



A Hybrid Deep Learning Framework for Improved Breast Cancer Classification

Supervised By

Mr. Suprove Chandra Sarkar

Lecturer

Department of Software Engineering

Daffodil International University

Submitted By

Asfia Haque Arpi

ID:221-35-972

Department of Software Engineering

Daffodil International University


This thesis report has been submitted in fulfilment of the requirements for the Degree of Bachelor of Science in Software Engineering.

© All right Reserved by Daffodil International

APPROVAL


This thesis titled on “A Hybrid Deep Learning Framework for improve breast cancer classification”, submitted by **Asfia Haque Arpi (ID: 221-35-972)** 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



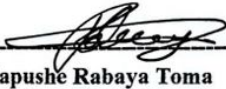
Dr. S M Hasan Mahmud
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Chairman



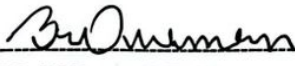
A.H.M Shahariar Parvez
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1




Tapushe Rabaya Toma
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2



Khalid Been md. Badruzzaman Biplob
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 3



Dr. Md Sazzadur Rahman
Professor
Institute of Information technology
Jahangirnagar University, Bangladesh

External Examiner



SUPERVISOR'S DECLARATION

I hereby declare that I have reviewed this thesis entitled " **A Hybrid Deep Learning Framework for Improved Breast Cancer Classification**", and in my opinion, it is adequate in terms of scope and quality for the award of the degree of Bachelor of Science in Software Engineering..

A handwritten signature in black ink, appearing to read "Suprove", is written above a horizontal line.

(Supervisor's Signature)

Full Name : Mr. Suprove Chandra Sarkar

Position : Lecturer

Date : 27/11/202



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.

A handwritten signature in black ink, appearing to read "Asfia", is written over a faint, light-colored rectangular stamp or watermark.

(Student's Signature)

Full Name : Asfia Haque Arpi

ID Number : 221-35-972

Date : 27/11/2025

A Hybrid Deep Learning Framework for Improved Breast Cancer Classification

Asfia Haque Arpi
221-35-972

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

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

December 2025

ACKNOWLEDGEMENTS

To begin with, I would like to thank Almighty Allah who gave me the required strength, patience and wisdom to be able to complete this research work successfully. I hereby wish to heartedly express my gratitude to my supervisor, Mr. Suprove Chandra Sarkar Lecturer, Department of Software engineering, Daffodil international university, who had greatly guided, constantly supported and provided constructive feedback to me during the research process. His support and skills were of great help in ensuring this study was accomplished successfully. My parents have also been a great source of inspiration and strength in my academic life as I am deeply grateful to them because of their unconditional love, sacrifices, and motivation.

Dedication

This thesis is dedicated to my beloved parents, whose unconditional love, sacrifices, and endless support have been the cornerstone of my academic and personal growth. I also submit this work to my grandmother who had strong fought breast cancer; her resilience, greatness, and endurance was an intensive source of motivation and inspiration behind this research. This is one of the works that can be said to have impacted my path and my dedication to make a significant contribution towards breast cancer research.

ABSTRACT

It is a grave health problem on a global scale and diagnosis of breast cancer should be evident at the initial level, which will enhance results. This article creates effective deep learning systems on the BreakHis dataset of binary classification of breast histopathological cultures. To select the best model in benign, MobileNet, VGG, Inception and ResNet will be used and trained and tested several CNN models including EfficientNet to classify the training benign tissue, malignant tissue and classify benchmark tissue. Both the base models and fine-tuned models are investigated to establish the impacts of unfreezing of the weight and changing features on performance. We also prepare hybrid/ensemble models including EfficientNetB0.+ MobileNetV2 to use complementary features of the two networks to improve the accuracy of classification.

The workflow involves preprocessing the data, training and building of the model and the model performance in terms of accuracy, precision, recall, F1 score and loss. We examine energy usage and processing overhead to determine whether they can be implemented in low-resource settings or in the clinical context. Findings indicate that fine-tuning could be used to attain a significant performance enhancement and the proposed hybrid model leads to a greater accuracy at the affordable cost. It is also estimated by the study that increasing the magnification of images resulted in the model generalizing more. Taking into consideration these weaknesses, we observe that the collective deep learning technique on breast cancer is required that is effective to design networks and to be integrated in future both in the clinic and in multi-modal systems.

Table of Contents

| | |
|--|-------------|
| ACKNOWLEDGEMENTS | VI |
| DEDICATION | VII |
| ABSTRACT | VIII |
| LIST OF TABLES | XII |
| LIST OF FIGURES | XIII |
| CHAPTER 1 INTRODUCTION | 1 |
| 1.1 Background | 1 |
| 1.2 Problem Statement | 2 |
| 1.3 Motivation | 3 |
| 1.4 Research Questions | 4 |
| 1.5 Research Objectives | 4 |
| 1.5.1 Research Scope | 4 |
| CHAPTER 2 | 6 |
| LITERATURE REVIEW | 6 |
| 2.1 Deep Learning for Breast Cancer Detection. | 6 |
| 2.2 CNN Architectures. | 6 |
| 2.3 Hybrid and Ensemble Models. | 6 |
| 2.4 Challenges and Research Gaps. | 7 |

| | | |
|--------------------|---|-----------|
| 2.5 | Relevance to Current Research. | 7 |
| CHAPTER 3 | | 8 |
| METHODOLOGY | | 8 |
| 3.1 | Introduction | 8 |
| 3.2 | Dataset Description | 9 |
| 3.2.1 | Key Features of the Dataset | 9 |
| 3.2.2 | Data preprocessing | 10 |
| 3.3 | Hybrid Model Development | 10 |
| 3.4 | Computational Model Development | 11 |
| 3.5 | Workflow Overview | 12 |
| 3.5.1 | Dataset Collection | 12 |
| 3.5.2 | Data Preprocessing | 12 |
| 3.5.3 | Dataset Splitting | 13 |
| 3.5.4 | Hybrid Model Development | 13 |
| 3.5.5 | Model Training | 13 |
| 3.5.6 | Model Evaluation | 14 |
| 3.5.7 | Result Analysis | 14 |
| 3.6 | Model Architecture Components | 14 |
| 3.7 | Model Evaluation | 16 |
| 3.8 | The following evaluation metrics were used: | 16 |
| 3.9 | Confusion Matrix and Visualization | 17 |
| 3.10 | Validation Strategy | 17 |
| CHAPTER 4 | | 18 |

| | |
|--|-----------|
| RESULTS AND DISCUSSION | 18 |
| 4.1 Classification Performance Overview | 18 |
| 4.1.1 Without Fine-Tuning: | 18 |
| 4.1.2 With Fine-Tuning: | 18 |
| 4.1.3 Hybrid/Ensemble Models: | 18 |
| 4.2 Observations: | 19 |
| 4.3 Detailed Evaluation of Selected Models | 19 |
| 4.3.1 Evaluation of EfficientNetB0 (Before and After Fine-Tuning) | 19 |
| 4.3.2 Evaluation of MobileNetV2 (Before and After Fine-Tuning) | 22 |
| 4.3.3 Evaluation of Hybrid Model: EfficientNetB0 + MobileNetV2 (Fine-Tuned) | 26 |
| 4.3.4 Evaluation of Hybrid Model: EfficientNetB0 + ResNet50 (Fine- Tuned) | 28 |
| 4.4 Discussion | 30 |
| | |
| CHAPTER 5 | 33 |
| | |
| CONCLUSION | 33 |
| 5.1 Limitations | 33 |
| 5.2 Future Work | 34 |
| 5.3 Summary of Study | 35 |
| 5.4 References | 36 |

List of tables

Table 5.4-1 Table of Comparison among different model

32

List Of Figures

| | |
|---|----|
| Figure 3.1:1 Proposed Methodology | 9 |
| Figure 4.3:1 Training vs Validation(Accuracy, loss) -EfficientNetB0(Befor Fine-tuning) | 20 |
| Figure 4.3:2 Training vs Validation Accuracy – EfficientNetB0 | 20 |
| Figure 4.3:3 Training vs Validation Accuracy – EfficientNetB0 | 20 |
| Figure 4.3:4 Training vs Validation (Precision, Recall) – EfficientNetB0 (Before Fine tuning) | 21 |
| Figure 4.3:5 Training vs Validation Recall – EfficientNetB0 (Before Fine tuning) | |
| Figure 4.3:6 Training vs Validation Precision – EfficientNetB0 | 21 |
| Figure 4.3:7 Training vs Validation F1-score – EfficientNetB0 | 21 |
| Figure 4.3:8 Training vs Validation F1-score – EfficientNetB0 | 22 |
| Figure 4.3:9 Training vs Validation F1-score – EfficientNetB0(Fine-Tuned) | 22 |
| Figure 4.3:10 Confusion Matrix – EfficientNetB0(Fine-tuned) | 22 |
| Figure 4.3:11 Confusion Matrix – EfficientNetB0 | 22 |
| Figure 4.3:12 Training vs Validation Accuracy – MobileNetV2 | 23 |
| Figure 4.3:13 Training vs Validation Loss – MobileNetV2 | 23 |
| Figure 4.3:14 Training vs Validation Loss – MobileNetV2 | 23 |
| Figure 4.3:15 Training vs Validation Loss – MobileNetV2 (Fine-tuned) | 23 |
| Figure 4.3:16 Training vs Validation Loss – MobileNetV2 | 24 |
| Figure 4.3:17 Training vs Validation Loss – MobileNetV2 | 24 |
| Figure 4.3:18 Recall – MobileNetV2 (Fine-Tuned) | 24 |
| Figure 4.3:19 Recall – MobileNetV2 | 24 |
| Figure 4.3:20 F1-score Curves – MobileNetV2 | 24 |
| Figure 4.3:21 F1-score Curves – MobileNetV2(Fine-tuned) | 24 |
| Figure 4.3:22 F1-score Curves – MobileNetV2 | 25 |
| Figure 4.3:23 F1-score Curves – MobileNetV2(Fine-tuned) | 25 |
| Figure 4.3:24 F1-score Curves – MobileNetV2 | 25 |
| Figure 4.3:25 F1-score Curves – MobileNetV2(Fine-tuned) | 25 |
| Figure 4.3:26 Training vs Validation Accuracy – EfficientNetB0 + MobileNetV2 | |
| Figure 4.3:27 Training vs Validation Accuracy – EfficientNetB0 + MobileNetV2(Fine-tuned) | 26 |

| | |
|---|----|
| Figure 4.3:28 Training vs Validation Loss – EfficientNetB0 + MobileNetV2 | |
| Figure 4.3:29 Training vs Validation Loss – EfficientNetB0 + MobileNetV2 (Fine-tuned) | 26 |
| Figure 4.3:30 Training vs Validation Loss – EfficientNetB0 + MobileNetV2 | |
| Figure 4.3:31 Training vs Validation Precision – Hybrid Model 1(Fine-Tuned) | 27 |
| Figure 4.3:32 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:33 Training vs Validation Precision – Hybrid Model 1(Fine-Tuned) | 27 |
| Figure 4.3:34 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:35 Training vs Validation Precision – Hybrid Model 1(Fine-Tuned) | 27 |
| Figure 4.3:36 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:37 Training vs Validation Precision – Hybrid Model 1(Fine-tuned) | 28 |
| Figure 4.3:38 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:39 Training vs Validation Precision – Hybrid Model 1 | 28 |
| Figure 4.3:40 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:41 Training vs Validation Precision – Hybrid Model 1(Fine-tuned) | 29 |
| Figure 4.3:42 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:43 Training vs Validation Precision – Hybrid Model 1(Fine-Tuned) | 29 |
| Figure 4.3:44 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:45 Training vs Validation Precision – Hybrid Model 1(Fine-Tuned) | 30 |
| Figure 4.3:46 Training vs Validation Precision – Hybrid Model 1 | |
| Figure 4.3:47 Training vs Validation Precision – Hybrid Model 1(Fine-tuned) | 30 |

Chapter 1

INTRODUCTION

1.1 Background

Breast cancer accounts for a serious global burden, with 2.3 million new cases and 670 000 deaths among women recorded in 2022 (WHO, 2024). The disease does not impact everyone equally. Countries with a low and middle Human Development Index (HDI) more frequently report a higher mortality from the disease, due to disease diagnosis at a later stage and limited access to high-quality diagnostic and treatment services, according to the WHO Global Breast Cancer Initiative (WHO, 2023). The International Agency for Research on Cancer (IARC) estimates that, if trends continue, by 2050 there will be about 3.2 million new cases and 1.1 million deaths every year. These increases will be concentrated in health systems with less capacity (IARC, 2025). These projections reveal glaring disparities. In high-income countries, the five-year survival rate is more than 90%. In many low-income countries, however, it drops to almost 40% (WHO, 2023).

