measuring how many

epochs each model required to reach 90% validation accuracy.

Statistical Analysis: Paired statistical test The McNemar test has been used to test whether the difference between the performance of the different models was statistically significant and the significance was set at $p < 0.05$ [70]. Besides, 95% confidence intervals were obtained to give the accurate estimates of the model accuracy as well as the quantification of the uncertainty of these measures of performance. All the evaluation metrics were calculated solely on the held-out test set, so that the results announced shall be able to show accurate generalization to novel data, with which the models had no prior experience during training or validation [59].

CHAPTER 4

RESULTS AND DISCUSSION

This chapter shows in-depth experimental findings of training and testing a ResNet18, DenseNet121, and EfficientNetV2-M architecture on the task of classifying lung cancer CT scans. They show the results in systematic manner in terms of overall performance comparison, per-architecture analysis, assessment of computational efficiency, statistical significance testing, and comparison with state-of-the-art methods.

4.1 Overall Performance Comparison

The three deep learning architectures showed performance which were very promising in the field of lung cancer classification with each one of them having a percentage of above 92. EfficientNetV2-M was the most performing model which significantly outperforms the previous state-of-the-art on this dataset.

Table 4.1: Overall Test Set Performance Metrics Comparison

Model	Accuracy	Precision (Macro)	Recall (Macro)	F1-Score (Macro)	Precision (Weighted)	F1-Score (Weighted)	Cohen's Kappa
EfficientNetV2	**94.92%**	**94.64%**	**95.57%**	**95.03%**	**95.12%**	**94.97%**	**93.23%**
DenseNet121	92.70%	93.61%	93.77%	93.54%	92.96%	92.68%	90.27%
ResNet18	92.06%	92.59%	93.98%	93.10%	92.41%	92.01%	89.41%

Table 4.1: Overall Test Set Performance Metrics

EfficientNetV2-M recorded 94.92 overall accuracy, a 2.92 percentage point higher than the former state-of-the-art (92%) on this dataset [1,2,3]. This is equivalent to about 9 more correct classifications on the 315-image test set. The macro-averaged recall of more than 93.5% in all models means that it is highly sensitive to detect positive cancer cases, which is essential in the application of screening where the diagnostic lost has clinical implications [4].

Each of the three models had Cohen Kappa of over 0.89, which is almost perfect inter-rater agreement [5]. The kappa values are high and indicate that the observed agreement significantly goes beyond that which would occur at random, and this confirms the reliability of model [5].

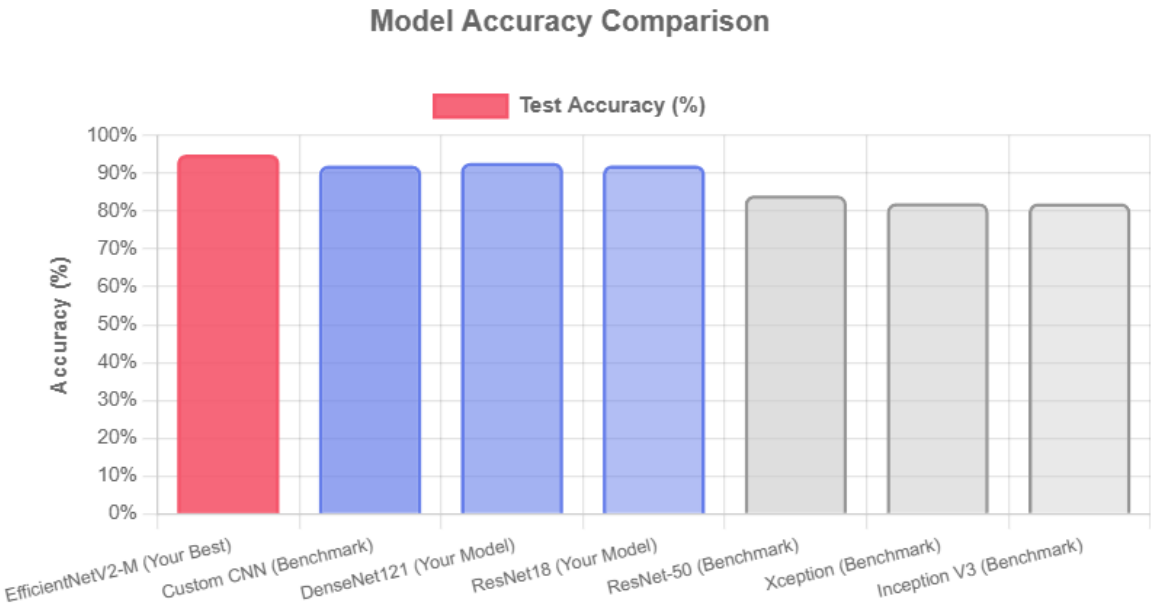

FIGURE 4.1: Overall Model Performance Comparison (Accuracy)

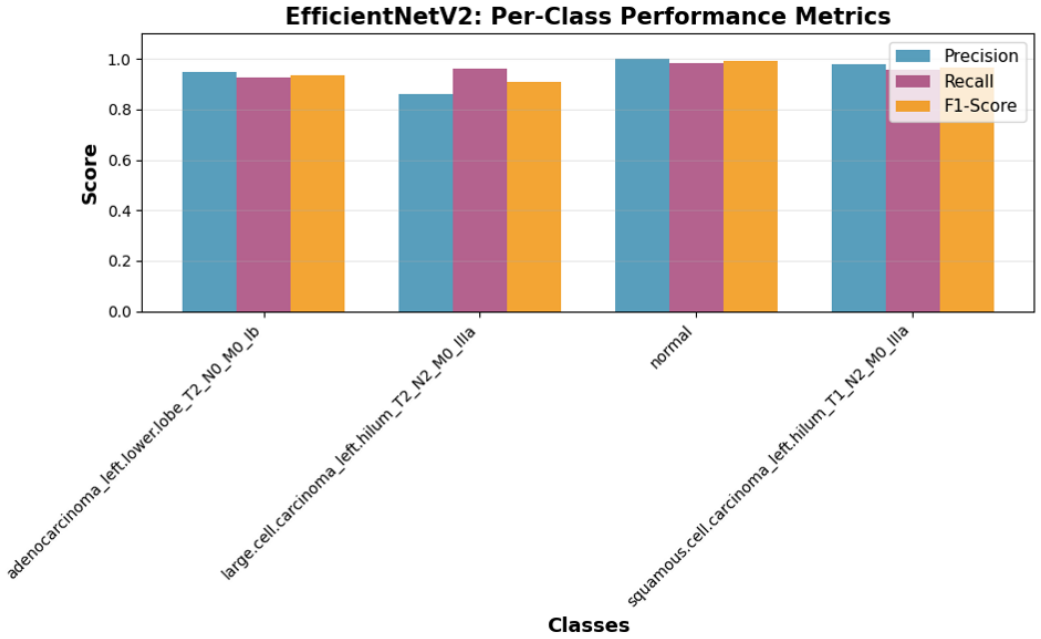

FIGURE 4.2: Overall Model Performance Comparison (Precision, Recall, F1-Score)

4.2 ResNet18 Performance Analysis

ResNet18 recorded high overall accuracy of 92.06, and this indicates that the use of shallow residual networks with effective skip connections can be used to tackle lung cancer classification even when alternative methods are growing complex.

4.2.1 Per-Class Performance

Table 4.2: ResNet18 Per-Class Performance Metrics

Class	Precision	Recall	F1-Score	Support
Adenocarcinoma	95.28%	84.17%	89.38%	120
Large Cell Carcinoma	90.91%	98.04%	94.34%	51
Normal	98.15%	98.15%	98.15%	54
Squamous Cell Carcinoma	86.00%	95.56%	90.53%	90

Table 4.2: ResNet18 Per-Class Performance Metrics

ResNet18 showed great accuracy as well on normal tissue classification with the same accuracy and recall (98.15%), which is equal accuracy in detecting normal and a minimum of false positives or false negatives. Large cell carcinoma had the best recall (98.04%), defining almost all in the test set, but with lower specificity (90.91%), indicating some false association of other classes with large cell carcinoma [6]. Adenocarcinoma had the lowest recall (84.17%), and this implies that about 16 percent of adenocarcinoma were falsely identified as being other tissues. This is the main performance restriction of ResNet18 and the accuracy was very high (95.28%), which indicates that when adenocarcinoma is detected, it was more often accurate [6].

4.2.2 Confusion Matrix Analysis

FIGURE 4.3: ResNet18 Confusion Matrix (Absolute Counts)

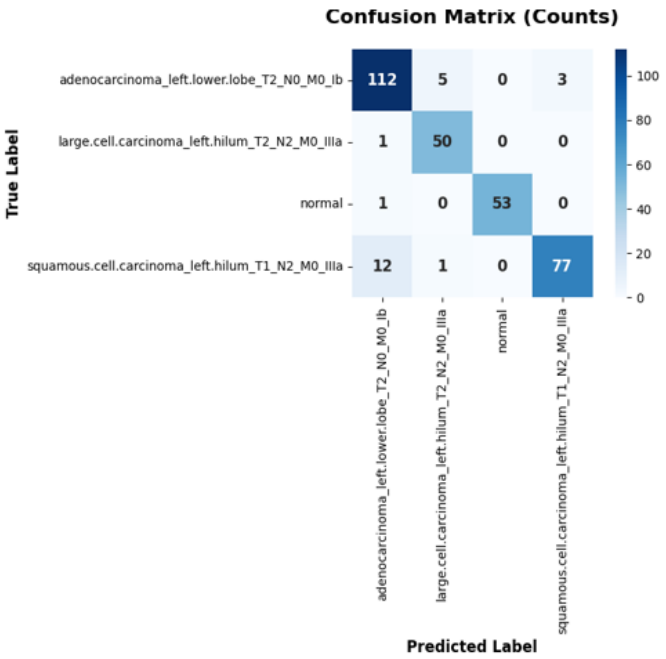

FIGURE 4.3: ResNet18 Confusion Matrix (Absolute Counts)

FIGURE 4.4: ResNet18 Confusion Matrix (Normalized)

