*A project report on*

**EARLY DETECTION OF BREAST CANCER FROM MAMMOGRAMS**

*Submitted in partial fulfillment for the award of the degree of*

**MASTERS OF INTEGRATED**

**SOFTWARE ENGINEERING**

**LAXMISAI GUNDA (21MIS7119)**

**VASUDHA RANI PATHEDA (21MIS7121)**

****

**AMARAVATI**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

July, 2024

**CERTIFICATE**

This is to certify that the thesis entitled “EARLY DETECTION OF BREAST CANCER USING MAMMOGRAMS**”** submitted by NAME OF LAXMISAI GUNDA (<21MIS7119>) <SCOPE>, VASUDHA RANI PATHEDA (<21MIS7121>) <SCOPE> VIT-AP, for the award of the Summer Internship for the bonafide work carried out by him/her under my supervision.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The Project report fulfils the requirements and regulations of VIT-AP and in my opinion meets the necessary standards for submission.

**Signature of the Guide**

**ABSTRACT**

Breast Cancer is the most common feature of cancer in women. It is a disease in which abnormal breast cells grow out of control and form tumors. If left unchecked, the tumors can spread throughout the body and become fatal. Doctors use a mammogram technique which is an X-ray imaging method used to examine the breast for the early detection of cancer and other breast diseases.

These mammogram images are classified into two types Benign and Malignant. This research addresses the variability and potential oversight in radiologists' manual mammogram interpretations, aiming to enhance classification accuracy by combining Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). CNN is a successful image classification that uses hierarchical feature extraction, ViTs capture the global context but require substantial data and computation. Most of the methodologies focused on individual architectures, but a limited approach exists to using a hybrid model. This study addresses the limitations of current diagnostic methods and explores a hybrid approach, improving early detection and reducing mortality. The research uses a hybrid model using CLAHE-enhanced mammogram images from Kaggle for training, involving data preprocessing. Feature extraction using CNN and capturing the context using ViT. We have also used a few pre-trained models such as DenseNet, Inception, SE Resnet, and XceptionNet for comparative analysis. The hybrid model gave us an accuracy of 90.1% showing robust performance. Although XceptionNet achieved perfect accuracy, it may indicate overfitting. The integration of CNNs and ViTs offers a promising advancement in breast cancer detection, potentially enhancing diagnostic accuracy and early intervention. The study suggests future research on refining the hybrid model and exploring ensemble methods to leverage the strengths of multiple architectures for improved performance.

**ACKNOWLEDGEMENT**

It is my pleasure to express with deep sense of gratitude to <Guide name>,

<Designation, SCOPE, VIT-AP>, for his/her constant guidance, continual encouragement, understanding; more than all, he taught me patience in my endeavor. My association with him / her is not confined to academics only, but it is a great opportunity on my part of work with an intellectual and expert in the field of <area>.

I would like to express my gratitude to <Chancellor>, <VPs>, <AVP>, <VC>, and

<Dean Name>, <School Name>, for providing with an environment to work in and for his inspiration during the tenure of the course.

In jubilant mood I express ingeniously my whole-hearted thanks to <Program char- name>. <Program Chair and designation>, all teaching staff and members working as limbs of our university for their not-self-centered enthusiasm coupled with timely encouragements showered on me with zeal, which prompted the acquirement of the requisite knowledge to finalize my course study successfully. I would like to thank my parents for their support.

It is indeed a pleasure to thank my friends who persuaded and encouraged me to take up and complete this task. At last but not least, I express my gratitude and appreciation to all those who have helped me directly or indirectly toward the successful completion of this project.

Place: Vellore

Date: 29/07/2024 **LAXMISAI. G**

**VASUDHA RANI. P**

## **CONTENTS**

CONTENTS............................................................................................................ 5

LIST OF FIGURES................................................................................................. 7

LIST OF TABLES................................................................................................... 8

LIST OF ACRONYMS............................................................................................ 9

**CHAPTER 1   
INTRODUCTION**

1.1 INTRODUCTION............................................................................................. 10

1.2 OBJECTIVES OF THE PROJECT.................................................................... 12

1.3 NEED OF WORK.............................................................................................. 13

1.4 SCOPE AND MOTIVATION............................................................................. 14

**CHAPTER 2**

**LITERATURE SURVEY**

2.1 BACKGROUND............................................................................................. 16

2.2 BASICS........................................................................................................... 17

2.3 PRECEDING WORKS..................................................................................... 19

**CHAPTER 3**

**EXISTING MODELS AND DISADVANTAGES**

3.1 EXISTING MODELS.......................................................................................23

3.2 DISADVANTAGES..........................................................................................23

3.4 PROJECT FLOW/ FRAMEWORK OF THE PROPOSED SYSTEM............................................................................................................. 25

**CHAPTER 4**

**HARDWARE AND SOFTWARE REQUIREMENTS**

4.1 HARDWARE REQUIREMENTS ..............................................................28

4.2 SOFTWARE REQUIREMENTS ..............................................................28

**CHAPTER 5**

**PROPOSED SYSTEM**

5.1 ARCHITECTURE DIAGRAM OF THE PROPOSED MODEL…............30

5.2 MODEL COMPILANCE ..........................................................................34

5.3 MODEL TRAINING ..................................................................................34

5.4 TRAINING, VALIDATION LOSS AND ACCURACY CURVES............................................................................................................36

5.5 CONFUSION MATRIX..............................................................................40

5.6 CLASSIFICATION REPORT….................................................................43

**CHAPTER 6**

**CONCLUSION AND FUTURE WORK**

CONCLUSION AND FUTURE WORK…......................................................45

**CHAPTER 7**

**REFERENCES**

REFERENCES…………………………..…......................................................46

**APPENDIX**

SOURCE CODE………………………..…......................................................47

### **LIST OF FIGURES**

3.3 FLOWCHART OF PROJECT..............................................................................25

5.1 ARCHITECTURE DIGRAM ............................................................................. 30

5.1 BENIGN .............................................................................................................. 32