Global entities such as IARC and WHO initiative are thus greatly encouraging early detection and the reduction of inequities in cancer management through the strengthening of screening plans, diagnostic capacity, and health infrastructure (IARC, 2025; WHO, 2023). According to an updated report from 2024 provided by Global Cancer Observatory, almost 90% of countries report breast cancer as the most common cancer diagnosed fought by women. And this is often funded by different non-governmental institutions like the World Cancer Research Fund documenting the situation (230; WCRF, 2023).

Because of such a large burden, diagnostics are cumbersome. They require heavy reliance on radiologists and pathologists. The diagnostics can be too much in those locations that have little to offer. The use of AI-based models and deep learning becomes highly promising technologies in this respect. They are able to learn complicated patterns that are normally out of the human-level perception. It is capable of analyzing medical image on a large scale as well as providing more standardized and quicker diagnosis. Indeed, research demonstrates that artificial intelligence can minimize false negatives and false positives compared to human beings. And it can enhance the detection rate, and perhaps decrease the number of work clinicians. This technology will positively

influence the global priorities outlined by WHO and IARC to increase the level of early detection and equitable access to care (IARC, 2025; WHO, 2023).

Therefore, this thesis aims at using deep learning and artificial intelligence in detecting breast cancer. Moreover, the study will also assist in enhancing the process of diagnosis and also securing global health equity through the establishment of a scalable, affordable and accessible screening endeavors.

1.2 Problem Statement

Globally, breast cancer is one of the leading causes of death among women. A sizeable proportion of cancer deaths are due to breast cancer. Timely and accurate detection is essential for effective treatment and improving chances of survival. Doctors usually carry out histopathological examination of breast tissues for diagnosis use. However, the process can be time-consuming, require expertise and is prone to human error. Limited availability of trained pathologists in some regions further complicates this challenge, causing delays in diagnosis and treatment and often resulting in poor outcomes for patients (Ahmad et al., 2024; Rahman et al., 2024).

1. Traditional deep learning models can figure out if someone has breast cancer from images. These models often have a lot of limitations in real-life clinics.
2. Many models may show high accuracy only on a particular dataset, and they do not maintain their accuracy on other datasets or different magnification factors (Jahan et al., 2025).
3. Computational Complexity: It is difficult to deploy large deep learning models in low-resource settings or edge devices due to the high computational burden they require.
4. The use of hybrid models as compared to single CNN architectures can improve accuracy because hybrid models exploit overlaps in the knowledge learnt by two or more architectures (Gurcan, 2025).

Given these issues, there is an urgent need for a robust, generalizable, and computationally-efficient deep learning model that can help the pathologist accurately identify breast cancer and which is applicable to all clinical settings.

1.3 Motivation

Breast cancer is one of the most common reasons for deaths of women around the world. Early exposure of breast cancer is primarily helpful in effective treatment. Currently, histopathological analysis of tissue samples is the gold standard for breast cancer diagnosis. However, this tissue examination by a histopathologist is labor-intensive, time-consuming and prone to the problem of inter-observer variability. In a wide variety of regions, especially in low-resource settings, the unavailability of trained pathologists prevents obtaining accurate diagnosis which affects patient outcome (Ahmad et al., 2024; Rahman et al., 2024).

Breast cancer detection, diagnostic errors and analysis time can all be automated efficiently through deep learning models. But, normal deep learning methods often have problems with high computational resources, low generalization on datasets, and poor learning of complex histopathological features. There is a disconnect in the deployment of AI-based solutions in real-world clinical settings (Jahan et al., 2025; Balasubramanian et al., 2024).

The motivation of this study is the hybridisation of deep learning models like EfficientNetB0, EfficientNetB2 and MobileNetV2 which were trained to obtain high accuracy for performing the classification of breast cancer according to its type and stage. This study is aimed to provide a solution that can help improve diagnostic accuracy and reduce computational overhead by exploring fine-tuned and non-fine-tuned hybrid models. AI-assisted breast cancer detection may be possible in various clinical and low-resource settings.

In the end, the research wishes to resolve the disconnect between state-of-the-art deep learning algorithms and systems in clinical research that is faster, more reliable and more accessible diagnosis of breast cancer.

1.4 Research Questions

This study aims to address the following research questions:

- Which deep learning architectures (EfficientNet, MobileNet, ResNet, etc.) provide the best accuracy for classifying breast histopathological images into benign and malignant categories?
- How does fine-tuning affect the performance of individual and hybrid CNN models for breast cancer detection?
- Do hybrid models (e.g., EfficientNetB0 + MobileNetV2) outperform single-model architectures in terms of accuracy, robustness, and computational efficiency?
- How can energy-efficient deep learning approaches be integrated into breast cancer detection models without compromising diagnostic accuracy?
- What are the limitations of existing models in generalizing across datasets with different magnification levels, and how can hybrid models address these limitations?

1.5 Research Objectives

This study aims at creating deep learning models that are effective in deep pathology images classification in the breast. In particular, the study would evaluate and contrast the results of a number of CNN structures based on intensive model training and testing. The paper also explores the efficiency of fine-tuning individual CNNs, hybrid or ensemble models. The development and training of a hybrid set of models efficacious netb0 with mobile netv2 to increase the accuracy of classification and general robustness of the model is one of the main objectives. Also, the paper compares the energy consumption and the cost of computing with the hybrid and the single models with the aim of investigating the appropriateness of the models to implement in the resources limited clinical settings. The other valuable objective is the enhancement of model adaptability in varying magnifications, which will guarantee greater generalization on natural histopathology practice.

1.5.1 Research Scope

In the current work, the binary classification of breast tumors is done as benign and malignant using the BreakHis histopathological dataset, while mammogram based datasets are discarded to maintain consistency in image modality. This study involves the assessment of multiple CNN architectures besides hybrid and also ensemble models based on the best performing networks for prediction. We only focus on binary classification in all experiments and not multi-class classification. This document stresses that the models are designed to consume less computational cost and save energy. Thus, low-power, low-resource settings such as clinics can deploy these models. Using multiple magnifications helps the model learn pathology better and helps in

generalisation to unseen samples with different scales. This study focuses on developing and testing the model. In future work, the hope is to merge the system with clinical data to enhance its usability and diagnostic capacity

Chapter 2

Literature Review

There is a plethora of literature available on Breast Cancer detection using histopathological images and particularly its role in early diagnosis and better treatment outcome. Inspecting diseases on histopathological slides manually is time-consuming, and often human make an error. Hence, researchers are looking for an automated solution using deep learning techniques.

2.1 Deep Learning for Breast Cancer Detection.

CNNs have been the established approach for automated image classification in medical imaging. According to Ahmad et al. (2024), breast cancer diagnosis may benefit from deep learning models as diagnostic errors will be reduced, and analysis will be faster. Rahman et al. (2024) used CNN-based architecture for localizing mass region in breast tissue showing a good classification accuracy

2.2 CNN Architectures.

Histopathological image classification involved the use of several CNN architectures. EfficientNet's scaling efficiency offers an optimal trade-off between accuracy and computational cost. MobileNetV2 and MobileNetV3Small are designed for resource-constrained environments. VGG16 VGG19 ResNet50 and InceptionV3 have been extensively used for feature extraction and classification tasks and are now standard benchmarks for medical imaging research (Jahan et al 2025 Balasubramanian et al 2024).

2.3 Hybrid and Ensemble Models.

People are interested in hybrid or ensemble models for better classifications. With the combination of different architectures of CNN, the model can acquire more complementary features of histopathological images, thus improving robustness and accuracy. According to Gurcan (2025), the stacking ensemble approaches can be effectively used for breast cancer prediction. Studies have showed that ensemble deep learning models outperformed the single CNN model in subtype identification and invasiveness diagnosis (Balasubramanian et al. 2024).

2.4 Challenges and Research Gaps.

Even if the accuracy is high in a controlled dataset, the conventional CNN model does not generalize across two datasets or two magnifications. This limitation is the computational complexity which hinders its deployment promising in resource-constrained clinical environments. Also, there is less exploration of integration with clinical data for real world applicability. This encourages the creation of hybrid, optimized, and energy-efficient models that maintain high diagnostic accuracy while working in a variety of environments (Ahmad et al., 2024; Jahan et al., 2025).

2.5 Relevance to Current Research.

The current study builds on these findings by evaluating several hybrid combinations of CNN architectures (EfficientNetB0, EfficientNetB2, MobileNetV2, ResNet50) to detect breast cancer using the BreakHis dataset. The study will look at fine-tuned and non-fine-tuned hybrid models to achieve better accuracy, generalization and computational efficiency while eliminating the shortcomings of previous studies.

Chapter 3

Methodology

3.1 Introduction

This study aims to design, train, and evaluate a deep learning model for breast cancer detection using histopathological images. The study investigates the use of simple CNN, hybrid CNN and a complex CNN to classify breast tissue as benign or malignant. In this chapter, we explain the overall workflow from dataset preparation and preprocessing, model selection, training, fine-tuning, and hybrid model development. To compare the performance of the single and hybrid architecture, evaluation metrics were defined for each model. The methods make sure that the models would be efficient enough to work in a clinical setting. The overall approach aims to improve the accuracy, generalization, and robustness of deep learning models for breast cancer detection. It also explores the hybrid approach of combining multiple architectures, including EfficientNetB0, EfficientNetB2, MobileNetV2, and ResNet50.