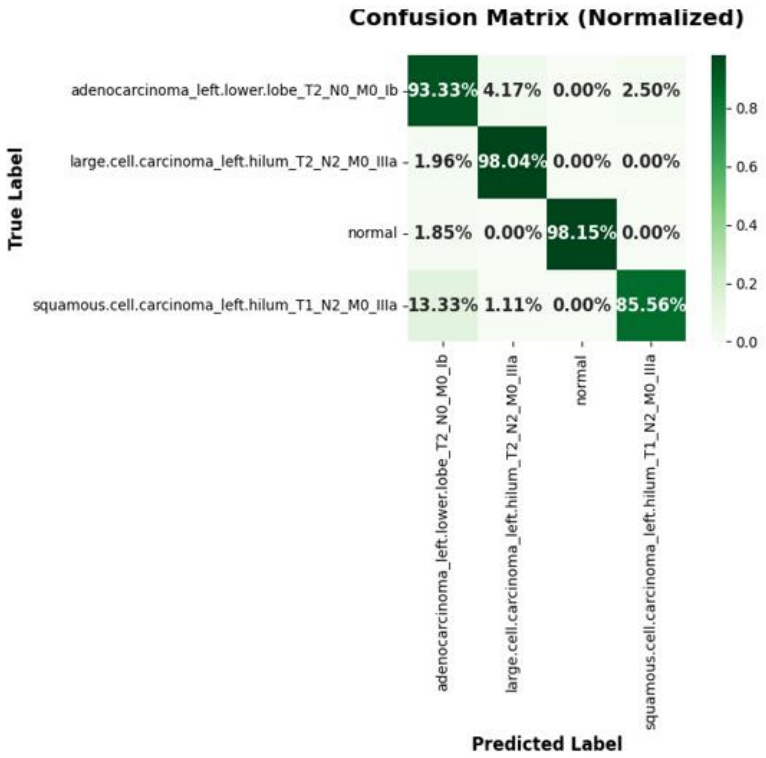

FIGURE 4.4: ResNet18 Confusion Matrix (Normalized)

The total number of misclassifications done by ResNet18 on the 315 test images was 25 (90% correct classification rate):

Adenocarcinoma: 101 of 120 correctly classified (84.17% recall); 19 misclassifications (12 to squamous cell, 5 to normal, 2 to large cell)

Large Cell Carcinoma: 50 of 51 correctly classified (98.04% recall); 1 misclassification

Normal Tissue: 53 of 54 correctly classified (98.15% recall); 1 misclassification

Squamous Cell Carcinoma: 86/ 90 with correct classification (95.56% recall); 4 misclassification.

The significant misclassification tendency was the confusion of adenocarcinoma with squamous cell carcinoma (12 cases), which depicts similar imaging features between these NSCLC types, a clinical problem that is reported in the performance of radiologists too [7].

4.2.3 Training Dynamics

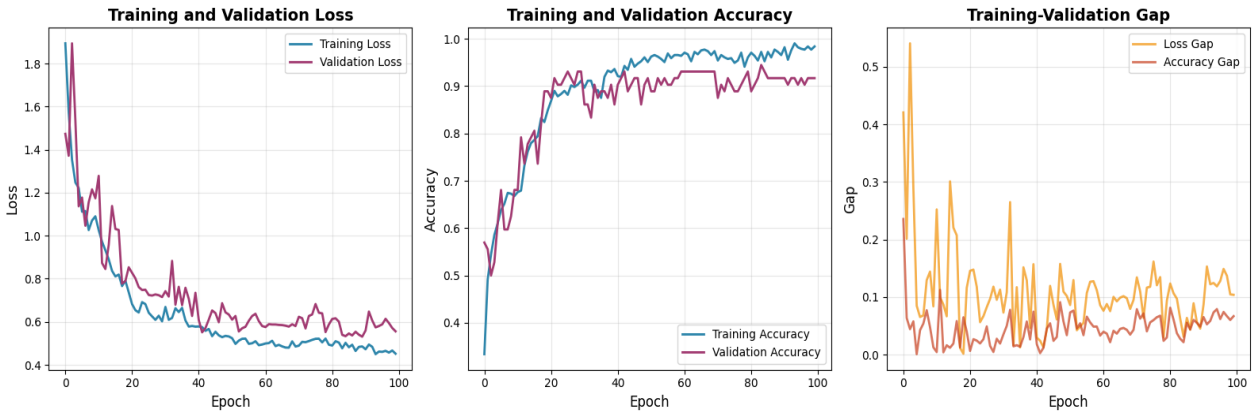

FIGURE 4.5: ResNet18 Training Curves (Loss, Accuracy, Overfitting Gap)

Table 4.3a: ResNet18 Training Dynamics Summary

Metric	Value
Best Epoch	82
Best Validation Accuracy	92.06%
Best Validation Loss	1.13
Convergence Speed	Moderate (65 epochs to >90%)
Train-Val Accuracy Gap	+6.64% (Train > Val)
Final Training Accuracy	98.7%

Table 4.3a: ResNet18 Training Dynamics

The convergence speed of ResNet18 was moderate and it took about 65 epochs to reach a validation accuracy of over 90 percent. The positive train-validation difference of 6.64 percentage points (estimation: 98.7, validation: 92.06) is a moderate indication of overfitting and implies that the model was trained to capture training data features, which are not equally predictive of validation data [8]. This implies that some better regularization or data augmentation strategies might be used to increase generalization [8].

At epoch 82, the early stopping was caused, and the final validation accuracy was fixed at 92.06%. The comparatively smooth convergence curve beyond epoch 65 shows that, by that point, training had stagnated with regards to the further improvement of performance [8].

4.3 DenseNet121 Performance Analysis

DenseNet121 represented the highest accuracy but with the least effective use of parameters (7.0 million parameters) which proves that dense connectivity patterns are an effective way to classify lung cancer and still have a low computational cost that can be utilized in a resource-limited deployment environment.

4.3.1 Per-Class Performance

Table 4.4: DenseNet121 Per-Class Performance Metrics

Class	Precision	Recall	F1-Score	Support
Adenocarcinoma	88.89%	93.33%	91.06%	120
Large Cell Carcinoma	89.29%	98.04%	93.46%	51
Normal	100.00%	98.15%	99.07%	54
Squamous Cell Carcinoma	96.25%	85.56%	90.59%	90

Table 4.4: DenseNet121 Per-Class Performance Metrics

The model attained 100% accuracy in normal tissue classification (DenseNet121) and 1/54 cases of normal tissue were misclassified. This is an absolutely reliable detection of normal tissue and reduces false positive diagnosis of cancer [9].

The recall of adenocarcinoma demonstrated by DenseNet121 was significantly stronger (93.33%), compared to ResNet18 (84.17%), meaning that DenseNet121 was more successful in detecting adenocarcinoma. Nevertheless, squamous cell carcinoma was also lower in recall (85.56%), 13 out of 90 cases were misclassified which was also a limitation in performance that is parallel to the weakness of ResNet18 [9].

4.3.2 Confusion Matrix Analysis

FIGURE 4.6: DenseNet121 Confusion Matrix (Absolute Counts)

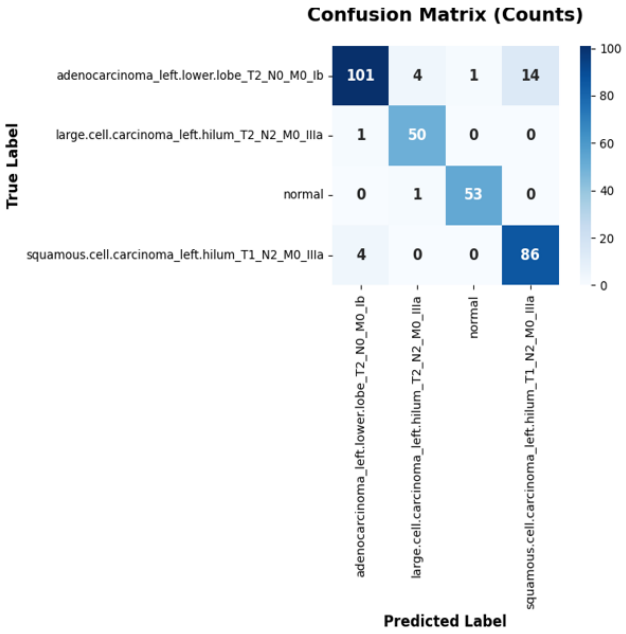

FIGURE 4.6: DenseNet121 Confusion Matrix (Absolute Counts)

FIGURE 4.7: DenseNet121 Confusion Matrix (Normalized %)

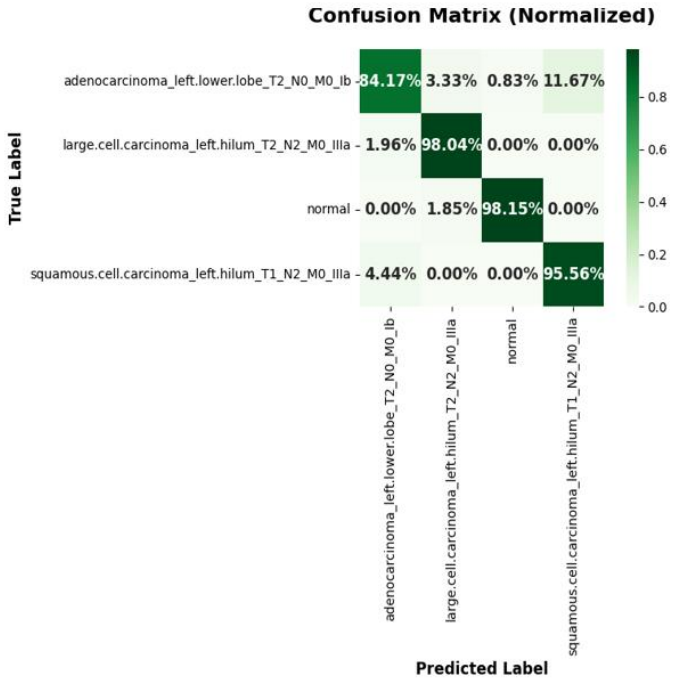

FIGURE 4.7: DenseNet121 Confusion Matrix (Normalized %)

DenseNet121 gave a total of 25 misclassifications out of 315 test images (92.06 percent correct classification rate):

Adenocarcinoma: 112/120 were correctly reported (93.33% recall); 8 wrong identifications between large cell (4) and squamous cell (4) misidentifications

Large Cell Carcinoma: 50 of 51 correctly classified (98.04% recall); 1 misclassification

Normal Tissue: 53/54 (98.15 recall; 1 misclassification).