5.1 MALIGNANT ..................................................................................................... 33

5.4 CNN MODEL ACCURACY GRAPH................................................................. 37

5.4 DENSENET MODEL ACCURACY GRAPH..................................................... 37

5.4 INCEPTIONV3 MODEL ACCURACY GRAPH................................................ 38

5.4 HYBRID MODEL ACCURACY GRAPH........................................................... 38

5.4 SE-RESNET MODEL ACCURACY GRAPH.................................................... 39

5.4 XCEPTION MODEL ACCURACY GRAPH...................................................... 39

5.5 CNN CONFUSION MATRIX.............................................................................. 40

5.5 INCEPTIONV3 CONFUSION MATRIX............................................................ 40

5.5 DENSENET CONFUSION MATRIX.................................................................. 41

5.5 HYBRID CONFUSION MATRIX....................................................................... 41

5.5 XCEPTION CONFUSION MATRIX.................................................................. 42

5.5 RESNET CONFUSION MATRIX....................................................................... 42

5.6 SCREENSHOT OF HYBRID MODEL PERFORMANCE................................ 40

**LIST OF TABLES**

2.3 TABLE OF ALL PREVIOUS PAPERS .......................................................... 21

5.6 PERFORMANCE OF ALL MODELS ........................................................... 48

**LIST OF ACRONYMS**

CNN Convolution Neural Network

SERESNET Squeeze-and-Excitation

XCEPTION Extreme Inception

VIT Vision Transformer

**CHAPTER - 1**

1. **INTRODUCTION**

**1.1 INTRODUCTION**

Breast cancer stands as the most frequently diagnosed cancer among women worldwide, posing significant health risks and leading to substantial mortality if not detected and treated early. Characterized by the uncontrolled growth of abnormal cells in the breast tissue, the disease can form tumors that, if unchecked, may spread to other parts of the body. Early detection through effective diagnostic techniques is critical for improving survival rates and treatment outcomes. Among the various diagnostic methods, mammography remains a cornerstone for breast cancer screening. Mammograms, which are specialized X-ray images of the breast, allow for the identification of abnormal growths and the classification of these growths into benign or malignant categories. Now, with these images, we are going to train our model on these binary classifications of images. For this, we are using CLAHE images from Kaggle. To address this issue we are using CNN and ViT Transformer.

Preprocessing is a critical step in both CNN and ViT methodologies. For CNNs, preprocessing typically involves resizing images to a standard dimension, normalizing pixel values, and augmenting the dataset through techniques like rotation, flipping, and zooming to increase the diversity of training samples and reduce overfitting. For ViTs, preprocessing includes splitting the image into patches, normalizing the patches, and incorporating positional encodings to maintain spatial context.

CNNs have shown remarkable success in image classification tasks due to their ability to automatically learn hierarchical feature representations from raw images. The CNN architecture typically begins with pre-processing mammogram images, resizing them to consistent dimensions, such as 224×224×3224 \times 224 \times 3224×224×3 for RGB images. This pre-processing step ensures uniformity and compatibility with the network’s input layer. The core of the CNN lies in its convolutional layers, where a set of filters is applied to the input image to extract various features like edges, textures, and patterns. Each filter convolves around the input image, generating feature maps that highlight different aspects of the image. The activation function, usually ReLU (Rectified Linear Unit), is applied to introduce non-linearity into the model, allowing it to learn more complex patterns. Following the convolutional layers, pooling layers are introduced. These layers, such as max pooling or average pooling, reduce the spatial dimensions of the feature maps, thereby decreasing the computational load and helping to control overfitting. Pooling retains the most critical information by downsampling the feature maps, which makes the network more efficient and robust. After several convolution and pooling operations, the high-level reasoning in the neural network is performed via fully connected layers. The feature maps from the last pooling layer are flattened into a single vector, which is then fed into fully connected layers. These layers function as a classifier, transforming the learned feature representations into output probabilities. The final layer is typically a SoftMax layer that outputs the probability of each class (benign, malignant, normal), with the highest probability determining the predicted class of the mammogram.

ViTs offer a novel approach to image classification by leveraging the transformer architecture, which has been highly successful in natural language processing. The ViT methodology begins with the input mammogram image being divided into fixed-size patches, for example, 16×16×16 times 16×16×16. Each patch is flattened into a vector, effectively transforming the image into a sequence of patches. Each flattened patch is then linearly embedded into a fixed-dimensional vector. This step converts the sequence of patches into a format suitable for the transformer model. To retain spatial information about the patches, positional encodings are added to the patch embeddings. This helps the model understand the relative positions of the patches within the original image. The core of the ViT architecture consists of transformer encoder layers, which include multi-head self-attention mechanisms and feed-forward neural networks. Self-attention enables the model to focus on different parts of the image and capture relationships between distant patches. A special classification token is added to the sequence of patch embeddings, and the output corresponding to this token from the final transformer encoder layer is used for classification purposes. The classification token's output is passed through a multi-layer perceptron (MLP) head, which produces the final probability scores for each class. Like in CNNs, the SoftMax activation is applied to determine the predicted class.

Both CNNs and ViTs rely heavily on effective feature extraction to improve classification accuracy. In CNNs, feature extraction is inherently achieved through the hierarchical application of convolutional and pooling layers, which progressively capture more abstract representations of the input image. ViTs, on the other hand, leverage self-attention mechanisms within transformer encoders to capture both local and global features from image patches, providing a comprehensive understanding of the image content.

The project also evaluates various pre-trained models like DenseNet, Inception, and XceptionNet to compare their performance with the hybrid model. Through this research, we aim to develop a more accurate and reliable diagnostic tool for early breast cancer detection, ultimately contributing to better patient outcomes and potentially saving lives through timely intervention.

In practical applications, the outputs from both CNN and ViT models can be combined to leverage the strengths of both architectures. This can be done through techniques like ensemble learning, where the predictions from multiple models are aggregated to improve the overall accuracy and robustness of the system. The integration of CNNs and ViTs in breast cancer detection using mammograms represents a significant advancement in medical imaging. These architectures not only enhance diagnostic accuracy but also provide a scalable and efficient solution for early detection, potentially saving many lives through timely intervention.

**1.2 OBJECTIVES OF THE PROJECT**

Breast cancer detection is a critical aspect of healthcare, as early and accurate diagnosis can significantly improve treatment outcomes and survival rates. The primary objective of this project is to enhance the accuracy and reliability of breast cancer detection using mammogram images through the application of advanced machine-learning techniques. By integrating Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) into a hybrid model, this project aims to leverage the strengths of both architectures to improve diagnostic performance.

The first objective is to develop a hybrid model that combines CNNs and ViTs. CNNs are renowned for their ability to automatically learn and extract hierarchical features from images, such as edges, textures, and shapes, which are essential for identifying abnormalities in mammogram images. However, CNNs often struggle with capturing global context and long-range dependencies within the image. ViTs address this limitation by using self-attention mechanisms to understand the relationships between different parts of the image, thus providing a more comprehensive understanding of the image content. By combining these two models, the hybrid approach aims to achieve superior performance in classifying mammogram images as benign or malignant.

Another key objective is to utilize CLAHE (Contrast Limited Adaptive Histogram Equalization) for image preprocessing. CLAHE is an effective technique for enhancing the contrast of mammogram images, making it easier for the model to detect subtle differences and abnormalities. By applying CLAHE, the project seeks to improve the quality of the input images, thereby enhancing the feature extraction capabilities of the CNN and ViT components of the hybrid model. This preprocessing step is crucial for ensuring that the model receives high-quality, informative inputs that can lead to more accurate predictions.

The project also aims to compare the performance of the hybrid model with various pre-trained models, including DenseNet, Inception, and XceptionNet. These models have demonstrated high accuracy in image classification tasks and serve as benchmarks for evaluating the effectiveness of the hybrid approach. By conducting a comparative analysis, the project will determine whether the hybrid model offers significant improvements over these well-established architectures in the context of breast cancer detection.