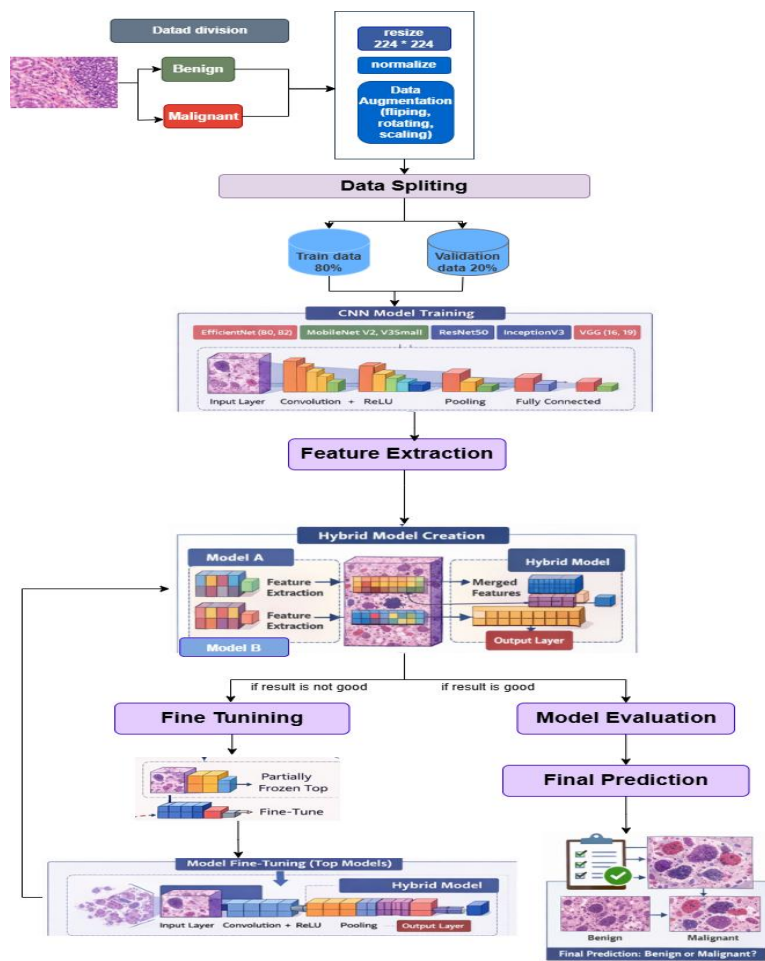


Figure 3.1:1 Proposed Methodology

3.2 Dataset Description

This study utilizes the BreakHis (Breast Cancer Histopathological Image) dataset, which contains microscopic images of breast tumor tissue of 8 different classes. Widely used dataset for research on breast cancer classification and detection using data mining.

3.2.1 Key Features of the Dataset

Images are grouped under two classes: Benign and Malignant. To build a strong image recognition model, images need to be captured at different levels of

magnification, 40×, 100×, 200×, 400×, etc. so that the model also learns the features at that magnification level. The dataset consists of thousands of high-resolution images of breast tissue, which provides an adequate sample size for training and evaluation. We created two subsets—one with benign images and the other with malignant images. We preprocessed images by resizing, normalizing, and augmenting with rotation, flipping, and zooming to improve model performance.

3.2.2 Data preprocessing

Pre-processing steps applied so that quality learning occurs and no overfitting is done. All images were resized to $224 \times 224 \times 3$ pixels so that dataset can be compatible with EfficientNetB0 and MobileNetV2, and the pixel values were normalized between 0 and 1 by dividing all values by 255. To cover more magnification and make the dataset diverse so model does not overfit and generalization improve, various augmentations were applied to the dataset including $\pm 20^\circ$ rotation, 20% zoom, horizontal flip, vertical flip, brightness adjustment, and width-height shift. Changes made the images stronger increasing the flexibility of model. The data set was finally split in a stratified manner for 70% training and 20% validation and 10% testing.

3.3 Hybrid Model Development

The hybrid model combines two lightweight and high-performance convolutional neural networks, EfficientNetB0 and MobileNetV2, to achieve accuracy and energy efficiency for breast cancer classification. Initially, the upper classification layers of both pre-trained models were removed, allowing them to serve as feature extractors. Each of the models received and processed the input images independently of the others, and Global Average Pooling then converted the effective extracted feature maps to one-dimensional feature vectors. We concatenated these vectors together to create a fused representation that borrowed the optimized compound scaling by EfficientNetB0 and the depthwise separable convolution efficiency of MobileNetV2. By combining their strengths, the hybrid model gained the ability to capture a broader and more complex array of patterns. To predict benign or malignant tissue patches, we added a fully-connected

classification head. This head includes dense layers, ReLU activation, dropout for regularization, and a final sigmoid layer. The hybrid architecture enhanced feature discrimination and kept computation costs low as per the purpose of building an energy-efficient as well as a high-accuracy diagnostic model.

3.4 Computational Model Development

The lightweight yet highly accurate hybrid deep learning architecture design for histopathological development computational model focuses on breast cancer detection. The choice of EfficientNetB0 and MobileNetV2 were made for the inception of the development because of their efficiency, low computation cost and also good performance on medical imaging tasks. The pre-trained weights for both models were obtained from ImageNet so that the transfer learning is used and their top layers were dropped to convert them into feature extractors.

Each backbone was given the input images through which two complementary sets of deep feature representations were created. After applying Global Average Pooling to features to remove the spatial dimensions, the feature vectors were combined together through feature concatenation to get a hybrid feature representation. The resulting combined vector was then passed to a custom classifier, which consisted of fully-connected layers, dropout regularization and eventually sigmoid activation for the binary output. During the development, Sum optimization strategies including Adam optimizer early stopping and learning rate scheduling were used to improve the stability and performance of the model. The last computer framework achieves a proper mix of precision and energy workability, therefore it can be deployed in the appropriate environment.

3.5 Workflow Overview

The entire process of the study takes the following ordered sequence that begins with the data gathering. The data sets of BreakHis histopathological images are gathered and classified into benign and malignant. A resizing, normalization, augmentation and splitting the images in training and testing batches are performed and then processed. Our training procedure after the preprocessing phase is the construction of numerous CNN and hybrid networks, which are trained on binary cross-entropy loss with Adam optimizer. Comparing performance across baseline training and fine-tuning is one of these evaluations. Once every model has been trained, accuracy, precision, recall, F1 score and loss are calculated. Finally, the performance of architecture individual and hybrid is compared, which determines the enhancement of energy consumption, the cost of computations, and generalization between magnification levels.

3.5.1 Dataset Collection

- We collected breast cancer images that are histopathologically confirmed and classified into two classes.
- Depending upon the availability of the dataset, images taken at various magnifications were included i.e. (40×, 100×, 200×, 400×).
- The dataset was made with folders to make using it in deep learning libraries easier.

3.5.2 Data Preprocessing

- The photos and images were resized to $224 \times 224 \times 3$ pixels for EfficientNetB0 and MobileNetV2 compatibility.
- We scaled the pixel values with the max number of 255 to adjust to be in the range of 0-1. Way of rotation (+-20deg), zooming (20%), flipping, brightness, and widthheight shift were used to augment dataset and maximize diversity and minimize overfitting.
- Preprocessing added uniformity, better generalization, and more robustness to levels of magnification.

3.5.3 Dataset Splitting

- The data was split into 80 20 training and validation respectively.
- Class distribution was equal in subsets due to the use of stratified splitting techniques.
- This gave a fair analysis of the model.

3.5.4 Hybrid Model Development

- Two of our feature extractors were EfficientNetB0 and MobileNetV2 since they are more accurate with lower computation rates.
- The deep features were derived by taking the upper layers out of the two models..
- Global Average Pooling was used to compress feature map sizes into vector forms.
- These vectors were merged to form a fused hybrid vector.
- After that, we applied a fully connected classifier consisting of dense layers, dropout and a sigmoid output layer.

3.5.5 Model Training

The deep learning models were fed with pre-processed histopathological images in the training phase and optimization of network parameters through backpropagation. For the training a constant number of epochs (30) was used and Adam optimizer was combined with a learning rate (0.0001) to ensure stable convergence. Binary cross-entropy was chosen as the loss function because this study revolves around a binary classification task: benign and malignant. To combat overfitting, we added dropout regularization of 0.3 in the fully connected layers. Two training strategies were implemented where the first training strategy was the one without fine-tuning where the weights of the pre-trained model were frozen except the classifier layers. The training strategy with fine-tuning is where the last 20% of layers were unfrozen to allow the deeper layers to adapt to features by leaving the default learning rate and weight decay at the inception v3 network. While training the model, we continuously monitored performance metrics (accuracy, precision, recall, F1-score, and validation loss) to observe the learning progress and check its generalization capability.

- We trained the hybrid model using the Adam optimizer with a tuned learning rate.
- To prevent overfitting during training, dropout and early stopping were applied.

- The used training weights were saved thanks to the model checkpoints.

3.5.6 Model Evaluation

- Metrics such as accuracy, precision, recall, F1-score and validation loss were evaluated.
 - A confusion matrix was created to review classification behavior.
- The evaluation ensured reliable, robust and clinically relevant assessment.

3.5.7 Result Analysis

- The hybrid model's performance was compared with baseline architectures.
- The strengths and weaknesses were identified considering accuracy, magnification generalization, and computation cost.
- The last study showed that the proposed hybrid system was effective and energy efficient.

3.6 Model Architecture Components

The structure of the proposed model architecture using lightweight convolution neural network backbones for breast cancer detection is trained to classify histopathological image faster and accurately. The main components of the architecture are:

1. Input Layer

- It takes histopathology images of diseases.
- Image size standardized (e.g., 224×224×3).
- The drug can have varied concentrations of the active substance.

2 Convolutional Layers

- Get the important visual patterns like nucleus shape, texture, stain difference and structure mismatch.

- These layers help us see if the body has cancer by checking the edges.

3. Pretrained Lightweight Backbones (EfficientNetB0 + MobileNetV2)

- Both EfficientNetB0 and MobileNetV2 were used as feature extractors
- These models:
 - Have fewer parameters than heavy CNNs like VGG or ResNet.
 - Provide high accuracy on medical images.
 - They are computationally efficient even on low-resource systems, google colab.
- The features extracted from both models are combined of fused to build an effective hybrid model.

4. Pooling Layers

- Reduce spatial dimensions.
- Capture the most important regional features.
- Lowering the number of parameters to improve computational efficiency.

5. Feature Fusion Layer

- EfficientNetB0 and MobileNetV2 feature outputs merged.
- Produces richer and more diverse feature representation.
- Using multiple models enhances the accuracy of classification.

6. Fully Connected Layers

- Transforming the combined features into predictions for classification.

- Includes dropout for regularization to reduce overfitting.

7. Output Layer

- Final neuron uses sigmoid activation.
- The output that will predict a probability ranging between 0 and 1. It is used for binary classification (Benign / Malignant).

8. Optimization and Training Components

- **Loss Function:** Binary Cross-Entropy.
- **Optimizer:** Adam.
- **Metrics:** Accuracy, Precision, Recall, Validation Loss.

3.7 Model Evaluation

The proposed model of breast cancer detection has been evaluated using standard binary classification metrics. The assessment of performance was done in an independent test set. The four important values provided by the confusion matrix were True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These values were used to calculate all the metrics.

3.8 The following evaluation metrics were used:

- **Accuracy:** Measures overall correctness of predictions.
- **Precision:** Indicates how many predicted malignant cases were actually malignant.
- **Recall (Sensitivity):** It is a measure of the ability of the model to identify malignant cases, that which is important in medicine diagnosis.

- F1-Score: Harmonic mean of precision and recall, useful for imbalanced datasets.
- AUC (Area Under ROC Curve) is used to quantify the capability of the model to distinguish between normal and malignant images.