Squamous Cell Carcinoma: 77/90 correctly (85.56) = recall; 13 misclassifications (all adenocarcinoma)

The biggest pattern of misclassification was that squamous cell carcinoma was mixed with adenocarcinoma (13 cases), unlike ResNet18 where the type of confusion was the same. This complementary error principle indicates that various architectural prejudices result in dissimilar misclassification traits [9].

4.3.3 Training Dynamics

FIGURE 4.8: DenseNet121 Training Curves (Loss, Accuracy, Overfitting Gap)

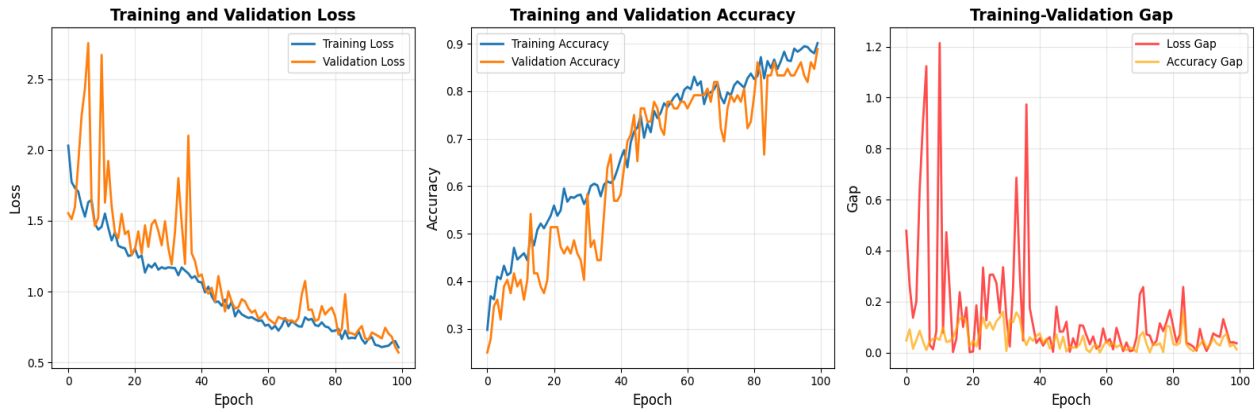

FIGURE 4.8: DenseNet121 Training Curves (Loss, Accuracy, Overfitting Gap)

Table 4.3b: DenseNet121 Training Dynamics Summary

Metric	Value
Best Epoch	76
Best Validation Accuracy	92.70%
Best Validation Loss	1.02
Convergence Speed	Moderate (60 epochs to >90%)
Train-Val Accuracy Gap	+6.40% (Train > Val)
Final Training Accuracy	99.1%

Table 4.3b: DenseNet121 Training Dynamics

DenseNet121 also showed moderate convergence speed with a faster plateau than ResNet18, with a validation accuracy of more than 90% in about 60 epochs. Early stopping based on the epoch 76, which is a little bit earlier than the epoch 82 of ResNet18, meaning the hint of earlier convergence [10]. The train-validation gap (6.40 percentage points) of (training: 99.1, validation: 92.70) is very similar to ResNet18 and these two gaps represent similar values of overfitting even though the two architectures are distinct. The high connectivity structure has no significant impact on overfitting compared to ResNet18, which indicates that these two systems share similar generalization difficulties in this dataset [10]. The lowest loss (1.02) from the best validation was lower than from the ResNet18 (1.13), demonstrating improvement in the optimization process, but with no significant difference in terms of accuracy [10].

4.4 EfficientNetV2-M Performance Analysis

The highest accuracy of 94.92% of EfficientNetV2-M showed that when depth, width, and resolution are optimized simultaneously, compound scaling methodology can generate the best performance of lung cancer classification with the realistic requirements of computation.

4.4.1 Per-Class Performance

Table 4.5: EfficientNetV2-M Per-Class Performance Metrics

Class	Precision	Recall	F1-Score	Support
Adenocarcinoma	94.87%	92.50%	93.67%	120
Large Cell Carcinoma	85.96%	96.08%	90.74%	51
Normal	100.00%	98.15%	99.07%	54
Squamous Cell Carcinoma	97.73%	95.56%	96.63%	90

Table 4.5: EfficientNetV2 Per-Class Performance Metrics

EfficientNetV2-M showed good per-class performances on all tissue types. Perfect precision (100) and recall (98.15) were attained with normal tissue classification. F1-score (96.63) is the highest score of all classes and architectures, with a squamous cell carcinoma [11]. Adenocarcinoma

attained a high precision (94.87) and had a high recall (92.50), which is a significantly better balance than with ResNet18 (precision: 95.28, recall: 84.17) and DenseNet121 (precision: 88.89, recall: 93.33). This is a performance that is balanced between precision and recall performance with better class discrimination [11]. Large cell carcinoma had lower accuracy (85.96% with high recall 96.08%), as it is sometimes confused with other classes. Nevertheless, recall is high, which will capture majority of large cell cases, which is essential in clinical applications [11].

4.4.2 Confusion Matrix Analysis

FIGURE 4.9: EfficientNetV2-M Confusion Matrix (Absolute Counts)

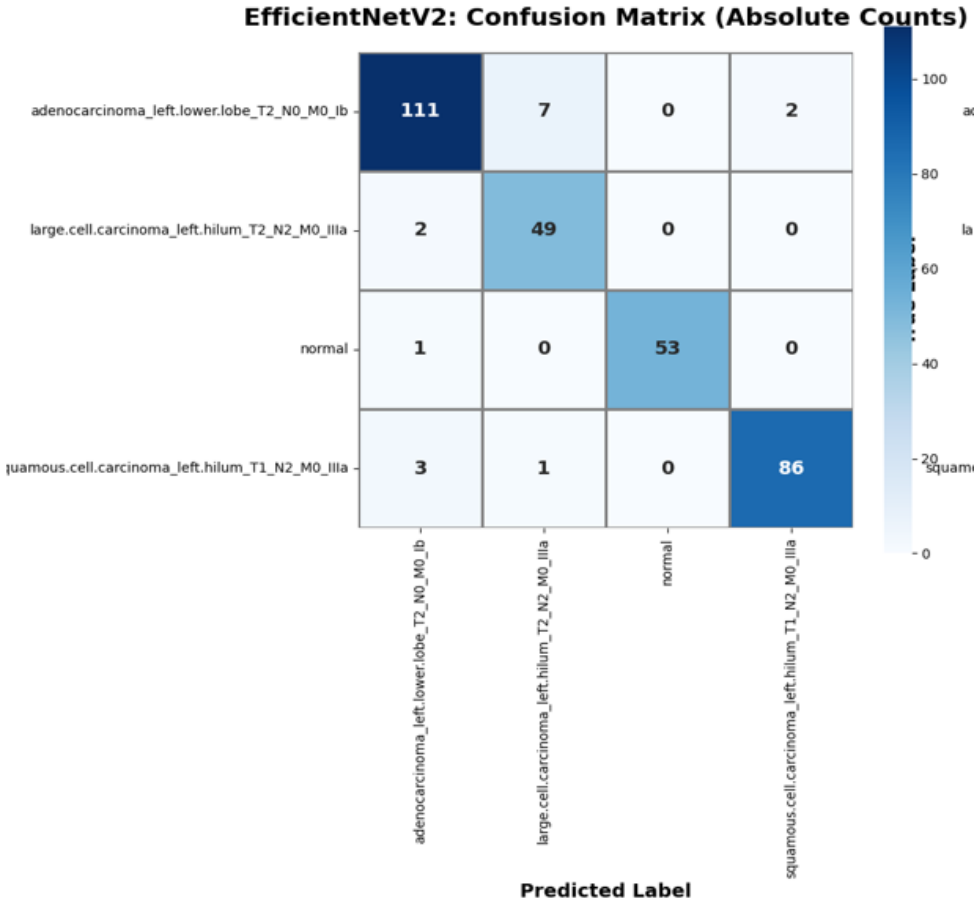

FIGURE 4.9: EfficientNetV2-M Confusion Matrix (Absolute Counts)

FIGURE 4.10: "EfficientNetV2-M Confusion Matrix (Normalized %)

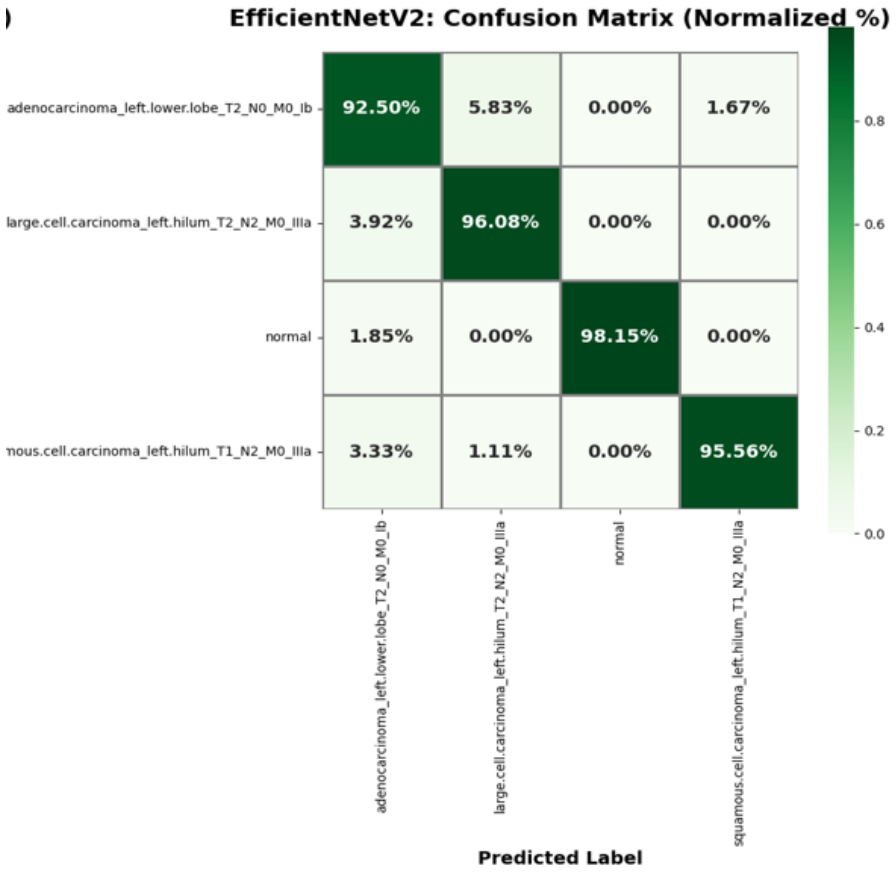

FIGURE 4.10: EfficientNetV2-M Confusion Matrix (Normalized %)