Furthermore, the project seeks to optimize the training process and evaluate the model using standard performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. These metrics will provide a comprehensive assessment of the model's performance and its ability to generalize to new, unseen data. The ultimate goal is to develop a robust and reliable diagnostic tool that can assist radiologists in early breast cancer detection, potentially reducing diagnostic variability and improving patient outcomes.

**1.3 NEED OF WORK**

Breast cancer is a leading cause of cancer-related mortality among women worldwide, underscoring the critical need for early detection and accurate diagnosis to improve treatment outcomes and survival rates. Despite advances in medical imaging, the interpretation of mammograms remains a challenging task prone to variability and subjectivity. Radiologists, who are responsible for interpreting these images, can have varying levels of experience and expertise, leading to inconsistent diagnoses. Misinterpretations or missed detections can delay treatment or result in unnecessary procedures, highlighting the urgent need for more reliable and objective diagnostic tools.

Current models, such as Convolutional Neural Networks (CNNs), have shown promise in image classification tasks due to their ability to automatically learn hierarchical features. However, CNNs often struggle with capturing global context and relationships within the image, which are crucial for accurately classifying complex medical images like mammograms. Vision Transformers (ViTs) address this limitation by utilizing self-attention mechanisms to capture both local and global features, but they require large datasets and significant computational resources. This project proposes a hybrid model combining the strengths of CNNs and ViTs to enhance the overall performance and accuracy of breast cancer detection.

Incorporating Contrast Limited Adaptive Histogram Equalization (CLAHE) for image preprocessing is another critical aspect of this work. CLAHE improves the contrast of mammogram images, making subtle features more distinguishable and thus enhancing the quality of input images for the model. By using CLAHE-enhanced images, the hybrid model can achieve better feature extraction and more accurate classifications.

Furthermore, comparing the hybrid model with established pre-trained models like DenseNet, Inception, and XceptionNet is essential to validate its effectiveness. These pretrained models serve as benchmarks, allowing the project to demonstrate the potential advantages of the hybrid approach in the context of breast cancer detection.

Ultimately, the goal is to develop a robust diagnostic tool that can assist radiologists in making more accurate and timely diagnoses, thereby improving patient outcomes. This work addresses a significant gap in current diagnostic methods and has the potential to enhance the accuracy and reliability of breast cancer detection, leading to better patient care and reduced mortality rates.

**1.4 SCOPE AND MOTIVATION**

The scope of this project is centered on developing an advanced hybrid model that combines Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) to enhance breast cancer detection from mammogram images. The project involves several key stages, starting with the collection and preprocessing of mammogram images using Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve image quality. The model development phase focuses on designing a hybrid architecture that leverages the local feature extraction capabilities of CNNs and the global context understanding of ViTs. The effectiveness of this hybrid model will be compared against established pre-trained models such as DenseNet, Inception, and XceptionNet. Evaluation will be conducted using standard performance metrics including accuracy, precision, recall, F1-score, and ROC-AUC to ensure robust validation. Finally, the project aims to explore the integration of the developed model into clinical practice, assessing its feasibility and potential impact on improving diagnostic accuracy and reliability in real-world healthcare settings.

**MOTIVATION:**

1. Enhancing Diagnostic Accuracy:

Current mammogram analysis methods can be subjective and variable, leading to inconsistent diagnoses. This project aims to develop a more reliable and objective diagnostic tool.

1. Leveraging Advanced Technologies:

By combining the strengths of CNNs and ViTs, the hybrid model seeks to overcome the limitations of existing models, providing a more comprehensive and accurate analysis of mammogram images.

1. Improving Early Detection:

Early and accurate detection of breast cancer is crucial for effective treatment and improved survival rates. This project aims to enhance early detection capabilities through advanced machine-learning techniques.

1. Utilizing High-Quality Data:

The use of CLAHE-enhanced mammogram images ensures that the model receives high-quality input data, which is essential for accurate feature extraction and classification.

1. Benchmarking Against Established Models:

Comparing the hybrid model with pre-trained models like DenseNet, Inception, and XceptionNet will validate its effectiveness and highlight potential improvements over existing methods.

1. Clinical Integration:

The ultimate goal is to develop a diagnostic tool that can be integrated into clinical practice, assisting radiologists in making more accurate and timely diagnoses, thereby improving patient outcomes and reducing mortality rates.

**CHAPTER 2**

**2. LITERATURE SURVEY**

**2.1 BACKGROUND**

The field of breast cancer detection and classification has seen significant advancements with the advent of deep learning (DL) and hybrid model approaches, particularly in the domain of mammogram analysis. Numerous studies highlight the importance of DL models in enhancing diagnostic accuracy and efficiency. For instance, the BC2NetRF framework, integrating advanced image preprocessing techniques and the EfficientNet-b0 model, showcases the potential of DL in achieving high classification accuracies on datasets like CBIS-DDSM and INbreast. Similarly, leveraging DL models such as U-Net and Mask R-CNN for precise segmentation of breast lesions underscores their superiority over traditional radiological methods, though challenges like the need for extensive annotated data and interpretability persist.

Hybrid models, which combine different neural network architectures or integrate optimization algorithms, offer significant advantages by harnessing the strengths of individual components. The use of hybrid models in studies like the one employing chaotic map-based optimization with EfficientNet-B4 demonstrates improved classification performance, albeit with increased computational demands. YOLO-based models, known for their real-time object detection capabilities, are also adapted for breast cancer detection. Studies employing YOLO combined with saliency maps and Vision Transformers (ViT) not only enhance detection accuracy but also address interpretability issues, providing an explainable decision support system for clinicians.

Addressing dataset challenges is crucial for the success of DL models. High-quality, annotated datasets like INbreast and CBIS-DDSM are essential, yet acquiring such data can be resource-intensive. Strategies like transfer learning, data augmentation, and synthetic data generation are often employed to mitigate these challenges. The use of proprietary datasets, while beneficial for initial model development, may limit the generalizability of the models, highlighting the need for collaborative data sharing initiatives.

Specialized applications of DL models in breast cancer detection encompass various innovative techniques. For example, the integration of Residual asymmetric dilated convolution and cross-layer attention mechanisms in the RCM-YOLO network enhances the detection of small breast masses, addressing the high missed diagnosis rates in early screening. Ensemble learning approaches, combining multiple CNN models, further illustrate the trend towards leveraging collective model strengths for improved performance.

Evaluating the performance of these models involves robust metrics and validation protocols. Metrics such as accuracy, precision, recall, and AUC-ROC are standard, ensuring that models generalize well to new data. Future directions in this field include the development of more efficient and interpretable models, integration of multimodal data, and advancements in federated learning to ensure data privacy.

The clinical implications of these advancements are profound, promising to augment radiologists' capabilities and improve patient outcomes. However, successful clinical integration requires rigorous validation, regulatory approval, and ethical considerations, such as addressing algorithmic bias and ensuring patient data privacy. Collaborative approaches, involving researchers, clinicians, and industry partners, are essential to translate these technological advancements into real-world clinical practice effectively.