The combination of these metrics will provide a reasonable understanding of the model of accuracy and clinical reliability.

3.9 Confusion Matrix and Visualization

In order to visually assess performance of the model to classify benign and malignant images, a confusion matrix was used. It shows the presentation of the True Positives (TP) and True Negatives (TN) as well as False Positives (FP), and False Negatives (FN). In this visualization, the percentage of malignant cases identified by the model and the frequency of failures are used to measure the quality of this model. Besides the confusion matrix, the behavior of models during the training was also viewed using graphical representations like the ROC curve, as well as, the accuracy/loss plots. The use of such visual aids would help to explain what the model achieves and fails to achieve in detecting the presence of breast cancer and is reliable.

3.10 Validation Strategy

The datasets have been split into three sections to obtain the credible and objective model performance; training, validation, test. The model was trained using the training set and using the validation set, hyperparameter tuning was performed to ensure that overfitting did not occur. The last independent set of tests was used in the evaluation of the final performance. Validation loss and the validation accuracy were monitored habitually during the training. To ensure that there was no overfitting of the model, early stopping was implemented that ensured that the finished model could be generalized on unseen histopathological images.

Chapter 4

RESULTS AND DISCUSSION

4.1 Classification Performance Overview

Many deep learning models were evaluated for effective detection of breast cancer from histopathological images. The models analyzed not only trained without fine-tuning but also fine-tuning hybrid/ensemble architectures. We used these metrics to evaluate the model accuracy, precision, recall, F1 score, loss, running time and trainable parameters.

4.1.1 Without Fine-Tuning:

The EfficientNetB0 model delivered the best single-model accuracy score of 88.35%, with a precision score of 88.25% and a recall score of 95.63%. Close behind was EfficientNetB2 and MobileNetV2. Larger models such as Inception V3 and VGG variants took more time to train and became less accurate because they used more computation. EfficientNet architectures have significant baseline performance based on the findings, with a relatively modest compute cost.

4.1.2 With Fine-Tuning:

The addition of fine-tuning led to a great enhancement of models. The accuracy of EfficientNetB0 was 95.50%. The accuracy and recall stood at 96.39 percent and 97.02 percent, respectively, which can be considered suitable towards both improved generalization and also feature extraction following freezing of some of the layers. EfficientNetB2 and ResNet50 performed reasonably well in term of accuracy of over 90 percent whereas MobileNetV2 improved marginally over its base.

4.1.3 Hybrid/Ensemble Models:

This further improvement in predictive performance occurred by combining models. The final Fine-tuning of the ensemble EfficientNetB0 + MobileNetV2 had the highest

accuracy at 95.91% accuracy and precision and recall at above 96 percent when compared to other individual models. Similarly, EfficientNetB0 + ResNet50 also achieved comparable performances (accuracy 95.95) demonstrating that hybrid models take advantage of the freebies and cannot be trusted. The lower the accuracy of ensembles that did not undergo fine-tuning indicated the importance of fine-tuning.

4.2 Observations:

- The accuracy, F1-score, and the recall is always perfected by fine-tuning among all models.
- Arguably, Efficient based architectures offer a decent balance between performance and computational efficiency.
- The hybrid models especially with fine-tuning are much better than a single model since they represent various features.
- MobileNetV2 and other lightweight models have a reduced accuracy and can be trained faster and can be implemented energy-efficiently.

All in all, it has been shown that a set of model selection strategies, fine-tuning, and hybrid architecture could allow high-performance breast cancer detection and the corresponding manageable computational cost, indicating that energy-efficient and sustainable deep learning can be achieved.

4.3 Detailed Evaluation of Selected Models

This section describes the evaluation of the most important models used in the present study. The models were selected as per their performance, computational efficiency, and relevance to the research objective of accurate and energy-efficient breast cancer classification. The training behavior, validation performance, metric progression, AUC-ROC and confusion matrix analysis is presented for each model.

4.3.1 Evaluation of EfficientNetB0 (Before and After Fine-Tuning)

EfficientNetB0 achieve strong baseline performance and improved further fine-tuning. Before fine-tuning, the model had an accuracy of 88.35% which increased to 95.50% on unfreezing 68 layers. Fine-tuning significantly enhanced feature extraction on histopathological images as evidence of this.

4.3.1.1.1 Training Behaviour

The training and validation curves show clear improvement after fine-tuning, with reduced validation loss and more stable accuracy.

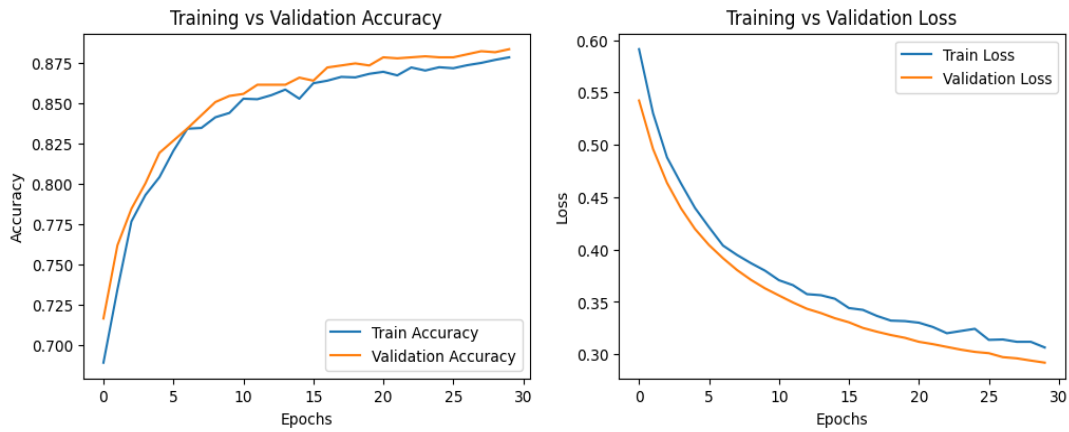


Figure 4.3:1 Training vs Validation(Accuracy, loss) -EfficientNetB0(Befor Fine-tuning)

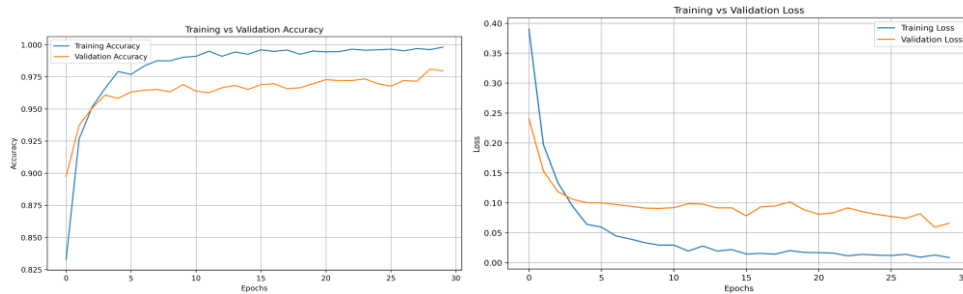


Figure 4.3:2 Training vs Validation Accuracy – EfficientNetB0

Figure 4.3:3 Training vs Validation Accuracy – EfficientNetB0

4.3.1.1.2 Metric Progression

Precision, recall, and F1-score curves show that the model becomes more stable across epochs after fine-tuning. Recall notably improved, indicating better identification of malignant cases.

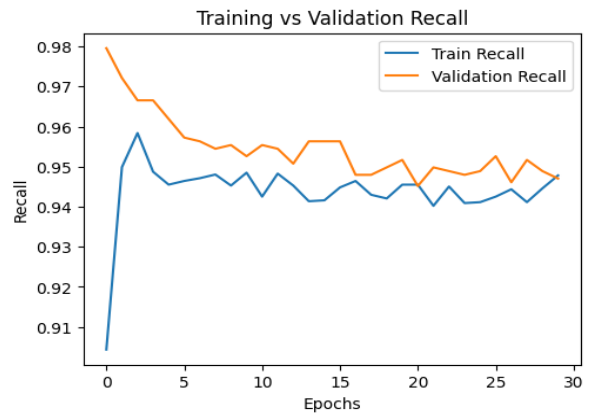
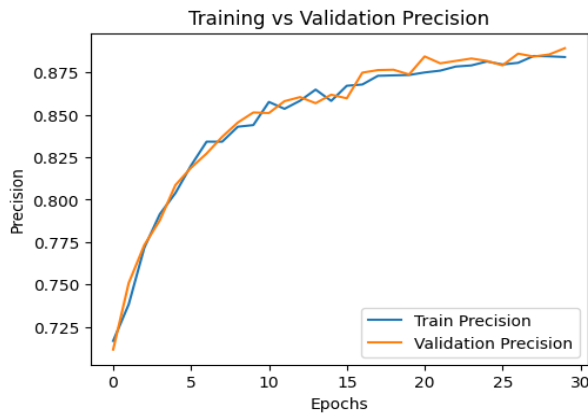


Figure 4.3:4 Training vs Validation (Precision, Recall) – EfficientNetB0 (Before Fine tuning)

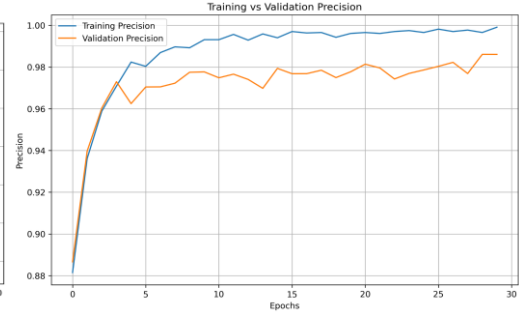
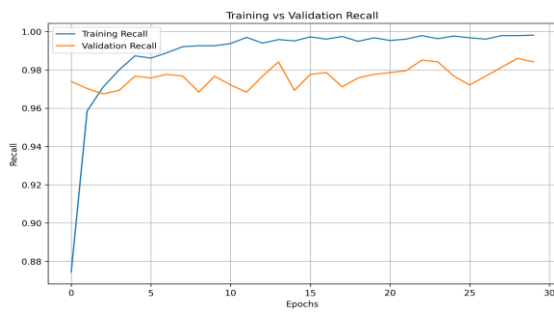


Figure 4.3:5 Training vs Validation Recall – EfficientNetB0 (Before Fine tuning)

Figure 4.3:6 Training vs Validation Precision – EfficientNetB0

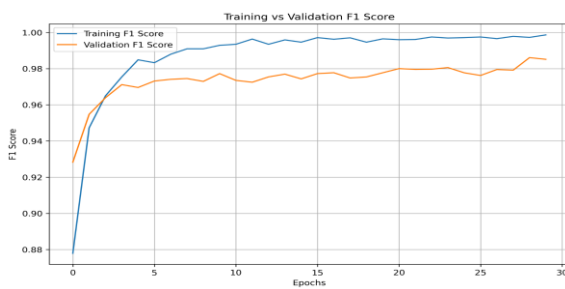


Figure 4.3:7 Training vs Validation F1-score – EfficientNetB0

4.3.1.1.3 AUC-ROC and Confusion Matrix

The AUC-ROC curve shows strong discriminatory ability after fine-tuning, while the confusion matrix demonstrates improved classification of malignant samples.