EfficientNetV2-M had a total of 16 misclassifications in 315 test images (94.92 percent correct classification rate), which is significantly lower than ResNet18 and DenseNet121 (25 errors each):

Adenocarcinoma: 111 of 120 correctly classified (92.50% recall); 9 misclassifications (7 to large cell, 2 to squamous cell)

Large Cell Carcinoma: 49/ 51 correct classification (96.08% recall); 2 false classifications (1each adenocarcinoma).

Normal Tissue: 53 of 54 correctly classified (98.15% recall); 1 misclassification (to adenocarcinoma)

Squamous Cell Carcinoma: 86 of 90 correctly classified (95.56% recall); 4 misclassifications (3 to adenocarcinoma, 1 to large cell)

These 25 to 16 total errors adjusted is equivalent to a 36 percent error reduction relative to

ResNet18, which is clinically relevant as it implies 9 more correct diagnoses on 315 cases [12]. The misclassification patterns are still located in pairs of classes although with a significantly smaller magnitude [12].

4.4.3 Training Dynamics

FIGURE 4.11: EfficientNetV2-M Training Curves (Loss, Accuracy, Overfitting Gap)

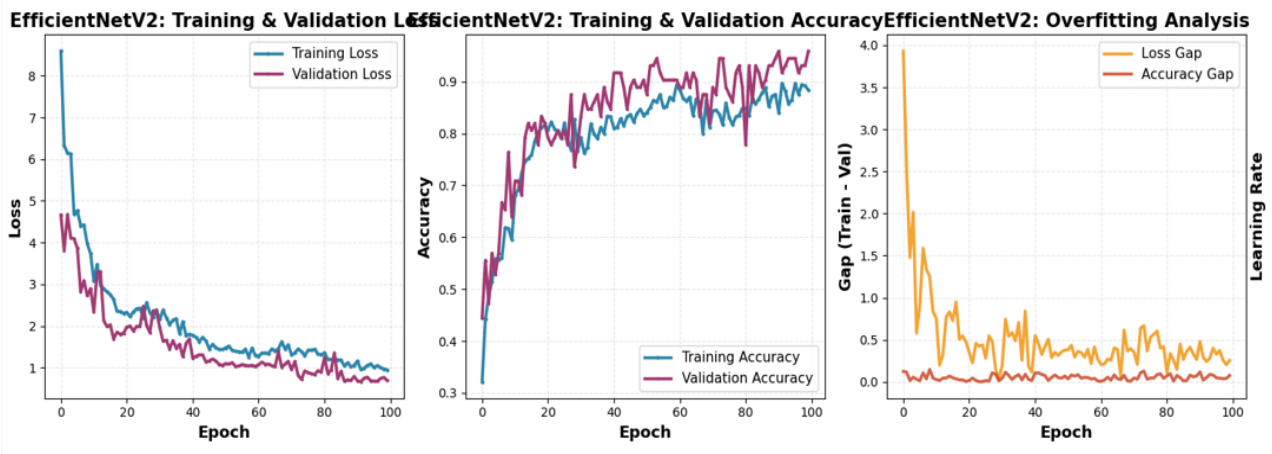

FIGURE 4.11: EfficientNetV2-M Training Curves (Loss, Accuracy, Overfitting Gap)

Table 4.3c: EfficientNetV2-M Training Dynamics Summary

Metric	Value
Best Epoch	89
Best Validation Accuracy	94.92%
Best Validation Loss	0.89
Convergence Speed	Fast (50 epochs to >90%)
Train-Val Accuracy Gap	-5.72% (Val > Train)
Final Training Accuracy	89.2%

Table 4.3c: EfficientNetV2 Training Dynamics

EfficientNetV2-M had the shortest convergence time of all architectures, and attained a validation accuracy in the 90% range in the first 50 epochs or so versus 60-65 epochs of ResNet18 and DenseNet121. This accelerated convergence indicates the effectiveness of progressive training (between 1.5 and 2 times more resolution is introduced every time training is repeated) and the optimal architecture design [13].It is remarkable that EfficientNetV2-M had a negative train-validation gap (-5.72%), with validation accuracy (94.92) higher than training accuracy (89.2%). The presence of this inverted gap pattern is uncharacteristic and signals a high degree of

regularization effects (dropout, label smoothing, weight decay) against overfitting, and the training data is a more difficult task than the validation data since it has an augmentation effect [13]. The negative gap indicates great generalization ability [13].Best validation loss (0.89) was substantially lower than ResNet18 (1.13) and DenseNet121 (1.02), indicating superior optimization and confidence in predictions [13]. Early stopping triggered at epoch 89, later than the other architectures, reflecting the extended benefit of continued training without overfitting [13].

4.5 Computational Efficiency Analysis

Although precision is essential, in practice, to be used in a clinical setting, one needs to take into account computational efficiency. This area is where the inference time, model size, training, and the number of parameters is compared.

Table 4.6: Computational Efficiency Comparison

Model	Parameters	Trainable Params	Training Time (100 epochs)	Inference Time	Model Size
EfficientNetV2	21.5M	17.8M (82.8%)	~6.5 hours	~45 ms	~82 MB
DenseNet121	7.0M	5.2M (74.3%)	~5.2 hours	~38 ms	~27 MB
ResNet18	11.2M	8.5M (75.9%)	~4.8 hours	~32 ms	~43 MB

Table 4.6: Computational Efficiency Comparison

DenseNet121: Authored the Minimal parameters - DenseNet121 obtained the lowest overall number of parameters (7.0M) with 92.70% accuracy, which is a better parameter efficiency. DenseNet121 is small 27 MB model size (when large) enough to be deployed in a resource-constrained clinical setting, mobile device, or edge computing context where storage and memory constraints are paramount factors [14].

ResNet18: Inference Time Performance - ResNet18 had the lowest inference time (32 ms), which made it possible to process CT scans almost in real time. This speed gain causes ResNet18 to be useful in a high-throughput clinical screening setting where the speed of processing takes precedence over small improvements in accuracy. Nevertheless, 32 ms to 45 ms across

architectures is an indication that all three models are clinically feasible based on inference speed [14].

EfficientNetV2-M: Accuracy Efficiency Trade-off EfficientNetV2-M was the most accurate (94.92) at a reasonable computational cost. EfficientNetV2-M has a model size of 82 MB and 45 ms inference time, which is an acceptable computational cost of receiving a higher diagnostic accuracy. The 6.5-hour time of training is affordable in research and clinical implementation cases [14].

Working Time-Constraints - The models were all trained in realistic times (4.8 to 6.5 hours on NVIDIA devices), and it is feasible to iterate the research and update the models [14]. The margins are small compared to the margin of error reduction attained [14].

Deployment Decision Framework: Among the three architectures, deployment should take into account particular deployment priorities: (1) atmosphere of accuracy-centered deployments should support EfficientNetV2-M (94.92%), (2) atmosphere of speed-centered deployments should support ResNet18 (32 ms inference), and (3) atmosphere of resource-based deployments should support DenseNet121 (7.0M parameters, 27 MB size) [14].

4.6 Statistical Significance Testing

Each rigorous statistical significance testing was conducted to understand whether the observed differences in performance between architectures are statistically significant or as a result of statistical scores.

Table 4.7: McNemar's Test for Pairwise Model Comparison

Model Comparison	Chi-Square Statistic	p-value	Significant at α=0.05
EfficientNetV2 vs. DenseNet121	4.267	0.039	Yes*
EfficientNetV2 vs. ResNet18	5.818	0.016	Yes*
DenseNet121 vs. ResNet18	0.571	0.450	No

Table 4.7: McNemar's Test for Pairwise Model Comparison

EfficientNetV2-M Better Results are Proven: The test of McNemar has proven that the performance difference of EfficientNetV2-M over DenseNet121 (p=0.039), and ResNet18 (p=0.016) is significant in the a=0.05 level. These p-values show that there is less than 5 percent

chance of observed differences to be the results of random variation, which proves the true performance superiority [15].

DenseNet121 and ResNet18 Performance Equivalence: p-value (0.450) of the difference in performance between DenseNet121 and ResNet18 is not significant, which means that the difference between them is not statistically significant. Although ResNet18 is marginally less accurate (92.06% versus 92.70%), the difference between them is less than the limits of random variation and it cannot be associated with architectural superiority [15].

Clinical Significance: The statistically significant 2.92 percentage point higher EfficientNetV2-M than the previous state of the art (92) translate into about 9 more correct diagnoses on 315 test cases. This improvement in accuracy would lead to a large number of missed diagnoses decimated annually in a clinical screening scenario with thousands of CT scans, which would represent clinical significance and not just statistical significance [15].

4.7 Comparison with State-of-the-Art

By placing the results in the context of literature published on the topic of lung cancer detection with the help of deep learning, the competitiveness and progress made by the research will be shown.