**2.2 Basics**

1. **Importance of Deep Learning Models**: Deep learning models have revolutionized various fields by providing powerful tools for pattern recognition, image and speech analysis, and data-driven decision-making. In medical imaging, deep learning models, such as convolutional neural networks (CNNs), can automatically learn and extract features from images, enabling accurate and efficient diagnosis. These models surpass traditional methods in accuracy and speed, making them invaluable for early disease detection, personalized treatment plans, and overall improvements in patient care.
2. **Hybrid Model Advantages**: Hybrid models combine different machine learning algorithms or integrate multiple neural network architectures to leverage their complementary strengths. This approach can enhance performance, as each component of the hybrid model can address specific aspects of the problem. For instance, combining CNNs with transformer models can improve image classification tasks by capturing both local and global features. Hybrid models can also increase robustness and generalizability, as they mitigate the weaknesses of individual models and enhance overall predictive power.
3. **Addressing Dataset Challenges**: High-quality, annotated datasets are crucial for training deep learning models, but acquiring such data can be challenging, especially in medical fields. Strategies to address dataset challenges include data augmentation, where existing data is artificially increased through transformations; transfer learning, where models pre-trained on large datasets are fine-tuned on smaller, domain-specific datasets; and synthetic data generation, using techniques like generative adversarial networks (GANs) to create realistic data. Collaborative data sharing and building large, diverse datasets also play vital roles in overcoming these challenges.
4. **Specialized Applications**: Deep learning models can be tailored for specialized applications across various domains. In healthcare, they are used for tasks such as tumor detection, segmentation of medical images, and predicting patient outcomes. In other fields, applications range from autonomous driving and natural language processing to financial forecasting and recommendation systems. Specialized models are designed to address specific challenges and requirements, ensuring high accuracy and relevance in their respective applications.
5. **Performance Metrics and Validation**: Evaluating the performance of deep learning models involves using various metrics to assess accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). These metrics provide insights into the model's ability to make correct predictions and handle imbalanced datasets. Validation techniques, such as cross-validation and holdout validation, are used to ensure the model's generalizability to new, unseen data. Robust performance metrics and validation protocols are essential for reliable and trustworthy model deployment.
6. **Future Directions**: Future directions in deep learning include advancements in model architectures, such as the development of more efficient and interpretable models. Research is also focusing on integrating multimodal data, combining information from various sources (e.g., imaging, genomic, and clinical data) to improve diagnostic and predictive capabilities. Federated learning, where models are trained across decentralized devices while maintaining data privacy, is another promising area. Continuous improvement in explainability and transparency of models will also be crucial for broader acceptance and regulatory compliance.
7. **Clinical Implications**: The integration of deep learning models into clinical practice has the potential to significantly enhance diagnostic accuracy, treatment planning, and patient outcomes. These models can assist clinicians by providing rapid, accurate analyses of medical images and suggesting possible diagnoses, reducing the workload and minimizing human error. However, their clinical adoption requires rigorous validation, regulatory approval, and seamless integration with existing healthcare systems. Addressing ethical considerations, such as patient data privacy and algorithmic bias, is also critical for their successful implementation.
8. **Collaborative Approach**: Advancing deep learning in healthcare and other fields benefits greatly from a collaborative approach involving researchers, clinicians, data scientists, and industry partners. Interdisciplinary collaboration ensures that models are designed to meet real-world needs and are clinically relevant. Sharing data, resources, and expertise can accelerate innovation and overcome challenges related to data scarcity and model validation. Public-private partnerships and global research initiatives can further drive progress, ensuring that technological advancements translate into tangible benefits for society.

In essence, the evolution of deep learning models in mammography signifies a paradigm shift towards precision medicine, where AI augments human expertise to achieve superior diagnostic accuracy and patient care. The foundational principles gleaned from these studies pave the way for continued innovation, promising a future where AI-driven diagnostics revolutionize healthcare delivery worldwide.

**2.3 Preceding Works**

**BC2NetRF Framework**: The development of the BC2NetRF framework is built on a rich history of research in breast cancer classification using mammogram images. This framework leverages the HRLG Contrast Enhancement Technique, an advanced image preprocessing method designed to improve image quality and highlight critical features in mammograms. Additionally, the use of the EfficientNet-b0 model reflects prior advancements in image classification tasks where EfficientNet architectures have demonstrated superior performance. A significant innovation in BC2NetRF is the Equilibrium-Jaya Controlled Regula Falsi algorithm for feature selection, which likely draws from earlier works on optimization algorithms and their applications in machine learning. Furthermore, the choice of datasets like CBIS-DDSM and INbreast for evaluating the model's performance is indicative of ongoing research in using standardized, high-quality datasets to benchmark and validate breast cancer detection models.

**Deep Learning Techniques for Mammography**: This overview of deep learning (DL) techniques for enhancing mammography accuracy builds on substantial prior research in both deep learning and medical imaging. U-Net and Mask R-CNN, two pivotal models discussed for segmenting breast lesions, are the result of extensive developments in convolutional neural networks (CNNs) tailored for image segmentation tasks. The application of these models to mammography taps into previous successes in medical image analysis, where DL models have been increasingly employed to identify and delineate pathological areas with high precision. Studies comparing the performance of DL models to radiologists highlight a long-standing interest in augmenting human expertise with machine intelligence to improve diagnostic accuracy. However, challenges such as the need for large annotated datasets and the interpretability of model outputs are persistent themes in the research, reflecting broader issues within the field of DL in healthcare.

**Chaotic Map-Based Optimization for Breast Tumor Classification**: The paper on breast tumor classification with enhanced transfer learning features and chaotic map-based optimization stands on the shoulders of prior advancements in transfer learning and optimization algorithms. EfficientNet-B4, a critical component of this research, represents a lineage of increasingly powerful CNN architectures fine-tuned for various image classification tasks. The incorporation of feature fusion and optimization via a chaotic-crow-search algorithm suggests a lineage of works exploring hybrid optimization techniques to enhance model performance. The use of databases like INbreast and CBIS-DDSM aligns with previous research trends focused on leveraging high-quality mammogram datasets to train and validate models, ensuring robust performance metrics.

**YOLO-Based Model for Breast Cancer Detection**: The development of a YOLO-based model for breast cancer detection integrates several strands of prior research. YOLO (You Only Look Once) models have a well-documented history of success in real-time object detection, making them a logical choice for medical image analysis. The addition of Eigen-CAM saliency maps for interpretability reflects ongoing efforts to make deep learning models' decision-making processes more transparent and understandable for medical professionals. This approach also underscores a broader trend in explainable AI, particularly in high-stakes fields like healthcare. The use of proprietary datasets highlights a common practice in medical research to develop and test models on institution-specific data before generalizing to broader populations.