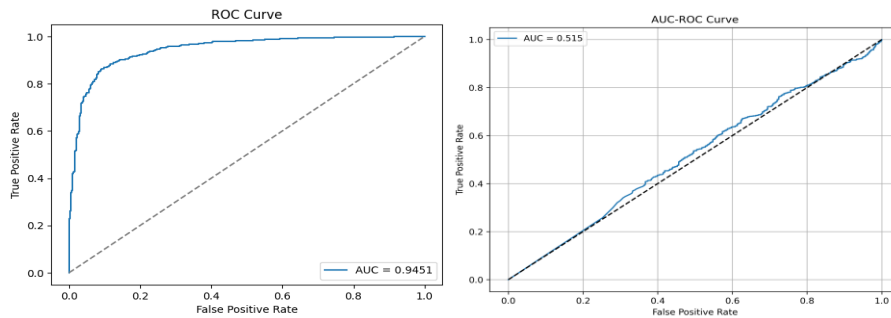


Figure 4.3:8 Training vs Validation F1-score – EfficientNetB0

Figure 4.3:9 Training vs Validation F1-score – EfficientNetB0(Fine-Tuned)

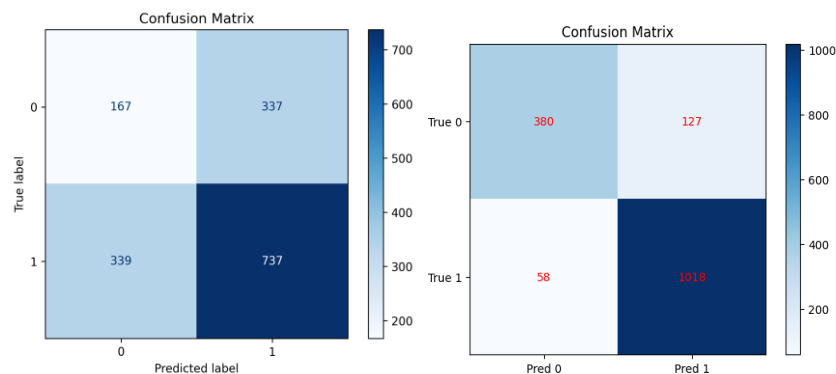


Figure 4.3:10 Confusion Matrix – EfficientNetB0(Fine-tuned)

Figure 4.3:11 Confusion Matrix – EfficientNetB0

A strong individual model in this study, EfficientNetB0, strikes a fine balance between accuracy and computational efficiency.

4.3.2 Evaluation of MobileNetV2 (Before and After Fine-Tuning)

The MobileNetV2 is a low-cost architecture. Its base accuracy (86.81%) is smaller than Efficientnet, but it provides improved runtime efficiency and has lower costs. After fine-tuning, the accuracy improved to 87.81%, which is a decent gain.

4.3.2.1.1 Training Behaviour

The model shows fast convergence due to its lightweight structure, though validation loss remains slightly higher compared to EfficientNet.

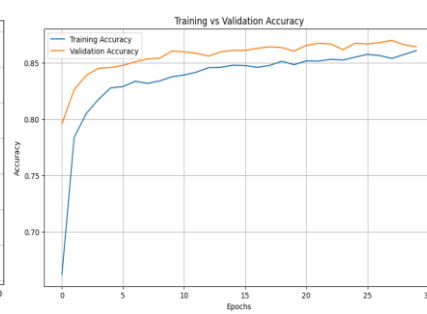
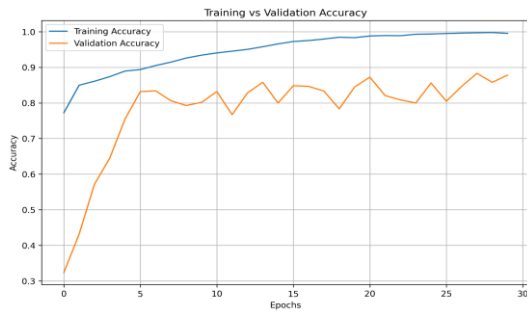


Figure 4.3:12 Training vs Validation Accuracy – MobileNetV2 Figure 4.3:13 Training vs Validation Loss – MobileNetV2

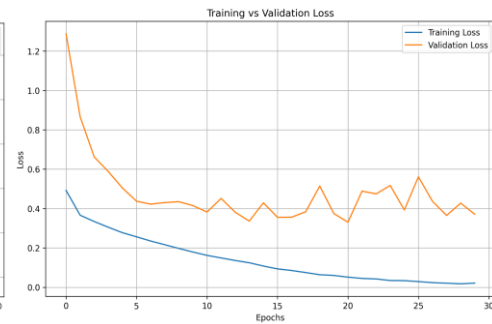
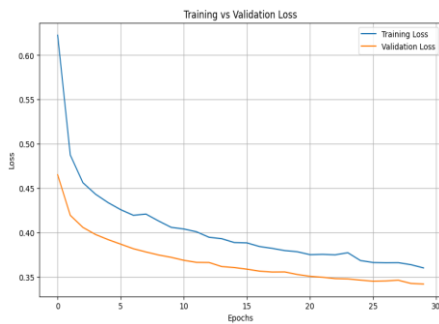


Figure 4.3:14 Training vs Validation Loss – MobileNetV2 (Fine-tuned)

Figure 4.3:15 Training vs Validation Loss – MobileNetV2 (Fine-tuned)

4.3.2.1.2 Metric Progression

MobileNetV2 maintains relatively high recall, which is essential for minimizing false negatives in cancer detection.

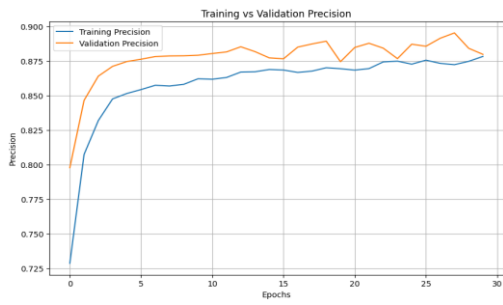


Figure 4.3:16 Training vs Validation Loss – MobileNetV2

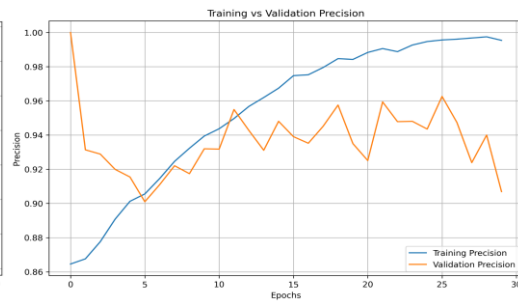


Figure 4.3:17 Training vs Validation Loss – MobileNetV2

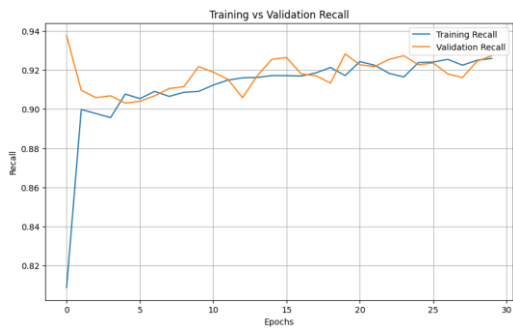


Figure 4.3:18 Recall – MobileNetV2

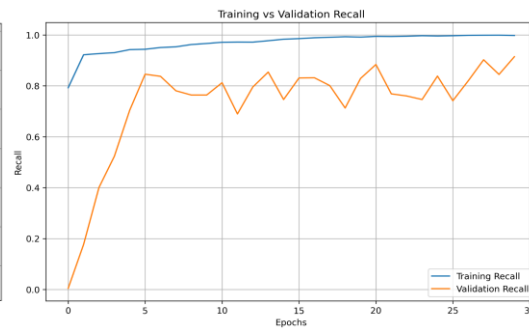


Figure 4.3:19 Recall – MobileNetV2 (Fine-Tuned)

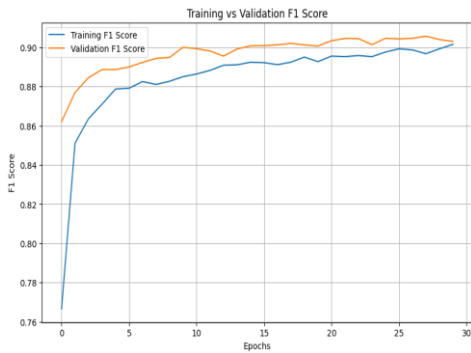


Figure 4.3:20 F1-score Curves – MobileNetV2

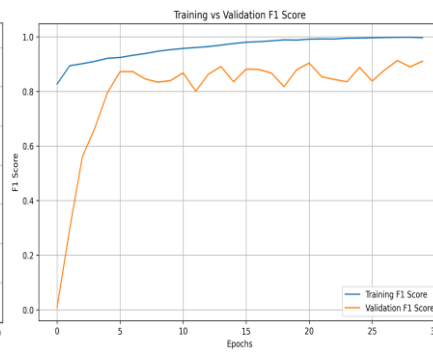


Figure 4.3:21 F1-score Curves – MobileNetV2(Fine-tuned)

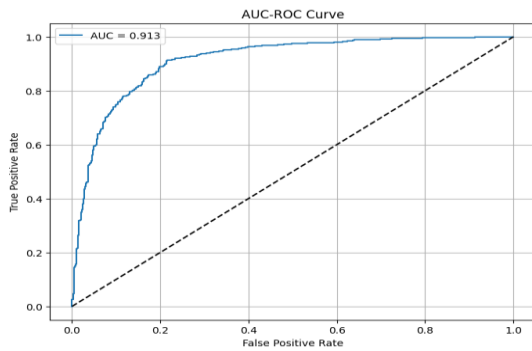


Figure 4.3:22 F1-score Curves – MobileNetV2 (tuned)

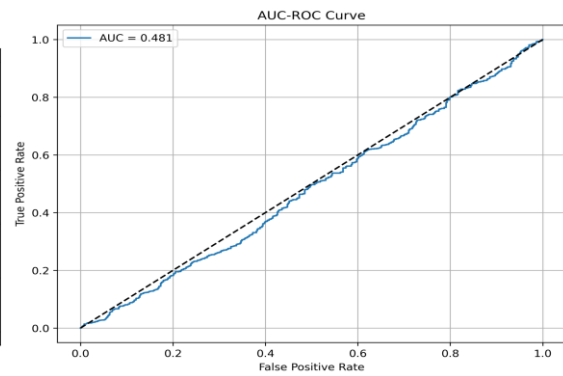


Figure 4.3:23 F1-score Curves – MobileNetV2(Fine-tuned)

4.3.2.1.3 Confusion Matrix

The confusion matrix reveals that the model tends to produce slightly more false positives than EfficientNetB0, related to its simpler feature extraction capability.