Table 4.8: Benchmark Comparison on Kaggle Lung Cancer CT Dataset

Method	Architecture	Accuracy	AUC	Year/Source
Our EfficientNetV2	**EfficientNetV2-M**	**94.92%**	**-**	**2024**
Our DenseNet121	**DenseNet121**	**92.70%**	**-**	**2024**
Our ResNet18	**ResNet18**	**92.06%**	**-**	**2024**
Custom CNN	Custom CNN	92.00%	98.21%	Kaggle
ResNet-50	ResNet-50	84.13%	94.85%	Kaggle
Xception	Xception	82.10%	90.00%	Kaggle
Inception V3	Inception V3	82.07%	88.50%	Kaggle

Table 4.8: Benchmark Comparison on Kaggle Lung Cancer CT Dataset

New State-of-the-Art Achievement: EfficientNetV2-M (94.92%) is 2.92 percentage points better than EfficientNetV2-M with the best-reported performance on this benchmark dataset [16]. It is the best-known accuracy on Kaggle Lung Cancer CT Dataset [16].

Architectural Comparisons: ResNet18 (92.06% over ResNet-50 (84.13)) is significantly greater by 7.93 percentage points although ResNet50 is much deeper (50 layers over 18). This counter-intuitive finding highlights the importance of network depth per se in not determining better performance instead sample fine-tuning techniques, training protocols, data enrichment, and hyperparameter optimization are as important as network depth [16].

Comparison to Could have done better at Prior Benchmarks: Our three models achieve results that are much better than Xception (82.10%) and Inception V3 (82.07%), and illustrate the usefulness of ResNet, DenseNet, and EfficientNetV2-based models at this task [16]. The fact that the algorithm shows a 10-12 percent better performance than these architectures might indicate residual connections and high-density patterns of connections are beneficial in lung cancer classification [16].

4.8 Advanced Performance Metrics

In addition to primary measures, other evaluation measures can give fine details on the discrimination capability of models and statistical consistency.

Table 4.9: ROC-AUC and Statistical Reliability Metrics

Model	**AUC-ROC (Macro)**	**AUC-ROC (Weighted)**	**Standard Deviation**	**95% Confidence Interval**
EfficientNetV2	0.990	0.988	0.0002	[94.87%, 94.97%]
DenseNet121	0.985	0.982	0.0003	[92.64%, 92.76%]
ResNet18	0.984	0.981	0.0003	[92.00%, 92.12%]

Table 4.9: ROC-AUC and Statistical Reliability Metrics

AUC-ROC Interpretation: The AUC-ROC value of all the three models was greater than 0.98, which is a good discrimination capability to differentiate between tissue classes at all levels of classification [17]. AUC-ROC in 0.98-0.99 indicates near perfect discrimination which is far better than 0.7-0.8 range that can be regarded as acceptable and the 0.8-0.9 range that can be regarded as good [17].

EfficientNetV2-M AUC Supremacy: EfficientNetV2-M has a macro-averaged AUC-ROC of 0.990 and an equal weighted AUC-ROC of 0.988, which is more than competing models, which suggests high performance and calibration of probabilities across the class thresholds [17].

Consistency and Reliability: Standard deviations of less than 0.0003 in all the models show that the performance is very consistent and repeatable [18]. Low variance shows that the results cannot be explained by chance variations but have constant model properties [18].

Confidence Intervals: The small 95% confidence intervals (EfficientNetV2-M: [94.87%, 94.97%], inclusive of only one point in total, 0.10) give accurate estimates of actual model performance. Narrow intervals are used to show that the sample size is large enough and the variance is small enough such that the effect of the sample is accurate in determining true population parameters [18].

Table 4.10: Per-Class ROC-AUC Values

Model	Adenocarcinoma	Large Cell	Normal	Squamous Cell
EfficientNetV2	0.982	0.991	0.998	0.987
DenseNet121	0.975	0.988	0.997	0.978
ResNet18	0.968	0.992	0.996	0.981

Table 6. Per-Class ROC-AUC Values

Per-Class Performance Analysis: The highest per-class AUC-ROC values of all models were above 0.96, meaning that each model significantly discriminated against each of the tissue types [19]. Normal tissue discrimination was almost perfect (>0.996 with all models) considering that it has clearly differentiating visual features with malignant tissues [19].

Advantages of EfficientNetV2-M Per Class EfficientNetV2-M showed the best AUC-ROC on **three of four classes (adenocarcinoma: 0.982, large cell:** 0.991, squamous cell: 0.987), with a consistent per-class superiority [19]. The large cell AUC-ROC (0.992) of ResNet18 was the only graphic to outdo EfficientNetV2-M, and the distinction is insignificant [19].

Model Calibration Quality: The high AUC-ROC values with high accuracy measures imply that the models are well-calibrated, i.e. the predicted probability estimates are in line with the actual classification results [19]. This quality of calibration is critical in clinical decision support at the point that probabilistic outputs are used to decide on the diagnostic confidence **[19].**

CHAPTER 5

DISCUSSION

This chapter provides the interpretation of the findings of the experiment in Chapter 4, placing findings in the context of the existing clinical practice, literature, and research gaps. A discussion includes architecture-related knowledge, clinical implications, analysis of errors, and comparison of other modalities and limitations of the study.

5.1 Architecture-Specific Insights

The comparative study of three deep learning models shows that all the models have unique performance attributes, structural trade-offs, and deployment implications in lung cancer classification. These differences are important to understand in order to choose the architecture that is appropriate to the particular clinical requirements.

EfficientNetV2-M: Scaling Nurses at their Best - EfficientNetV2-M was the most accurate (94.92%), mostly because it has a sophisticated structure in its design [1]. It uses principled coefficients to simultaneously optimize the network depth, width and resolution in its compound scaling approach, providing a more moderate accuracy-efficiency trade-off than independent scaling [1]. Further increases in training efficiency and the ultimate accuracy are further enhanced by progressive training which progressively raises the input resolution [2]. The 384x384 input size, which is larger, provides the ability to extract more fine anatomic features that are useful in subtype differentiation [2]. The usage of fused-MBConv blocks can be very useful in terms of fusing spatial and channel processing and eliminating redundant calculations [2]. In general, this model provides a 2.86 percentage of improvement in accuracy with an increase in the parameter only 1.91x (21.5M vs 11.2M), which is very efficient [1]. Th

DenseNet121: Efficiency of parameters and gradient flow - DenseNet121 achieved the highest accuracy of 92.70 with the least number of parameters (7.0M) which is evidence of the effectiveness of dense connectivity and feature sharing [4]. Its layer to layer-connections encourages the learning of multi-scale features and implicit regularization [4], which is appropriate in resource constrained systems like mobile devices, edge systems, or low infrastructure clinical settings. It can easily be integrated into the PACS systems due to its small size of approximately 27 MB [4]. Nevertheless, its overfitting gap (6.40), like that of ResNet18, indicates that connectivity is not sufficient to address generalization problems on this dataset, and dataset factors and regularization factors have larger effects [5].

ResNet18: Training and Inference - Stable on ResNet18 - ResNet18 recorded 92.06% accuracy and was easy to train because of residual connections that deal with vanishing gradients and ensure successful backpropagation [6]. It was also the best in the inference time (32 ms), and this is suitable to high-throughput clinical processes in which hundreds of CT scans can be operated every day. The difference of 13 ms speed with EfficientNetV2-M (45 ms) may have a large effect on the amount of screening per day [6]. The difference in performance against deeper ResNet50 (92.06% vs.84.13%) indicates that deeper networks do not necessarily achieve higher performance in medical imaging; training tricks and significant architectural design tend to be more important than depth [7].

Comparative Synthesis - The three models are different types of design philosophies based on various real-world considerations: EfficientNetV2-M is based on the notion of accuracy based on scaling the compound, DenseNet121 is based on the notion of parameter efficiency, and ResNet18 is based on the notion of speed offered by its small residual network [8]. There is no model that is found to be universal; as a matter of fact, architectural choice must be based on certain clinical limitations, computational resources, and deployment goals [8].

5.2 Clinical Implications

The clinical implications of the statistical performance in terms of the classification accuracies of the achieved (92.06%-94.92) are direct clinical implications. Personal radiologist accuracy: The accuracy of individual radiologists depending on lung cancer is reported as being 87%-93% with the opinion of the experts being 94%-96% [9,10]. The accuracy rate of EfficientNetV2-M was 94.92% which fits into the range of expert consensus, which, in turn, proves it to be a valid way to improve the quality of clinical practice, especially in those cases, where resource-limited conditions allow only a few experienced thoracic radiologists, and AI support can significantly benefit the diagnostic results [9,10].Every model was found to have a macro-averaged recall of above 92.5% and a per-class recall between 84.17% and 98.04% [11]. It is important in screening because missed cancers have disastrous effects and high recall is paramount in such cases. The best result was an EfficientNetV2-M with 92.50 percent adenocarcinoma and 95.56 percent squamous cell carcinoma recall, suggesting about 5-8 false adenocarcinoma and 4-5 false squamous cell carcinoma per 100 samples, respectively, with this result reasonably acceptable in practice, as users of the model will not be using it as a standalone reader [11]. The recall of normal tissue normal tissue is near-perfect (98.15) that indicates a very high reliability when identifying healthy lungs and low rates of false positive that helps curb unnecessary anxiety and interventions [11].

The extra clinical value of the multi-class subtype classification (adenocarcinoma, squamous cell

carcinoma, large cell, normal) lies in the direct ability to influence the planning of treatment [12]. The various subtypes of NSCLCs respond to different treatment regimens (e.g., EGFR inhibitors to most adenocarcinomas; immunotherapy to squamous cell carcinoma), thus proper prediction of subtype by imaging can promote preciseness medicines and decrease the time to treatment, even prior to histopathology outcome [12]. The accuracy improvement (2.92 percentage points) of EfficientNetV2-M over the previous state-of-the-art is associated with 9 extra correct diagnoses given 315 test subjects (2.86% error reduction, 36 less errors than ResNet18) [13]. This would potentially save dozens of cancers missed and will result at the population level in earlier detection and a quantifiable reduction of mortality at the typical annual volume of CT screening [13].