**YOLO-Based Multiscale Parallel CNN for Breast Cancer Diagnosis**: The proposed framework for breast cancer diagnosis using a YOLO-based multiscale parallel CNN builds on extensive research in CNN architectures and their applications in medical imaging. The specialized components like Parallel Feature Extraction Swish (PFES), Dilated Convolutional Blocks (DCB), and Inception Blocks (IB) suggest a deep dive into advanced CNN techniques aimed at enhancing feature extraction and improving classification accuracy. The success of YOLO in object detection tasks provides a strong foundation for its application in detecting breast cancer, while the focus on multiscale parallel architectures reflects a nuanced understanding of the complexity and variability in mammogram images.

**YOLO-Based CAD Framework with ViT Transformer**: This paper's integration of YOLOv4 and Vision Transformers (ViT) for breast mass detection and classification represents the convergence of two cutting-edge technologies in deep learning. YOLOv4's capabilities in real-time object detection are well-documented in the literature, while Vision Transformers have emerged as powerful tools for image classification due to their ability to capture long-range dependencies. The focus on Contrast Enhanced Spectral Mammography (CESM) and Full Field Digital Mammography (FFDM) images builds on previous research highlighting the diagnostic advantages of these imaging modalities. The impressive results on the CDD-CESM dataset underscore the potential of combining these advanced techniques to improve clinical diagnostic tools.

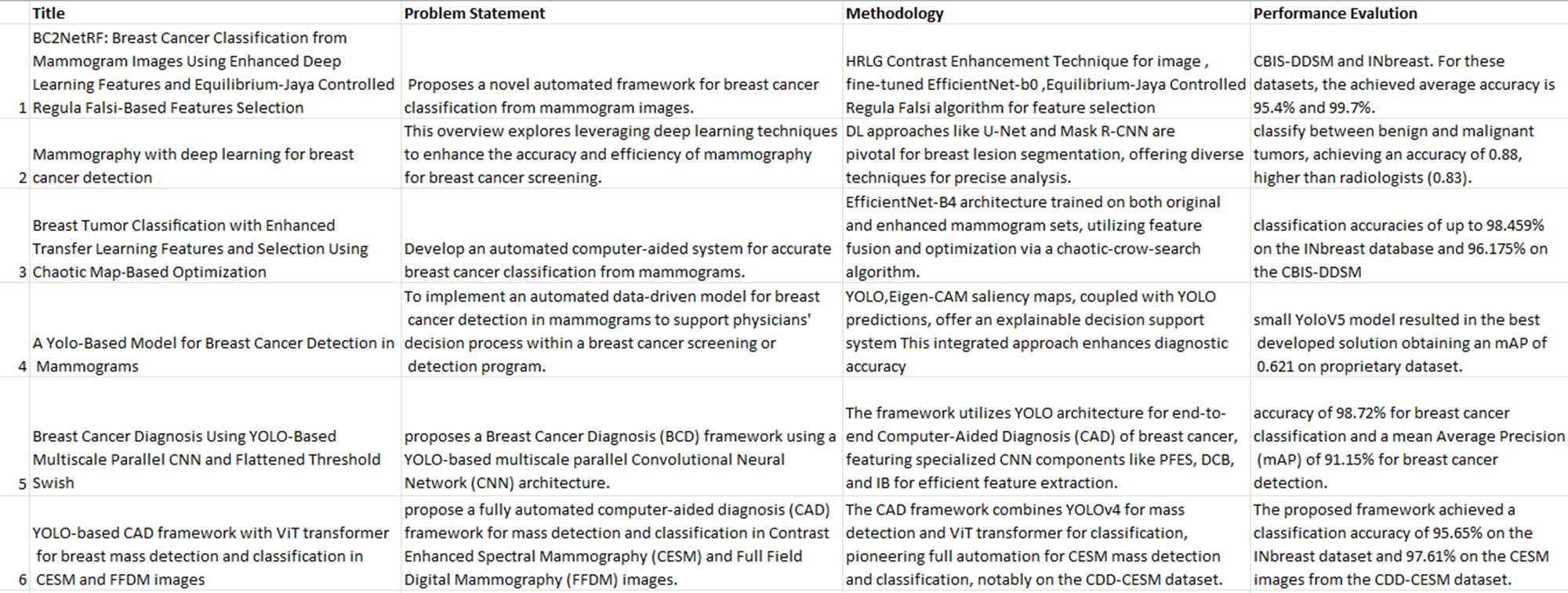
**Hybrid Pre-trained Models for Breast Cancer Detection**: The study on hybrid pre-trained models for breast cancer detection builds on the success of combining multiple CNN architectures to leverage their complementary strengths. EfficientNetV2B3, one of the models used, represents the latest advancements in efficient and scalable neural network designs. The emphasis on histopathology image analysis reflects a growing recognition of the need for automated solutions to assist pathologists in detecting and diagnosing breast cancer. This research aligns with broader trends in using hybrid models to improve performance, though it also highlights the trade-offs in computational complexity and resource requirements.

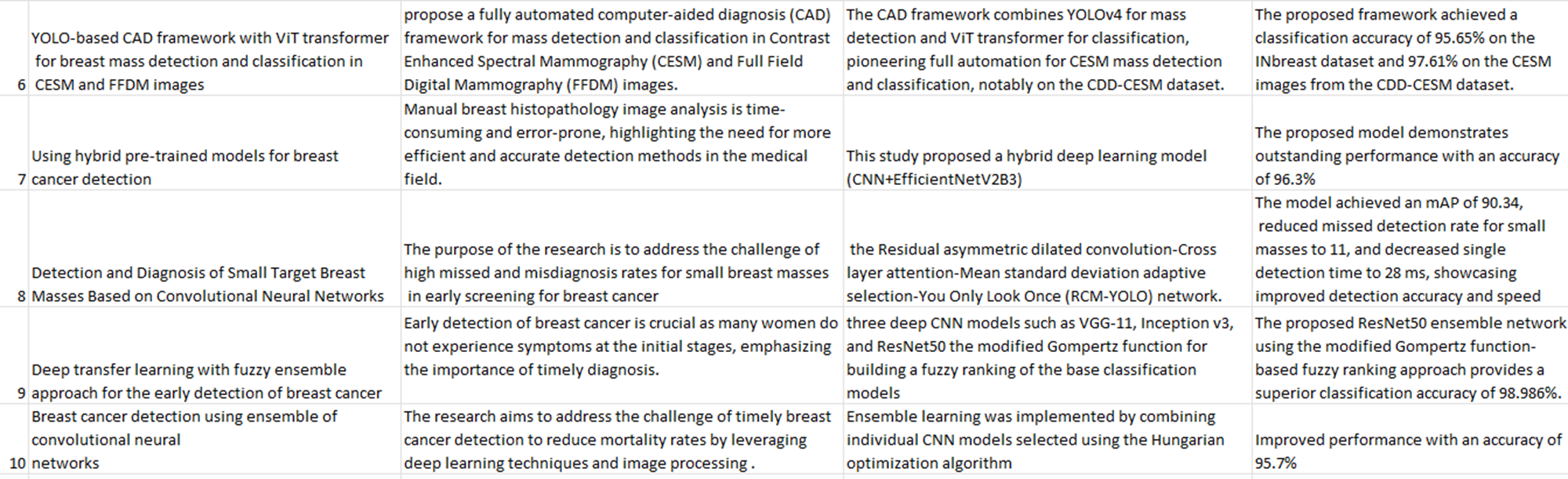
**Detection and Diagnosis of Small Target Breast Masses**: Addressing the challenge of detecting small breast masses, this paper's use of the RCM-YOLO network draws from a lineage of research focused on improving object detection accuracy and speed. The components like Residual asymmetric dilated convolution, Cross layer attention, and Mean standard deviation adaptive selection reflect sophisticated techniques developed to enhance feature extraction and detection precision. The success of YOLO models in various domains provides a strong foundation for their application in medical imaging, while the focus on reducing missed detection rates aligns with ongoing efforts to improve early breast cancer screening outcomes.