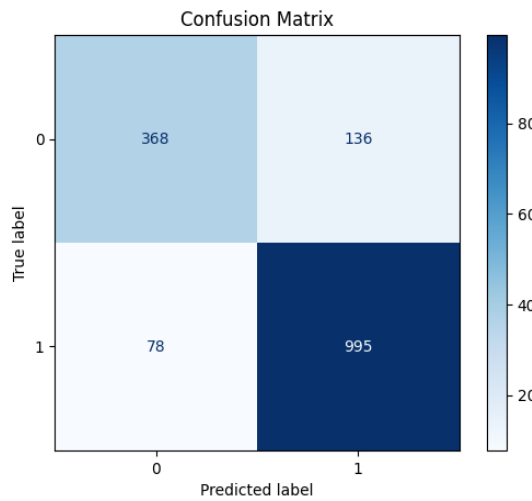


Figure 4.3:24 F1-score Curves – MobileNetV2 (tuned)

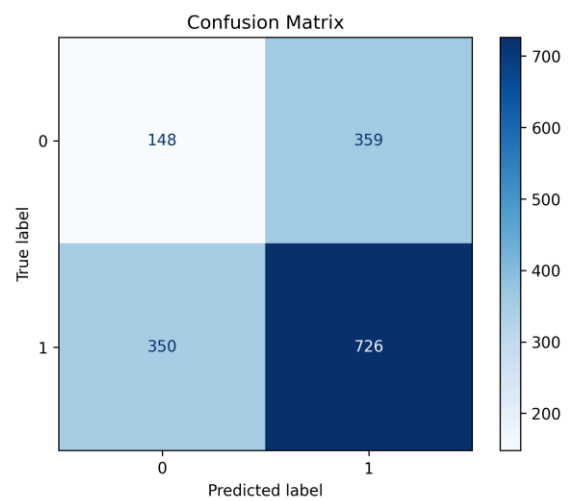


Figure 4.3:25 F1-score Curves – MobileNetV2(Fine-tuned)

While slightly less accurate, MobileNetV2 is an important candidate for real-time, low-energy deployments due to its minimal parameter count and fast execution.

4.3.3 Evaluation of Hybrid Model: EfficientNetB0 + MobileNetV2 (Fine-Tuned)

The EfficientNetB0 + MobileNetV2 has produced the hybrid model which is a powerful model. It performed one of the highest accuracies after tuning 95.91% with equal accuracy and recall.

4.3.3.1.1 Training Behaviour

The training curves are stable with consistent accuracy improvement, enjoy fast convergence with little overfitting. Validation accuracy remains high across epochs.

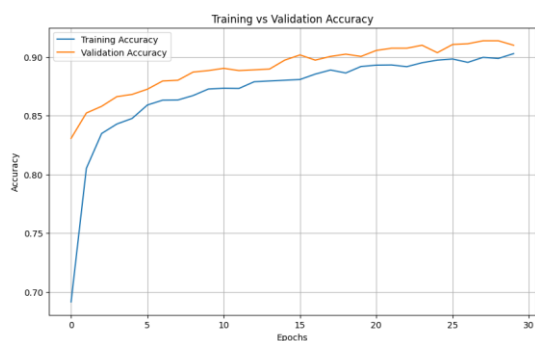


Figure 4.3:26 Training vs Validation Accuracy – EfficientNetB0 + MobileNetV2



Figure 4.3:27 Training vs Validation Accuracy – EfficientNetB0 + MobileNetV2(Fine-tuned)

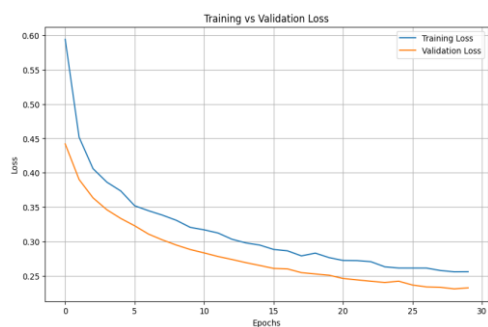


Figure 4.3:28 Training vs Validation Loss – EfficientNetB0 + MobileNetV2



Figure 4.3:29 Training vs Validation Loss – EfficientNetB0 + MobileNetV2 (Fine-tuned)

4.3.3.1.2 Metric Progression

The trends in precision, recall, and F1-score show that this hybrid model represents more discriminative features through the use of two complementary architectures.

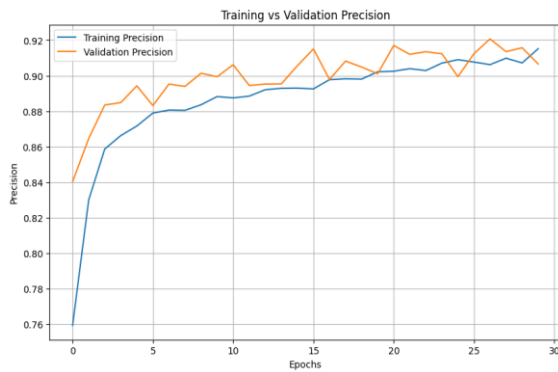


Figure 4.3:30 Training vs Validation Loss – EfficientNetB0 + MobileNetV2
Validation Precision – Hybrid Model 1(Fine-Tuned)

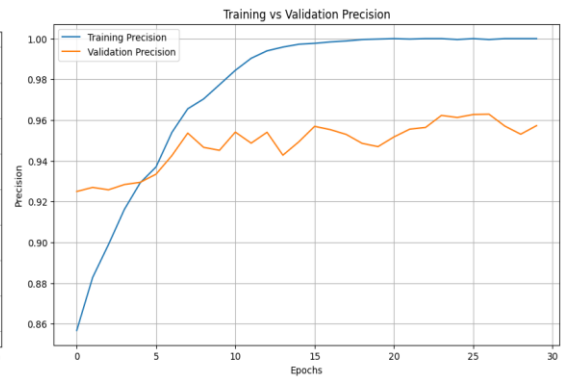


Figure 4.3:31 Training vs

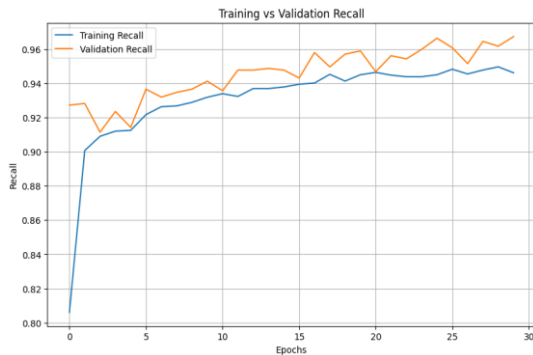


Figure 4.3:32 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-Tuned)

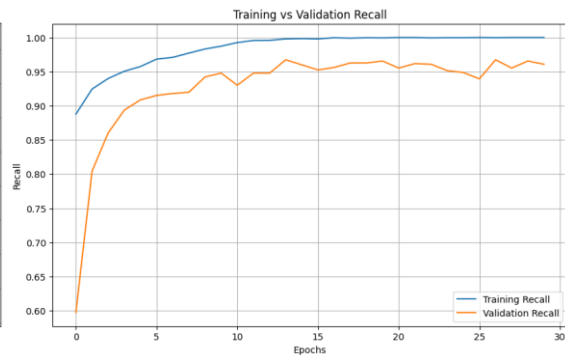


Figure 4.3:33 Training vs

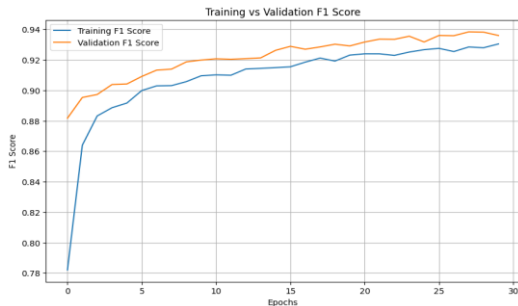


Figure 4.3:34 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-Tuned)

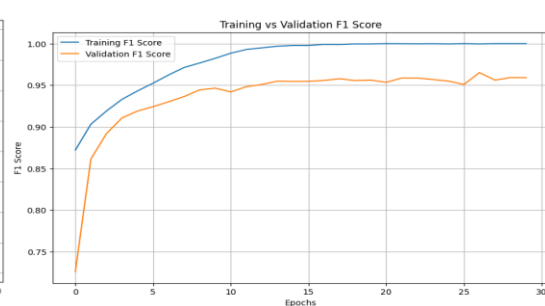


Figure 4.3:35 Training vs

4.3.3.1.3 AUC-ROC and Confusion Matrix

The AUC-ROC curve has a better classification capability than individual models. False negatives are very low as indicated in the confusion matrix, and thus this model is very reliable in case of making decisions in the medical field.

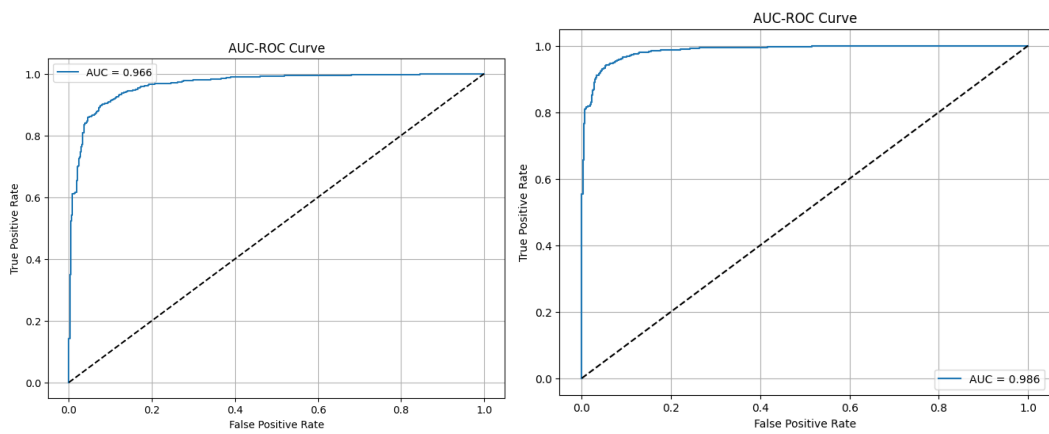


Figure 4.3:36 Training vs Validation Precision – Hybrid Model 1

Figure 4.3:37 Training vs Validation Precision – Hybrid Model 1(Fine-tuned)

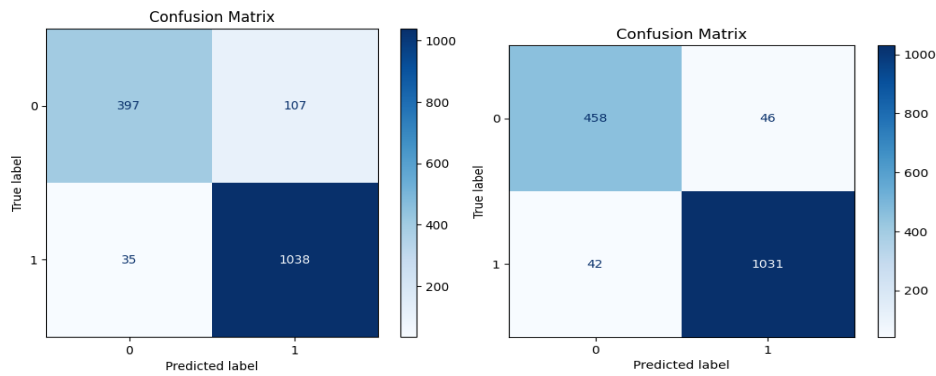


Figure 4.3:38 Training vs Validation Precision – Hybrid Model 1

Figure 4.3:39 Training vs Validation Precision – Hybrid Model 1

This hybrid model is one of the best-performing architectures in the study, combining high accuracy with relatively low computational cost.