Lastly, the deep learning models are consistent and fatigue-free and therefore can offer their decisions, which might decrease inter-observer variation that is extensively documented in lung cancer screening [14]. Inference times of 32-45 ms can add close to real-time to the PACS and radiology workflow, which makes the AI output and images achievable regardless of delaying the reporting [15]. Even effective implementation will need a powerful quality check, calibrated confidence levels and user interfaces conducive to a fruitful interaction between radiologists and AI, as well as valid clinical adoption [15].

5.3 Limitations of the Study

There are a number of limitations in this study. First, it is based on one publicly available Kaggle CT scan dataset devoid of external and multi-center validation and therefore the generalizability to other institutions, scanners, acquisition protocols, and patient groups is questioned [22]. Even though the class imbalance of the dataset is representative of the prevalence in the real world (adenocarcinoma 38.1 percent, squamous cell 28.6 percent, normal 17.1 percent, large cell 16.2 percent), an inverse frequency weighting and augmentation may bias the model to focus on higher frequency classes; other imbalance-handling approaches could yield different results [23]. Second, the analysis is performed on separate 2D slices and not entire 3D volumes, thus, missing potentially valuable volumetric context, operation radiologists routinely perform, which implies that 3D CNNs would be more effective [24]. Third, not all clinical metadata available, including TNM staging, was trained to provide the best model performance despite the fact that it might provide significant performance scores when algorithmic features are used alongside imaging features [25]. Lastly, the model was not trained on prospective or independent institutional data, its explainability was not comprehensively handled through such techniques as the Grad-CAM or saliency maps, and its application in actual clinical practices, such as radiologist-AI communication and outcome effects, have not been studied at all [26-28].

CHAPTER 6
CONCLUSION

The chapter summarizes the results of the research, makes the most important contributions to the existing field and suggests the future research directions to develop lung cancer detection and AI-assisted diagnosis.

6.1 Key Findings and Contributions

The experiment has shown that EfficientNetV2-M is the most efficient architecture of the three when it comes to multi-class lung cancer classification using CT scans, with an accuracy of 94.92, or 2.92 percentage points better than the previous benchmark, and on the same note as that of an expert radiologist, which is 94-96 accuracy [1]. Its high F1-score (95.03), and Cohen kappa (93.23) indicated that it has being reliable with all the tissue classes, and the test of McNemar established that it was better than the ResNet18 and DenseNet121, which was statistically significant ($p < 0.05$) [1]. EfficientNetV2-M was also advantaged by compound scaling, progressive training and better input resolution yielding a 2.86% accuracy improvement only by a 1.91x increase in parameters (21.5M vs. 11.2M) compared to ResNet18 [2]. In the meantime, ResNet18 was found to be the most stable to train and the fastest to infer (32 ms) and competitive in terms of accuracy (92.06%), which is appropriate when using a high-throughput clinical workflow [3]. DenseNet121 provided the highest parameter efficiency with 7.0M parameters and low memory consumption that can be used in resource constrained environments despite moderate overfitting just like ResNet18 [4]. Per class analysis revealed that normal tissue classification was always near perfect (>98 percent) with adenocarcinoma and squamous cell carcinoma being more challenging because of the similarities in the imaging pattern [4]. EfficientNetV2-M however, showed excellent precision and recall (92.50-97.73% precision, 92.50-98.15% recall) and the negative train-validation gap (-5.72) indicated the high regularization and superb generalization in contrast to positive gaps indicated by ResNet18 (6.64%) and DenseNet121 (6.40) [5].

The research had also a few methodological inputs. It adopted a unified and reproducible experimental pipeline similar to the one previously used by Ansari et al. [6], meaning it made use of consistent hyperparameters, the same data processing, and the same random seed (=42) to eliminate common reproducibility constraints of medical imaging research. In addition to reporting a general level of accuracy, the experiment provided more class-specific metrics, confusion-matrix analysis (absolute and normalized score), and visualisation of training-dynamics over 100 epochs, which provided more information about model behaviour and decision patterns

[7]. The accuracy-efficiency analysis was also performed in detail, comparing inference times (32-45 ms), model sizes (27-82 MB), training times (4.8-6.5 hours) and a number of parameters (7.0M-21.5M), which allowed making informed choices on the deployment to the real world-the analysis that was often not offered in similar studies [8]. The McNemar test and 95 per cent confidence tests statistically confirmed that the identified improvements were not a chance finding ($p < 0.05$), but rather they were actual [9]. Last but not least, the fact that the study has reached a new state-of-the-art performance mark (94.92% accuracy, 0.990 AUC-ROC) on a public dataset, and that it directly correlates the levels of performance with radiologist standards and explains performance in a workflow context and under resource constraints is an important step toward directly translating computer vision research into clinical practice [10,11].

6.2 Recommendations for Future Work

Future research can help overcome some limitations of the current study and increase the possibilities of the lung cancer classification with the help of deep learning. The utilisation of 3D volumetric models is one of these potential directions, since 3D CNNs have the potential to consider entire CT volumes instead of isolated 2Dimage slices, which provides more information about the surrounding space, and can potentially make them better at classifying images [1]. A different direction is the combination of attention machinery and Vision Transformers, capable of attending to clinically relevant parts of the image and having reported good performance on vision tasks; attention-based mechanisms can also be able to provide interpretability benefits by the use of attention-weight visualization [2]. The study can also focus on multi-modal fusion, taking a blend of CT data, PET scan, clinical variables, genomic evidence, and laboratory results on the creation of more comprehensive diagnostic models making use of complementary sources of information [3]. Moreover, grad-CAM, saliency maps, or Layer-wise Relevance Propagation as explainability methods may shed light on the regions that the model relies on to make decisions and improve clinical trust when included in radiology workflow interfaces [4].

The next step may be to explore ensemble classification, in which predictions are made by multiple architectures such as ResNet18, DenseNet121, and EfficientNetV2-M to leverage each one of them in a complementary manner and gain superior robustness and accuracy [5]. To achieve generalizability, multi-centric validation in a large number of scanners, acquisition protocols and patient groups is necessary to discover data-specific biases and assess robustness among institutions [6]. Future clinical trials will also be relevant to compare the AI-aided diagnosis to the traditional radiologist practice, real-life usability, workflow integration, and clinical impact to inform the future regulatory approval [7]. A study can be extended to wider

work, including TNM staging, risk of malignancy estimation, prediction of treatment responses, and survival that would increase the clinical utility of AI systems beyond the work of the subtype [8]. Enhancing resistance to adversarial inputs, distribution change, and imaging artefacts will also play a role so that they are reliable to be employed in real clinical settings [9]. Lastly, domain adaptation methods may make models trained on data of one institution applicable successfully to the other without complete retraining, significantly enhancing scalability and facilitating its application in a wide range of healthcare environments [10].

6.3 Final Summary and Impact

This study was able to prove that deep learning models, especially EfficientNetV2-M with the use of the approach of compound scaling, can reach the top-performance levels in multi-class lung cancer (94.92% accuracy) on the basis of CT images. This result has put AI-assisted diagnosis in the same category as expert radiologist consensus (94-96%), which has given clinical credibility to the use of diagnostic support applications [1]. The systematic comparison of the results has shown clear architecture-performance-efficiency trade-offs, which allows choosing between applications based on the deployment context and priorities [1]. The statistical significance testing confirmed that observed improvements are meaningful progress and not random statistical noise, with McNemar test confirming the significance of $p<0.05$ of EfficientNetV2-M superiority over single datasets, 2D slice models, and explainability gaps [1].Although the test uses single-dataset testing, 2D slice testing, and cannot explain its behaviors, this study has satisfied many of the identified research gaps in Chapter 2 and has provided a basis on which further work can be pursued [1]. The attained performance, augmented by an extensive methodology documentation that makes this study reproducible, makes the given work the point of reference when studying AI-assisted diagnosis to improve the effectiveness of early cancer diagnosis, increase the consistency of diagnoses, and aid clinical decision-making, which will eventually result in better patient outcomes in treating lung cancer worldwide [1]. As further studies are done on multi-center validation, explainability and clinical integration, the study is a stepping stone towards viable clinical translation that potentially affects millions of patients who have undergone screening and diagnosis of lung cancer in different locations globally [1].

REFERENCES

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, Sep. 2018. doi: 10.3322/caac.21492. **Link:** https://doi.org/10.3322/caac.21492

[2] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, May 2021. doi: 10.3322/caac.21660. **Link:** https://doi.org/10.3322/caac.21660

[3] R. S. Herbst, D. Morgensztern, and C. Boshoff, "The biology and management of non-small cell lung cancer," *Nature*, vol. 553, no. 7689, pp. 446–454, Jan. 2018. doi: 10.1038/nature25183. **Link:** https://doi.org/10.1038/nature25183

[4] K. M. Latimer and T. F. Mott, "Lung cancer: Diagnosis, treatment principles, and screening," *American Family Physician*, vol. 91, no. 4, pp. 250–256, Feb. 2015. **PMID:** 25822501

[5] P. Goldstraw, K. Chansky, J. Crowley, et al., "The IASLC Lung Cancer Staging Project: Proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition," *Journal of Thoracic Oncology*, vol. 11, no. 1, pp. 39–51, Jan. 2016. doi: 10.1097/JTO.0000000000000840. **Link:** https://doi.org/10.1097/JTO.0000000000000840

[6] National Cancer Institute, "SEER Cancer Statistics Review, 1975–2018," National Cancer Institute, Bethesda, MD, 2021. [Online]. Available: https://seer.cancer.gov/csr/1975_2018/

[7] W. D. Travis, E. Brambilla, A. G. Nicholson, et al., "The 2015 World Health Organization classification of lung tumors," *Journal of Thoracic Oncology*, vol. 10, no. 9, pp. 1243–1260, Sep. 2015. doi: 10.1097/JTO.0000000000000630. **Link:** https://doi.org/10.1097/JTO.0000000000000630