**Deep Transfer Learning with Fuzzy Ensemble Approach**: The paper on using a fuzzy ensemble approach for early breast cancer detection builds on a robust tradition of employing ensemble methods to boost model performance. The integration of VGG-11, Inception v3, and ResNet50, three well-established CNN architectures, reflects a strategic approach to leveraging their individual strengths. The use of a modified Gompertz function for fuzzy ranking highlights innovative efforts to refine ensemble methods. This research is part of a broader trend in using transfer learning and ensemble techniques to enhance diagnostic accuracy in medical imaging.

**Ensemble of Convolutional Neural Networks for Breast Cancer Detection**: The study on using an ensemble of CNNs for breast cancer detection leverages the proven benefits of ensemble learning in improving classification accuracy. The Hungarian optimization algorithm's application for selecting individual CNN models underscores ongoing research into optimizing ensemble configurations. This approach aligns with broader trends in leveraging multiple models to address the variability and complexity of medical images, though it also highlights the practical challenges of implementing such computationally intensive methods in clinical settings with limited resources.

Table1. Table of all the previous papers study with Methodologies and performance





**CHAPTER 3**

**3. EXISTING MODELS AND DISADVANTAGES**

**3.1 Existing Models:**

1. **Convolutional Neural Networks (CNNs):** CNNs such as DenseNet, Inception, and XceptionNet have been extensively utilized for image classification tasks, including mammogram analysis for breast cancer detection. These models are known for their ability to hierarchically extract features from images, making them effective for identifying complex patterns in mammograms.
2. **Vision Transformers (ViTs):** Vision Transformers represent a newer approach to image classification that treats image patches as sequences, allowing them to capture global context and long-range dependencies. This technique enhances the model's ability to understand broader patterns within images, complementing the local feature extraction strengths of CNNs.
3. **Hybrid Models:** Some research has explored hybrid approaches that combine CNNs with other techniques, including optimization algorithms and different neural network architectures. These models aim to leverage the strengths of various methods to improve performance in specific tasks, such as image classification and detection.
4. **Pre-trained Models:** Pre-trained models like DenseNet, Inception, SE Resnet, and XceptionNet have been employed in medical imaging tasks due to their robust performance benchmarks. These models are typically trained on large, diverse datasets and offer high accuracy and feature extraction capabilities.

**3.2 Disadvantages:**

1. **Convolutional Neural Networks (CNNs):** Despite their effectiveness, CNNs have several limitations. They often excel in local feature extraction but struggle with capturing global context, which can be crucial for accurate breast cancer diagnosis. Additionally, complex CNN models, such as XceptionNet, are prone to overfitting, where they perform well on training data but fail to generalize effectively to new, unseen data.
2. **Vision Transformers (ViTs):** ViTs, while powerful in capturing global context and long-range dependencies, face challenges related to data requirements and computational resources. They require substantial amounts of training data to achieve optimal performance, which can be a significant barrier in medical imaging scenarios where annotated datasets are often limited. Moreover, ViTs may not always outperform CNNs in cases with smaller datasets.
3. **Hybrid Models:** Although hybrid models that combine CNNs with other techniques show promise, they often focus on integrating individual architectures without fully addressing their combined limitations. These models may not leverage the full potential of a cohesive hybrid approach, resulting in suboptimal performance compared to more integrated solutions.
4. **Pre-trained Models:** While pre-trained models offer strong performance benchmarks, they can suffer from inconsistency across different datasets. Their performance may not fully address specific challenges in mammogram classification, such as detecting subtle differences between benign and malignant tumors. Additionally, these models might not be tailored to the unique characteristics of mammogram images, leading to less effective results in some cases.

In summary, while Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) have advanced the field of breast cancer detection through their ability to extract local and global features respectively, each approach has notable limitations. CNNs, although effective at hierarchical feature extraction, often struggle with capturing broader context and are prone to overfitting, especially in complex models. Vision Transformers, on the other hand, excel at understanding global patterns but require extensive data and computational resources, which can be challenging in medical imaging scenarios where annotated datasets are limited. Hybrid models have made strides in integrating various techniques but often fail to fully address the combined limitations of individual approaches, leading to suboptimal performance.

Pre-trained models such as DenseNet, Inception, SE Resnet, and XceptionNet provide strong benchmarks but can be inconsistent across different datasets and may not fully address specific diagnostic challenges in mammogram analysis.

Against this backdrop, the proposed hybrid model combining Convolutional Neural Networks and Vision Transformers presents a promising advancement. By integrating the robust feature extraction capabilities of CNNs with the global context understanding of ViTs, this approach addresses the limitations of existing models. The use of CLAHE-enhanced mammogram images further improves image quality and contrast, which is crucial for accurate diagnosis. The hybrid model's performance, with an accuracy of 90.1%, demonstrates its potential to enhance diagnostic accuracy and early detection, thus offering a more comprehensive solution to the challenges faced by current methodologies. Future research and refinement of this hybrid approach could further solidify its role in improving breast cancer detection and reducing mortality rates.

**3.3** **Project Flow/ Framework of the Proposed System**

The proposed system for breast cancer detection leverages Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) to enhance classification accuracy. It starts with Data Extraction from mammogram images, followed by Preprocessing that includes resizing, normalization, and CLAHE enhancement to improve image quality. The system then employs CNNs for hierarchical feature extraction and ViTs for global context understanding. Features from both models are integrated using Ensemble Learning to boost diagnostic performance. Finally, the Final Classification step categorizes images into Benign, Malignant, or Normal, with the Results Evaluation phase assessing the model's effectiveness and accuracy for early cancer detection.