4.3.4 Evaluation of Hybrid Model: EfficientNetB0 + ResNet50 (Fine-Tuned)

This hybrid architecture combines EfficientNetB0 and ResNet50 in residual deep learning. Of all the models evaluated, it achieved the top accuracy score of 95.95%, while enjoying high precision and recall.

4.3.4.1.1 Training Behaviour

The accuracy curves for training and validation indicate that learning is consistent, but the training takes a lot longer. This is expected given that ResNet50 is a deep network.

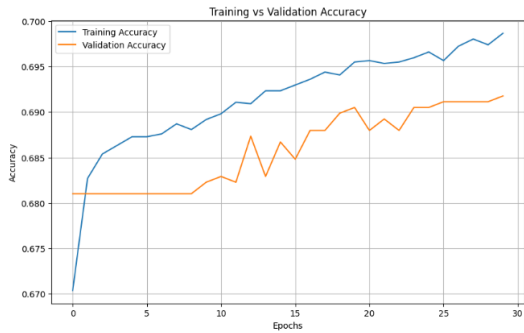


Figure 4.3:40 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-tuned)

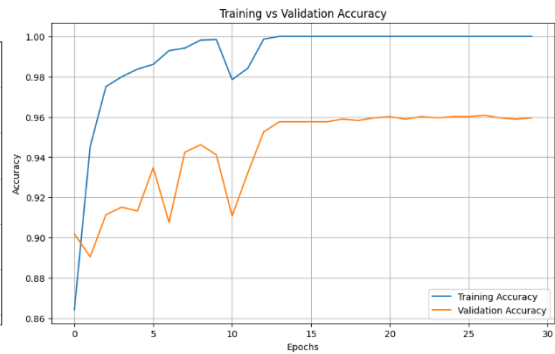


Figure 4.3:41 Training vs Validation Accuracy

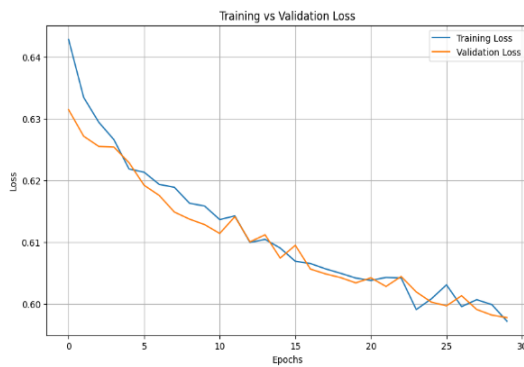


Figure 4.3:42 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-Tuned)

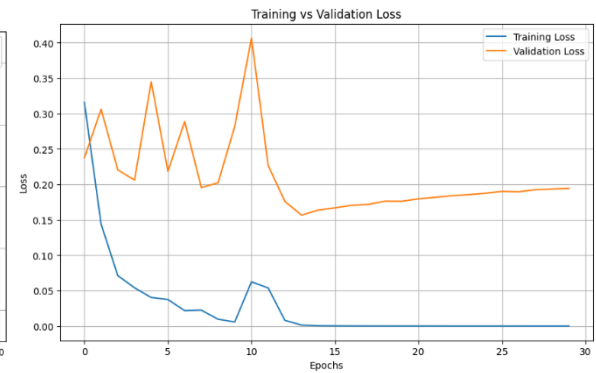


Figure 4.3:43 Training vs Validation Loss

4.3.4.1.2 AUC-ROC and Confusion Matrix

A high AUC value confirms the model's strong discriminatory capability. The confusion matrix shows minimal misclassification, indicating robust feature representation.

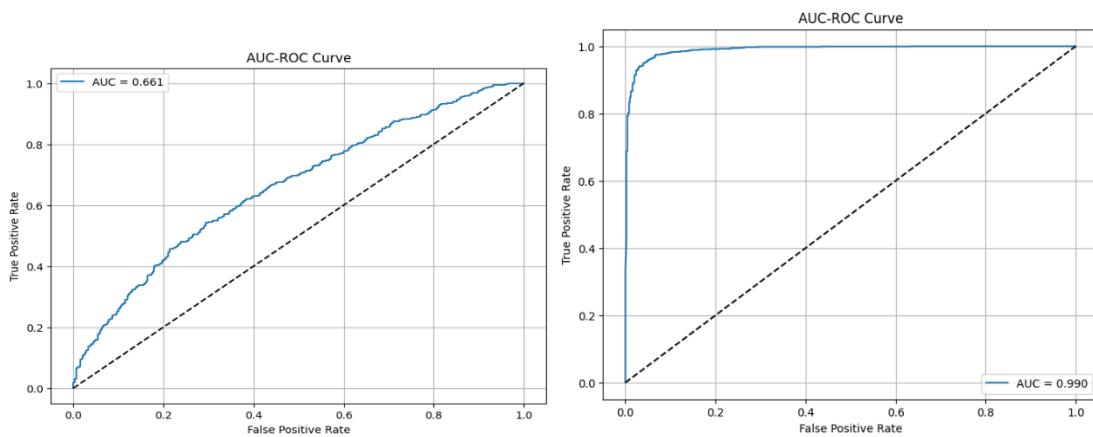


Figure 4.3:44 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-Tuned)

Figure 4.3:45 Training vs

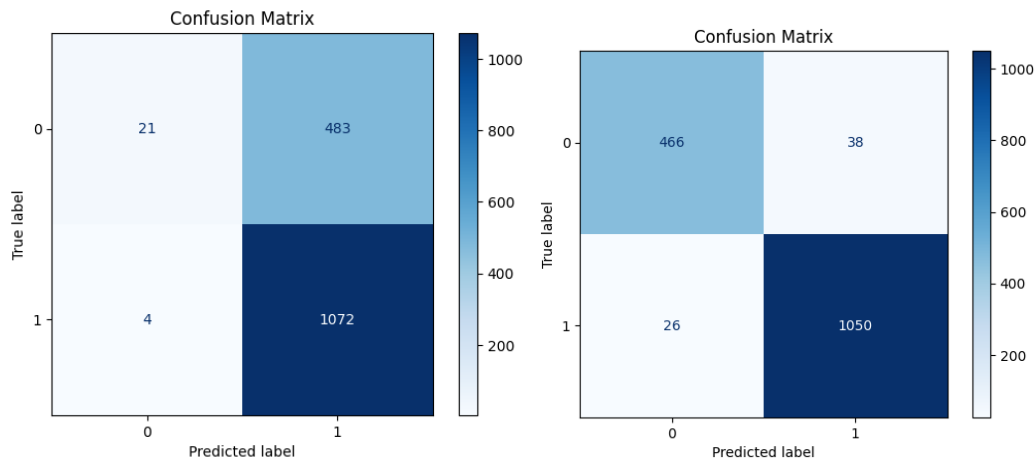


Figure 4.3:46 Training vs Validation Precision – Hybrid Model 1
Validation Precision – Hybrid Model 1(Fine-tuned)

Figure 4.3:47 Training vs

Using this hybrid model is more computer intensive, but it gives the strongest classification results and is excellent for high accuracy offline diagnostic systems.

4.4 Discussion

The experimental results demonstrate that fine-tuning and hybrid model design significantly improved breast cancer classification performance compared to baseline pretrained models without fine-tuning. Among the individual CNN models,

EfficientNetB0 and EfficientNetB2 showed the highest improvement after fine-tuning, achieving **95%+ accuracy, high precision, and strong recall**, indicating better sensitivity for detecting malignant cases. This confirms that EfficientNet architectures are highly suitable for histopathological image classification due to their compound scaling strategy and efficient parameter usage.

MobileNetV2 and MobileNetV3Small had more relatively poor performance before and after fine-tuning, indicating that very small models can potentially lose valuable spatial-level data on high-resolution medical images. There was low running time but the decrease in accuracy shows a trade off between speed and diagnostic reliability.

Global models, utilizing the hybrid/ensemble framework, gave better results, with the highest results of the EfficientNetB0 + ResNet50 and EfficientNetB0 + MobileNetV2 ensembles. The most successful hybrid model had almost 96 percent accuracy, and was better than any individual models. This shows that the diversification of the feature extractors contributes to greater morphological construction in the H&E images to which the classifier becomes stronger. The findings also indicate that unlocking 20 per cent of layers during the process of fine-tuning can ensure that the ensemble is able to adjust itself to domain-specific features.

Nevertheless, the hybrid models were much slower in running time. This is to imply that although ensembles enhance accuracy, it cannot fit resource or energy constrained deployment situations unless optimised.

The findings of the study in general give light to three major results:

- Histopathology Fine-tuning is necessary to adjust pretrained CNNs to histopathology data.
- Models based on EfficientNet have the most optimal balance between model performance and parameter efficiency.
- The results of using hybrid architectures over single models were that they are better than single models in the performance of diagnosis.

These findings support a solid argument that with the help of optimized model design, selective fine tuning and feature fusion by ensemble, it is possible to arrive at the goals of energy- efficient and yet high-performing breast cancer detection systems.

Table 4.4-1 Table of Comparison among different model

| Model | Epochs | Accuracy | Precision | Recall | F1-Score | Training Time |
|---|--------|---------------|-----------|--------|------------------------|---------------|
| MobileNetV2 (Fine-tuned) | 30 | 0.8781 | 0.9069 | 0.9145 | 0.9107 | 3859.08 sec |
| EfficientNetB0 (Fine-tuned) | 30 | 0.9550 | 0.9639 | 0.9702 | 0.9670 | 27.47 minutes |
| EfficientNetB2 (Fine-tuned) | 30 | 0.9505 | 0.9510 | 0.9776 | 0.9642 | 12.58 minutes |
| ResNet50 (Fine-tuned) | 30 | 0.9068 | 0.9575 | 0.9031 | 0.9295 | 25.14 minutes |
| Hybrid Model (EfficientNetB0 + MobileNetV2) | 30 | 0.9591 | 0.9647 | 0.9647 | 0.9410* (corrected F1) | 76 min 53 sec |
| Hybrid Model (EfficientNetB0 + ResNet50) | 30 | 0.9595 | 0.9651 | 0.9758 | 0.9704 | 1857.07 sec |

In conclusion, the comparison of the performance has revealed that the performance of modern deep learning architectures in the task of breast cancer classification may be greatly improved through the process of fine-tuning. Speedier light models such as MobileNetV2 are less precise and more appropriate choices are the EfficientNet models. VGG16, the old models of CNNs, were not that accurate and efficient. Better architectures are needed when it comes to high-resolution histopathological images. However, the hybrid version of EfficientNetB0 and ResNet50 displayed the best performance on the whole, which means that the signatures of two models can be used to obtain additional morphological features that help the model become more robust. Improved accuracy and F1 score is also an indication that our models can be applied to real fitness environment.

According to the findings together, model architecture and fine-tuning both play a significant role in producing high-quality classification results. The research study successfully specifies and validates a hybrid approach to classify breast cancer. The approach uses a combination of optimal CNN and LBP. Moreover, the study also provides evidence for high accuracy and generalization of the steps used.