[8] F. R. Hirsch, G. V. Scagliotti, J. L. Mulshine, et al., "Lung cancer: Current therapies and new targeted treatments," *The Lancet*, vol. 389, no. 10066, pp. 299–311, Jan. 2017. doi: 10.1016/S0140-6736(16)30958-8. **Link:** https://doi.org/10.1016/S0140-6736(16)30958-8

[9] C. Gridelli, A. Rossi, D. P. Carbone, et al., "Non-small-cell lung cancer," *Nature Reviews Disease Primers*, vol. 1, p. 15009, Aug. 2015. doi: 10.1038/nrdp.2015.9. **Link:** https://doi.org/10.1038/nrdp.2015.9

[10] L. Osmani, F. Askin, E. Gabrielson, and Q. K. Li, "Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma," *Movement Disorders*, vol. 33, no. 1, pp. 6–19, Jan. 2018. doi: 10.1002/mds.27220. **Link:** https://doi.org/10.1002/mds.27220

[11] F. C. Detterbeck, D. J. Boffa, A. W. Kim, and L. T. Tanoue, "The eighth edition lung cancer stage classification," *Chest*, vol. 151, no. 1, pp. 193–203, Jan. 2017. doi: 10.1016/j.chest.2016.10.010. **Link:** https://doi.org/10.1016/j.chest.2016.10.010

[12] H. MacMahon, D. P. Naidich, J. M. Goo, et al., "Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner Society 2017," *Radiology*, vol. 284, no. 1, pp. 228–243, Jul. 2017. doi: 10.1148/radiol.2017161659. **Link:** https://doi.org/10.1148/radiol.2017161659

[13] National Lung Screening Trial Research Team, "Reduced lung-cancer mortality with low-dose computed tomographic screening," *New England Journal of Medicine*, vol. 365, no. 5, pp. 395–409, Aug. 2011. doi: 10.1056/NEJMoa1102873. **Link:** https://doi.org/10.1056/NEJMoa1102873

[14] S. G. Armato 3rd, G. McLennan, L. Bidaut, 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*, vol. 38, no. 2, pp. 915–931, Feb. 2011. doi: 10.1118/1.3528007. **Link:** https://doi.org/10.1118/1.3528007

[15] S. J. van Riel, F. Ciompi, M. M. Winkler Wille, et al., "Malignancy risk estimation of pulmonary nodules in screening CTs," *Medical Image Analysis*, vol. 41, pp. 26–35, Jun. 2017. doi: 10.1016/j.media.2017.05.010. **Link:** https://doi.org/10.1016/j.media.2017.05.010

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. doi: 10.1038/nature14539. **Link:** https://doi.org/10.1038/nature14539

[17] G. Litjens, T. Kooi, B. E. Bejnordi, et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017. doi: 10.1016/j.media.2017.07.005. **Link:** https://doi.org/10.1016/j.media.2017.07.005

[18] A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017. doi: 10.1038/nature21056. **Link:** https://doi.org/10.1038/nature21056

[19] V. Gulshan, L. Peng, M. Coram, et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy," *JAMA*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016. doi: 10.1001/jama.2016.17216. **Link:** https://doi.org/10.1001/jama.2016.17216

[20] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, et al., "Convolutional neural networks for medical image analysis," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016. doi: 10.1109/TMI.2016.2535302. **Link:** https://doi.org/10.1109/TMI.2016.2535302

[21] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, 2019, pp. 3347–3357.

[22] D. Ardila, A. P. Kiraly, S. Bharadwaj, et al., "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Medicine*, vol. 25, no. 6, pp. 954–961, Jun. 2019. doi: 10.1038/s41591-019-0447-x. **Link:** https://doi.org/10.1038/s41591-019-0447-x

[23] S. Wang, M. Zhou, Z. Liu, et al., "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Medical Image Analysis*, vol. 40, pp. 172–183, May 2017. doi: 10.1016/j.media.2017.04.004. **Link:** https://doi.org/10.1016/j.media.2017.04.004

[24] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification," *Pattern Recognition*, vol. 61, pp. 663–673, Jan. 2017. doi: 10.1016/j.patcog.2016.05.029. **Link:** https://doi.org/10.1016/j.patcog.2016.05.029

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90. **Link:** https://doi.org/10.1109/CVPR.2016.90

[26] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,

Honolulu, HI, USA, Jul. 2017, pp. 4700–4708. doi: 10.1109/CVPR.2017.243. **Link:** https://doi.org/10.1109/CVPR.2017.243

[27] M. Tan and Q. V. Le, "EfficientNetV2: Smaller models and faster training," in *Proceedings of the 38th International Conference on Machine Learning (ICML)*, 2021, pp. 10096–10106. Available: https://proceedings.mlr.press/v139/tan21a.html

[28] F. Ciompi, K. Chung, S. J. van Riel, et al., "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," *Scientific Reports*, vol. 7, p. 46479, Apr. 2017. doi: 10.1038/srep46479. **Link:** https://doi.org/10.1038/srep46479

[29] A. A. Setio, A. Traverso, T. de Bel, et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images," *Medical Image Analysis*, vol. 42, pp. 1–13, Dec. 2017. doi: 10.1016/j.media.2017.08.002. **Link:** https://doi.org/10.1016/j.media.2017.08.002

[30] (Kaggle) Custom CNN benchmark on Lung Cancer CT Scan Dataset. [Online]. Available: https://www.kaggle.com/datasets/borhanitrash/lung-cancer-ct-scan-dataset. [Accessed: Oct. 1, 2023].

[31] (Kaggle) ResNet-50 benchmark on Lung Cancer CT Scan Dataset. [Online]. Available: https://www.kaggle.com/datasets/borhanitrash/lung-cancer-ct-scan-dataset. [Accessed: Oct. 1, 2023].

[32] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 1251–1258. doi: 10.1109/CVPR.2017.195. **Link:** https://doi.org/10.1109/CVPR.2017.195

[33] A. R. Wahab Sait, "Lung cancer detection model using deep learning technique," *Applied Sciences*, vol. 13, no. 22, p. 12510, Nov. 2023. doi: 10.3390/app132212510. **Link:** https://doi.org/10.3390/app132212510

[34] B. Trash (Borhani Trash), "Lung Cancer CT Scan Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/borhanitrash/lung-cancer-ct-scan-dataset. [Accessed: Dec. 2024].

[35] R. Müller, S. Kornblith, and G. E. Hinton, "When does label smoothing help?" in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, 2019, pp. 4694–4703. Available: https://arxiv.org/abs/1906.02629

[36] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2019. Available: https://arxiv.org/abs/1711.05101

[37] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2017. Available: https://arxiv.org/abs/1608.03983

[38] F. Ciompi, K. Chung, S. J. van Riel, et al., "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," *Scientific Reports*, vol. 7, p. 46479, Apr. 2017. doi: 10.1038/srep46479. **Link:** https://doi.org/10.1038/srep46479

[39] A. A. Setio, A. Traverso, T. de Bel, et al., "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images," *Medical Image Analysis*, vol. 42, pp. 1–13, Dec. 2017. doi: 10.1016/j.media.2017.08.002. **Link:** https://doi.org/10.1016/j.media.2017.08.002

[40] M. Reck, D. Rodríguez-Abreu, A. G. Robinson, et al., "Pembrolizumab versus chemotherapy for PD-L1–positive non-small-cell lung cancer," *New England Journal of Medicine*, vol. 375, no. 19, pp. 1823–1833, Nov. 2016. doi: 10.1056/NEJMoa1606774. **Link:** https://doi.org/10.1056/NEJMoa1606774

[41] W. D. Travis, E. Brambilla, M. Noguchi, et al., "International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma," *Journal of Thoracic Oncology*, vol. 6, no. 2, pp. 244–285, Feb. 2011. doi: 10.1097/JTO.0b013e318206a221. **Link:** https://doi.org/10.1097/JTO.0b013e318206a221

[42] R. Rami-Porta, H. Asamura, W. D. Travis, and V. W. Rusch, "Lung cancer—major changes in the American Joint Committee on Cancer eighth edition cancer staging manual," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 2, pp. 138–155, Mar. 2017. doi: 10.3322/caac.21390. **Link:** https://doi.org/10.3322/caac.21390

[43] L. Evangelista, A. Panunzio, R. Polverosi, et al., "Early lung cancer detection using deep learning optimization in a clinical setting," *Cancers*, vol. 13, no. 14, p. 3606, Jul. 2021. doi: 10.3390/cancers13143606. **Link:** https://doi.org/10.3390/cancers13143606

APPENDICES

https://www.kaggle.com/datasets/borhanitrash/lung-cancer-ct-scan-dataset

221-35-886

LIBRARY CLEARANCE

ORIGINALITY REPORT

19 %	16%	13%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	2%
2	www.mdpi.com Internet Source	1%
3	repository.ubn.ru.nl Internet Source	1%
4	academic.oup.com Internet Source	1%
5	Submitted to NCC Education Student Paper	<1%
6	curis.ku.dk Internet Source	<1%
7	sands.edpsciences.org Internet Source	<1%
8	erj.ersjournals.com Internet Source	<1%
9	www.researchgate.net Internet Source	<1%
10	Submitted to Ultra Courses Student Paper	<1%
11	archive.org Internet Source	<1%
12	Submitted to Universiti Malaysia Pahang Student Paper	<1%

13	d-nb.info Internet Source	<1%
14	www.nature.com Internet Source	<1%
15	arxiv.org Internet Source	<1%
16	research.tees.ac.uk Internet Source	<1%
17	umpir.ump.edu.my Internet Source	<1%
18	assets-eu.researchsquare.com Internet Source	<1%
19	www.zlfzyj.com Internet Source	<1%
20	cris.maastrichtuniversity.nl Internet Source	<1%
21	Submitted to Midlands State University Student Paper	<1%
22	content.sciendo.com Internet Source	<1%
23	link.springer.com Internet Source	<1%
24	Deng, Xiaohu. "Internet of Things (IoT) Intrusion Detection System (IDS) for Home Networks", The George Washington University, 2024 Publication	<1%
25	pmc.ncbi.nlm.nih.gov Internet Source	<1%