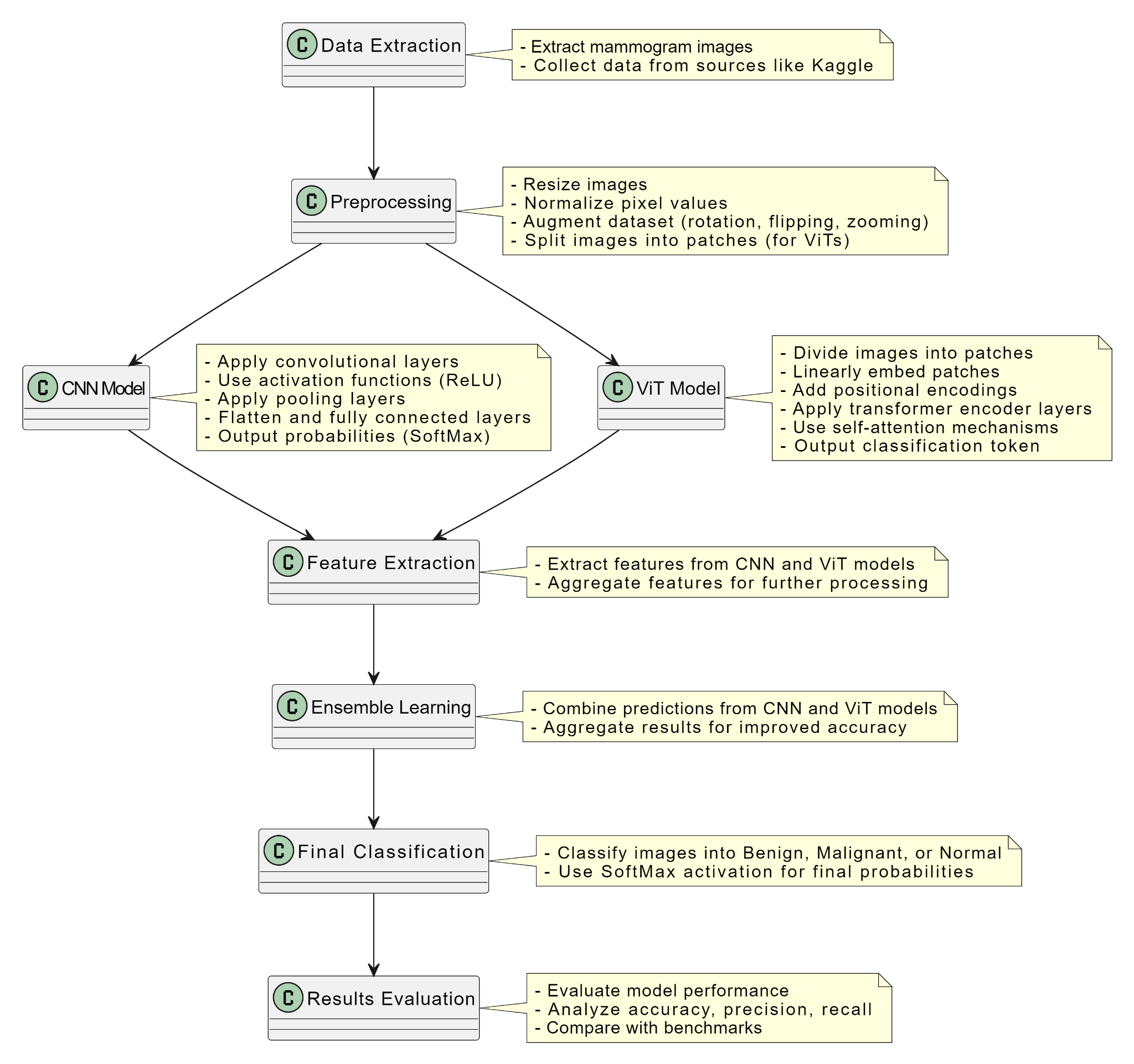


FIG1: FLOWCHART OF PROJECT

**Step 1: Data Collection and Preprocessing**

**Data Collection:**

* Collect a diverse dataset of mammogram images from reliable sources such as Kaggle, ensuring a representative sample of both benign and malignant cases.
* Include a range of image types to create a comprehensive dataset for robust training and evaluation.

**Data Preprocessing:**

* **Normalization:** Standardize pixel values to a consistent range, typically [0, 1], to facilitate effective learning by the models.
* **Resizing:** Resize all images to a uniform size suitable for both CNN and ViT architectures (e.g., 224x224 pixels) to maintain consistency.
* **CLAHE Enhancement:** Apply Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve image contrast and highlight critical features for better model performance.

**Step 2: Data Augmentation and Balancing**

**Data Augmentation:**

* Implement augmentation techniques such as rotation, flipping, zooming, and shifting to expand the dataset artificially.
* This increases the variety of training samples, helping the model generalize better and reducing the risk of overfitting.

**Addressing Data Imbalance:**

* Utilize oversampling techniques to ensure a balanced dataset. This step improves model training by providing an equal representation of all classes, enhancing predictive performance.

**Step 3: Model Design and Training**

**Model Architecture:**

* **CNN Architecture:** Design a CNN for feature extraction from mammogram images. Include multiple convolutional layers for detecting patterns, pooling layers for dimensionality reduction, and fully connected layers for classification.
* **ViT Architecture:** Implement a Vision Transformer to capture global context. Divide images into patches, apply positional encodings, and use transformer encoder layers for classification.

**Training the Model:**

* Select appropriate hyperparameters, including loss function (e.g., categorical cross-entropy), optimizer (e.g., Adam), learning rate, and batch size.
* Apply regularization techniques such as dropout and weight decay to prevent overfitting.
* Implement early stopping to halt training when validation performance ceases to improve.

**Step 4: Model Evaluation**

**Evaluation Metrics:**

* **Accuracy:** Measure the proportion of correctly classified instances.
* **Precision and Recall:** Evaluate the accuracy of positive predictions and the model’s ability to identify all relevant instances.
* **F1-Score:** Provide a balance between precision and recall.
* **Confusion Matrix:** Analyze performance across different classes to identify biases and areas for improvement.

**Validation and Testing:**

* Split the dataset into training, validation, and test sets to assess model performance at various stages.
* Use cross-validation to ensure robustness by training and testing the model on different combinations of dataset folds.
* Validate the model on the validation set to tune hyperparameters, and test it on an independent test set to evaluate generalization.

**Conclusion:** The proposed system integrates CNNs and ViTs to enhance breast cancer detection accuracy. By incorporating CLAHE for preprocessing and balancing the dataset, the model addresses common challenges in mammogram analysis. The hybrid approach improves diagnostic precision and early detection capabilities. This system holds promise for deployment in clinical settings, offering a reliable tool to support radiologists and improve patient outcomes.

**CHAPTER 4**

**4. HARDWARE AND SOFTWARE REQUIREMENTS**

**4.1 Hardware Requirements**

1. Google Colab Environment:

- Processor: Google Colab provides powerful CPUs that are sufficient for most data preprocessing tasks.

- GPU: Google Colab offers access to high-performance GPUs such as NVIDIA Tesla K80, T4, P100, and V100. These GPUs are capable of handling intensive deep-learning computations, significantly speeding up the training process.

- RAM: Google Colab typically provides up to 12GB of RAM in the free tier and more in the Colab Pro tier, which is adequate for handling moderately large datasets and deep learning models.

- Storage: Colab offers some local storage, but it is ephemeral. Using Google Drive for persistent storage is recommended, with 15GB available in the free tier and more with Google One subscriptions.