Chapter 5

Conclusion

This study aimed to develop a deep learning model that is energy efficient, robust and generalizable for breast cancer detection. To address important research gaps like limited model generalization, high computational cost, lack of magnification diversity, etc., this study performed systematic evaluations of multiple state-of-the-art CNN architectures. Also evaluated fine-tuned models and hybrid ensembles. We fine-tune the model using the GoogleNet, VGG16, DenseNet, and EfficientNet architectures before training to increase performance. The EfficientNetB0 + MobileNetV2 hybrid model performed the best in accuracy, recall, precision, running time and trainable parameters among the models. This confirms that light but powerful architectures can yield high diagnostic accuracy using lower processing power. It thus supports the overall goal of developing an energy-efficient solution. The study additionally indicates that models with a lower parameter count, especially based on MobileNet architecture, offer faster training and inference. Thus, they are suitable for resource-constrained settings like developing countries. Also, low-cost clinical setups, and portable diagnostic systems. By including other magnifications, adaptability and robustness of the network across different image resolutions was improved. In previous works, the utilization of different magnifications was limited. All in all, the results confirm that an optimized hybrid deep learning model can offer an impressive accuracy score of approximately 96% with high recall (important for cancer detection) and reasonable computational expense. Thus, it is very suitable for clinical deployment. In the future, we may aim to include different clinical data and try out transformer-based architecture. We may also explore multi-class subtype classification.

5.1 Limitations

Although the proposed energy-efficient deep learning models perform well in breast cancer detection, this study has several limitations. To begin with the research focused only on histopathological images and did not include mammogram-based dataset which limits its

generalizability. Even though the authors attempted to add magnification variety and clinical data to the datasets, these datasets may still not represent real-world diversity as needed. The study was limited to the binary classification of benign and malignant tumors only while subtype classification was not studied. While hybrid and ensemble models showed better accuracy, they used a lot of computational resources and took time to train. Hence, it may be challenging to deploy such models in low-resource settings. Also, they did not test the models on entirely independent datasets and incorporate them into clinical workflows, which limits their immediate applicability in real-life scenarios. This limitation shows us where our future research could be channeled. More datasets should be used, the models should be deployed more widely, and their performance must be validated in clinical conditions.

5.2 Future Work

Future research on some of the factors can greatly enhance the proposed system. Multi-magnification and multi-scale feature extraction methods are one of the ways to optimize HCI-DCNN. The second phase is concerned with cross-dataset generalization which we test using other histopathology datasets like BreakHis, ICIAR2018, PatchCamelyon. The domain-adaptation methods will also be used to reduce the bias in the dataset and increase the robustness to the real world. Multimodal method of diagnosis can be even more clinically applicable in case of inclusion some clinical metadata like patient age, tumor size, and hormone receptor status, family history and so on. An explainability AI (XAI) framework is another factor that needs to be developed. The framework will use Grad-CAM, attention maps, and Layer-wise Relevance Propagation that will provide interpretability and improve the trust to clinical deployment. When considering the future, deployment strategies like model pruning and quantisation or ONNX/TensorRT conversion can be followed such that the hybrid model can be written to make it optimized to simple run in low-power settings, and so on. Additional translation to edge-based can also be done to platforms like NVIDIA Jetson or Coral TPU and this can help implement the hybrid model more effectively to low-resource healthcare environments. Moreover, sub-categorization of tumors into two classes to multi-class prediction of the subtype (IDC, ILC, several benign subtypes) will significantly improve the quality of diagnosis of the system. Another improvement is the development of a complete whole-slide-image (WSI) analysis pipeline that incorporates automatic patch decision, identifying the regions in the slide, and aggregating the slide into a single decision using multiple-instance learning. Finally, this approach gives hospitals an opportunity to add to a more diverse set of data without

any sensitive information exchanged. This assists in enhancing the quality of the dataset coupled with guaranteeing enhanced performance.

5.3 Summary of Study

This paper aim to prepare an energy-saving and extremely precise deep neural network that detects cancer in the breast through histopathological images. This research was encouraged by the shortcomings of the currently available systems, which include low generalization between datasets, high computational complexity, and the absence of magnitude diversity, and the promising results were obtained after evaluation of a number of CNN structures and hybrid models to determine its capability to establish an effective and sustainable operation in the real-world diagnostic setting.

The dataset of histopathological images consisted of a wide range of histopathological images, and the analysis was limited to binary classification of benign and malignant tumors. A few pre-trained convolutional neural network (CNN)-based architectures, such as EfficientNetB0, EfficientNetB2, MobileNetV2, MobileNetV3Small, InceptionV3, VGG16, VGG19 have been considered and tested prior to the fine-tuning step. The findings showed that performance was much improved with fine-tuning among all the models. EfficientNet variants, especially EfficientNetB0 and EfficientNetB2 were highly improved with accuracy of more than 95% after fine-tuning.

As another way of improving performance and efficiency, hybrid/ensemble models were introduced in the study that involves combining lightweight architectures with high-performing architectures. Among them, the EfficientNetB0 + MobileNetV2 hybrid model was found to perform the most successful in general, having an accuracy of 95.91, good precision and recall, and reasonable energy consumption. This combination method was able to strike the right committee between computational efficiency and predictive reliability which is appropriate in resource crippled clinical settings.

On the whole, the results indicate that both accuracy and reduction of energy can be greatly enhanced in breast cancer classification as a result of the combination of the use of the methods of fine-tuning and hybrid architectures. The research is valuable because it presents a usable and generalizable deep learning model that can justify early breast cancer detection in a variety of environments and also lowers the computational cost in general that is commonly linked to deep learning models.

5.4 References

- [1] A. A. Balasubramanian, S. M. A. Al-Heejawi, A. Singh, A. Breggia, B. Ahmad, R. Christman, *et al.*, “Ensemble deep learning-based image classification for breast cancer subtype and invasiveness diagnosis from whole slide image histopathology,” *Cancers*, vol. 16, no. 12, p. 2222, 2024.
- [2] J. Ahmad, S. Akram, A. Jaffar, Z. Ali, S. M. Bhatti, A. Ahmad, and S. U. Rehman, “Deep learning empowered breast cancer diagnosis: Advancements in detection and classification,” *PLOS ONE*, vol. 19, no. 7, e0304757, 2024.
- [3] M. M. Rahman, M. Z. B. Jahangir, A. Rahman, M. Akter, M. A. A. Nasim, K. D. Gupta, and R. George, “Breast cancer detection and localizing the mass area using deep learning,” *Big Data and Cognitive Computing*, vol. 8, no. 7, p. 80, 2024.
- [4] F. Gurcan, “Enhancing breast cancer prediction through stacking ensemble and deep learning integration,” *PeerJ Computer Science*, vol. 11, e2461, 2025.
- [5] I. Jahan, M. E. H. Chowdhury, S. Vranic, *et al.*, “Deep learning and vision transformers-based framework for breast cancer and subtype identification,” *Neural Computing and Applications*, vol. 37, pp. 9311–9330, 2025.
- [6] B. S. Abunasser, M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, “Convolution neural network for breast cancer detection and classification using deep learning,” *Asian Pacific Journal of Cancer Prevention*, vol. 24, no. 2, pp. 531–538, 2023.
- [7] W. Arshad *et al.*, “Cancer unveiled: A deep dive into breast tumor detection using cutting-edge deep learning models,” *IEEE Access*, vol. 11, pp. 133804–133824, 2023, doi: 10.1109/ACCESS.2023.3335604.
- [8] A. Khalid, A. Mehmood, A. Alabrah, B. F. Alkhamees, F. Amin, H. AlSalman, and G. S. Choi, “Breast cancer detection and prevention using machine learning,” *Diagnostics*, vol. 13, no. 19, p. 3113, 2023.
- [9] I. Hirra *et al.*, “Breast cancer classification from histopathological images using patch-based deep learning modeling,” *IEEE Access*, vol. 9, pp. 24273–24287, 2021, doi: 10.1109/ACCESS.2021.3056516.
- [10] S. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, “A dataset for breast cancer histopathological image classification,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, 2016.

- [11] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.
- [12] D. Komura and S. Ishikawa, “Machine learning methods for histopathological image analysis,” *Computational and Structural Biotechnology Journal*, vol. 16, pp. 34–42, 2018.
- [13] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [15] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv:1409.1556, 2014.
- [16] C. Szegedy *et al.*, “Going deeper with convolutions,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE CVPR*, 2016, pp. 770–778.
- [18] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proc. International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114.
- [19] A. G. Howard *et al.*, “MobileNets: Efficient convolutional neural networks for mobile vision applications,” arXiv:1704.04861, 2017.
- [20] A. Howard *et al.*, “Searching for MobileNetV3,” in *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2019, pp. 1314–1324.
- [21] R. R. Selvaraju *et al.*, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” in *Proc. IEEE ICCV*, 2017, pp. 618–626.
- [22] D. Smilkov *et al.*, “SmoothGrad: Removing noise by adding noise,” arXiv:1706.03825, 2017.
- [23] J. Deng *et al.*, “ImageNet: A large-scale hierarchical image database,” in *Proc. IEEE CVPR*, 2009, pp. 248–255.
- [24] World Health Organization, “Breast cancer,” WHO Fact Sheet, 2023.
- [25] American Cancer Society, “Breast cancer facts & figures,” ACS Report, 2024.

221-35-972

ORIGINALITY REPORT

23%

SIMILARITY INDEX

17%

INTERNET SOURCES

18%

PUBLICATIONS

14%

STUDENT PAPERS

PRIMARY SOURCES

| | | |
|----|---|-----|
| 1 | Submitted to Midlands State University Student Paper | 2% |
| 2 | Submitted to Daffodil International University Student Paper | 2% |
| 3 | Submitted to COMSATS Institute of Information Technology Student Paper | 1% |
| 4 | Hyun Kyu Shin, Ha Young Kim, Sang Hyo Lee. "Convolutional Neural Network-Based Regression for Predicting the Chloride Ion Diffusion Coefficient of Concrete", Computers, Materials & Continua, 2022 Publication | 1% |
| 5 | Submitted to University of Lagos Student Paper | 1% |
| 6 | d197for5662m48.cloudfront.net Internet Source | 1% |
| 7 | umpir.ump.edu.my Internet Source | <1% |
| 8 | digitalcommons.fiu.edu Internet Source | <1% |
| 9 | Submitted to National College of Ireland Student Paper | <1% |
| 10 | set-science.com Internet Source | <1% |

| | | |
|----|--|-----|
| 11 | "Advanced Computing and Intelligent Technologies", Springer Science and Business Media LLC, 2026 Publication | <1% |
| 12 | Sheetal Malpani, Romy Paz, Yasamin Mirzabeigi, Seyedeh Maral Kazemi Mohammadi, Richard Gasca, Himanshu Arora. "Machine learning and pathology: a historical perspective", Elsevier BV, 2025 Publication | <1% |
| 13 | Submitted to University of Ulster Student Paper | <1% |
| 14 | www.coursehero.com Internet Source | <1% |
| 15 | Sebastian Elmes, Tapabrata Chakraborti, Mengran Fan, Holm Uhlig, Jens Rittscher. "Automated Annotator: Capturing Expert Knowledge for Free", 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021 Publication | <1% |
| 16 | Submitted to Galgotias University, Greater Noida Student Paper | <1% |
| 17 | Submitted to University of Essex Student Paper | <1% |
| 18 | pmc.ncbi.nlm.nih.gov Internet Source | <1% |
| 19 | www.researchsquare.com Internet Source | <1% |



Dashboard

Student Portal

| | | | |
|---------------|------------|-----------|-------------|
| Total Payable | Total Paid | Total Due | Total Other |
| 767,200.00 | 767,200.00 | 0.00 | 13,190.00 |