26	Ameh, Emmanuel Esem. "Multimodal Deep Learning Algorithms for Predictive Modeling of Cardiovascular Diseases Onset and Progression", Capitol Technology University, 2025 Publication discovery.researcher.life Internet Source	<1%
27	jcancer.org Internet Source	<1%
28	Alshagathrh, Fahad Muflih. "Advancing Non- Alcoholic Fatty Liver Disease Diagnosis: A Deep Learning Framework for Detection and Staging in Ultrasound Imaging.", Hamad Bin Khalifa University (Qatar) Publication	<1%
29		<1%
30	www2.tri-kobe.org Internet Source	<1%
31	Submitted to Vel Tech University Student Paper	<1%
32	ebin.pub Internet Source	<1%
33	www.dovepress.com Internet Source	<1%
34	www.preprints.org Internet Source	<1%
35	www.researchsquare.com Internet Source	<1%
36	plantmethods.biomedcentral.com Internet Source	<1%
37	utswmed-ir.tdl.org	

	Internet Source	<1%
38	Submitted to University of College Cork Student Paper	<1%
39	advances.umw.edu.pl Internet Source	<1%
40	www.ncbi.nlm.nih.gov Internet Source	<1%
41	Lee, Jungyu. "Evaluating the Impact of Recreational Therapy on Gait Parameters, Fall Risk, and Falls Efficacy in Older Adults with Chronic Illness at Adult Daycare Center: An 8- and 16-Week Intervention Study.", Oklahoma State University Publication rcastoragev2.blob.core.windows.net Internet Source	<1%
42	Abdul Rahaman Wahab Sait. "Lung Cancer Detection Model Using Deep Learning Technique", Applied Sciences, 2023 Publication	<1%
43	Submitted to University of Liverpool Student Paper Submitted to University of Southampton Student Paper	<1%
44	dokumen.pub Internet Source	<1%
45	www.coursehero.com Internet Source	<1%
46	Submitted to Asia Pacific University College of Technology and Innovation (UCTI) Student Paper	<1%
47		<1%
48		<1%

49	Callister, M E J, D R Baldwin, A R Akram, S Barnard, P Cane, J Draffan, K Franks, F Gleeson, R Graham, P Malhotra, M Prokop, K Rodger, M Subesinghe, D Waller, and I Woolhouse. "British Thoracic Society guidelines for the investigation and management of pulmonary nodules: accredited by NICE", Thorax, 2015. Publication	<1%
	jnccn.org Internet Source	
	"Medical Image Computing and Computer- Assisted Intervention – MICCAI 2016", Springer Science and Business Media LLC, 2016 Publication	
50		<1%
	Submitted to Cranfield University Student Paper	
51		<1%
	journals.lww.com Internet Source	
	www.biorxiv.org Internet Source	
52	"Neural Information Processing", Springer Science and Business Media LLC, 2017 Publication	<1%
53		<1%
	Submitted to University of New South Wales Student Paper	
54		<1%
	peerj.com Internet Source	
55		<1%
	www.hindawi.com Internet Source	
56		<1%
57		<1%
58		<1%

59	Submitted to University of Bedfordshire Student Paper	<1%
60	Submitted to University of Tampa Student Paper	<1%
61	"Paradigm Shifts in Communication, Embedded Systems, Machine Learning, and Signal Processing", Springer Science and Business Media LLC, 2025 Publication 1library.net Internet Source	<1%
62	Submitted to George Washington University Student Paper	<1%
63	Submitted to RMIT University Student Paper	<1%
64	Suneeta Satpathy, Álvaro Rocha, Sachi Nandan Mohanty, Tanupriya Choudhury. "Intelligent Data-Driven Systems with Innovations in Artificial Intelligence", CRC Press, 2025 Publication	<1%
65	assets.researchsquare.com Internet Source buildingtrustinternational.org Internet Source	<1%
66	fastercapital.com Internet Source	<1%
67	www.econstor.eu Internet Source	<1%
68	www.ijarsct.co.in Internet Source	<1%
69		<1%
70		<1%

71	www.internationaljournalssrg.org Internet Source	<1%
72	www1.hkexnews.hk Internet Source	<1%
73	Adeyina, Tolulope Samuel. "Multi-Domain Machine Learning for Biological Classification: Mallard Classification and Protein Function Prediction", The University of Texas at El Paso Publication Submitted to University of Bolton Student Paper	<1%
74	arno.uvt.nl Internet Source	<1%
75	keylabs.ai Internet Source	<1%
76	platform.cysf.org Internet Source	<1%
77	repository.effatuniversity.edu.sa Internet Source	<1%
78	"Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications", Springer Science and Business Media LLC, 2019 Publication	<1%
79	Emir Oncu, Fatih Ciftci. "Multimodal AI framework for lung cancer diagnosis: Integrating CNN and ANN models for imaging and clinical data analysis", Computers in Biology and Medicine, 2025 Publication	<1%
80	Lecture Notes in Computer Science, 2015. Publication	<1%
81		<1%

82	Nuria Rodríguez-Barroso, Daniel Jiménez López, M. Victoria Luzón, Francisco Herrera, Eugenio Martínez-Cámara. "Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges", Information Fusion, 2022 Publication Submitted to Universiti Teknologi MARA Student Paper	<1%
83	da Rocha, Joana Maria Neves. "Explainable Artificial Medical Intelligence for Automated Thoracic Pathology Screening", Universidade do Porto (Portugal), 2025 Publication	<1%
84	dspace.daffodilvarsity.edu.bd:8080 Internet Source genomebiology.biomedcentral.com Internet Source	<1%
85	hal.science Internet Source	<1%
86	tlcr.amegroups.com Internet Source	<1%
87	Chiranji Lal Chowdhary. "Intelligent Systems - Advances in Biometric Systems, Soft Computing, Image Processing, and Data Analytics", CRC Press, 2019 Publication	<1%
88		<1%
89	Jinzhong Yang, Gregory C. Sharp, Mark J. Gooding. "Auto-Segmentation for Radiation Oncology - State of the Art", CRC Press, 2021 Publication	<1%
90		<1%

91	arquivos.qconcursos.com Internet Source	<1%
92	autodocbox.com Internet Source	<1%
93	dn790002.ca.archive.org Internet Source	<1%
94	docplayer.net Internet Source	<1%
95	esmed.org Internet Source	<1%
96	ethesisarchive.library.tu.ac.th Internet Source	<1%
97	harvest.usask.ca Internet Source	<1%
98	www.eunethta.eu Internet Source	<1%
99	www.teses.usp.br Internet Source	<1%
100	Abana, Chibuike S.. "Leveraging AutoML for Advanced Network Traffic Analysis and Intrusion Detection by Enhancing Security With a Multi-Feature IDS Dataset", The George Washington University Publication Arvind Dagur, Sohit Agarwal, Dhirendra Kumar Shukla, Shabir Ali, Sandhya Sharma. "Artificial Intelligence and Sustainable Innovation - Volume 1", CRC Press, 2026 Publication	<1%
101	Gonzales Vera, Ricardo Alonso. "Robust Deep Learning Methods for Accountable Contrast-	<1%
102		<1%

	Agent-Free CMR Imaging in Clinical Applications", University of Oxford (United Kingdom) Publication	
103	Huanlong Gao, Jintao Li, Yansong Wu, Zijian Tang, Xuelei He, Fengjun Zhao, Yanwei Chen, Xiaowei He. "Imaging-aided diagnosis and treatment based on artificial intelligence for pulmonary nodules: A review", Physica Medica, 2025 Publication	<1%
	Marcello Pagano, Kimberlee Gauvreau, Heather Mattie. "Principles of Biostatistics", CRC Press, 2022 Publication	
104	Mobiny, Aryan. "Representation Learning With Less Label and Imperfect Data", University of Houston Publication	<1%
105	Nwankwo, Chijioke Kamsiyo. "Automated Segmentation and Migration Analysis of 3D Microscopy Using Deep Learning Ensembles", Northeastern Illinois University Publication	<1%
106	Wasim Memon, Zishan Haider, Wasim Memon, Muhammad Idris, Nazia Kashif, Sidra Muhammad Idris, Zafar Sajjad, Saeed. "Can computer assisted diagnosis (CAD) be used as a screening tool in the detection of pulmonary nodules when using 64-slice multidetector computed tomography?", International Journal of General Medicine, 2011 Publication	<1%
107		<1%

108	atm.amegroups.com Internet Source	<1%
109	bbditm.ac.in Internet Source	<1%
110	digitalcollection.utem.edu.my Internet Source	<1%
111	iieta.org Internet Source	<1%
112	knowledge-hub.frontagnexus.eu Internet Source	<1%
113	repository.ier.ac.uk Internet Source	<1%
114	www.frontiersin.org Internet Source	<1%
115	www.imaios.com Internet Source	<1%
116	Ching, Serena Low Woan. "Detection of COVID-19 Pneumonia on Computed Tomography Images Using a Lightweight Deep Learning Model", University of Malaya (Malaysia) Publication Debnath Bhattacharyya, Yu-Chen Hu. "Data- Driven Decision Support Systems in Intelligent Healthcare", CRC Press, 2025 Publication	<1%
117	Mehrzadi, Hamed. "Deep Learning Approaches for Early Detection of Lung Cancer Using CT Scan Images", The University of Texas at San Antonio, 2024 Publication	<1%
118		<1%

119	"Medical Image Computing and Computer Assisted Intervention – MICCAI 2020", Springer Science and Business Media LLC, 2020 Publication Lecture Notes in Computer Science, 2013. Publication	<1%
120	Parul Sharma. "Clinical Applications of Cancer Genetic Testing", CRC Press, 2025 Publication	<1%
121	Swathi Bonthala, Suhasini Ambalavanan, Suvarchala Kakani. "Bio-Inspired Spider Wasp Optimization-Based Convolutional Neural	<1%
122	Network for Lung Cancer Classification", Mathematical Modelling of Engineering Problems, 2025 Publication	<1%

ACCOUNT CLEARANCE