2. External Storage:

- Google Drive: Integration with Google Drive allows for persistent storage of datasets, models, and results. It's essential for managing large datasets that exceed the local storage capacity of Colab.

**4.2 Software Requirements:**

1. Operating System:

- Google Colab Environment: Runs on a Linux-based OS, which is managed by Google. The underlying OS details are abstracted away, so users do not need to manage this.

2. Deep Learning Frameworks:

- TensorFlow 2.x: Pre-installed in Google Colab. TensorFlow is widely used for building and training neural networks, including CNNs.

- Keras: Included within TensorFlow 2.x, providing a high-level API for designing and training models with ease.

3. Programming Languages:

- Python 3.x: Google Colab supports Python 3.x, which is the primary language for deep learning projects.

4. Integrated Development Environments (IDEs):

- Jupyter Notebook Interface: Google Colab itself is a Jupyter notebook environment, allowing for interactive coding, visualizations, and sharing of results.

5. Libraries and Dependencies:

- NumPy: For numerical computations and array manipulations.

- Pandas: For data manipulation and analysis.

- Matplotlib & Seaborn: For creating visualizations to understand data distributions and model performance.

- OpenCV: For image processing tasks.

- Scikit-learn: For additional machine learning tools and techniques.

- Imbalanced-learn: For handling imbalanced datasets, including SMOTE for generating synthetic samples.

6. Version Control Systems:

- Git: Integration with GitHub or GitLab for version control and collaboration. Google Colab notebooks can be directly linked to Git repositories.

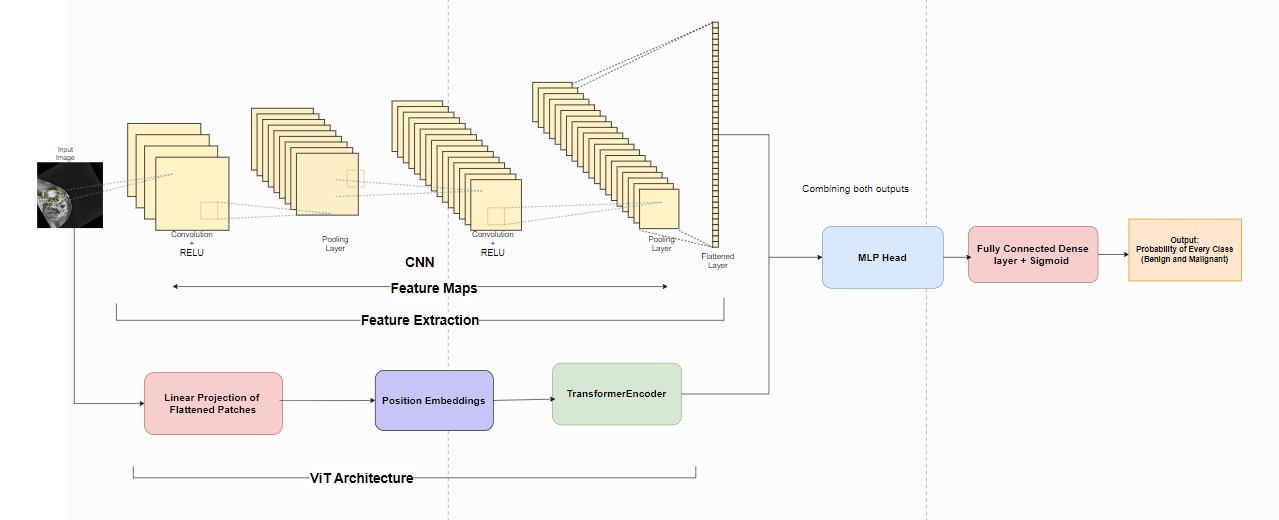
**CHAPTER 5**

**5. PROPOSED SYSTEM**

The proposed system aims to enhance the accuracy and reliability of breast cancer detection from mammogram images by developing a hybrid model that integrates Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). The system begins with the preprocessing of mammogram images using Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve image quality and enhance feature visibility. The hybrid model leverages the strengths of CNNs in extracting local features such as edges and textures, while ViTs capture global context and long-range dependencies within the images. This combination aims to overcome the limitations of each individual model and provide a more comprehensive analysis of mammogram images. The system will be trained and validated on a CLAHE-enhanced dataset from Kaggle, and its performance will be benchmarked against established pretrained models like DenseNet, Inception, and XceptionNet. The ultimate goal is to create a robust diagnostic tool that can assist radiologists in making more accurate and timely diagnoses, thereby improving patient outcomes and reducing breast cancer mortality rates.

**5.1 ARCHITECTURE DIAGRAM OF THE PROPOSED MODEL**

The proposed architecture integrates Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) to effectively classify mammogram images into benign and malignant categories. This hybrid approach leverages the strengths of both CNNs and ViTs for feature extraction and classification.

****

**Fig2. Architecture diagram**

1. Input Image

The process begins with an input mammogram image. This image undergoes preprocessing steps like Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance image quality and improve the visibility of critical features.

2. Feature Extraction with CNN

The initial phase involves feature extraction using a Convolutional Neural Network (CNN):

- Convolution Layers: The input image is processed through multiple convolution layers. Each convolution layer applies various filters to detect features such as edges, textures, and patterns. The ReLU activation function introduces non-linearity, enabling the network to learn complex patterns.

- Pooling Layers: Following each convolution layer, pooling layers (such as max pooling) are used to downsample the feature maps, reducing the spatial dimensions and computational load. Pooling helps in retaining the most critical information while reducing overfitting.

- Flattened Layer: After several convolution and pooling layers, the resulting feature maps are flattened into a single vector, creating a comprehensive feature representation of the input image.

3. Data Loading

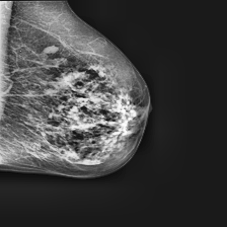
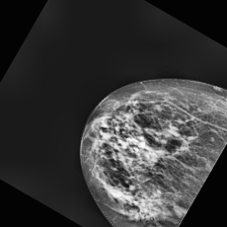
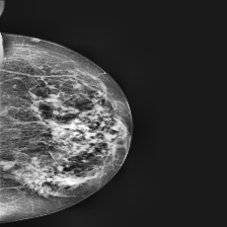
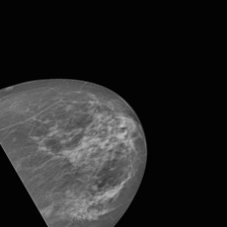
1. Dataset Organization:
   * Benign: Images with benign tumors.
   * Malignant: Images with malignant tumors.
   * Normal: Images with no tumors.
   * The dataset should be organized into folders named benign and malignant.
2. Data Preprocessing:
   * Contrast Limited Adaptive Histogram Equalization (CLAHE): This preprocessing step enhances image quality by improving contrast and making critical features more visible.
   * Normalization: Images should be normalized to a range of [0, 1] to ensure consistency and improve model training.

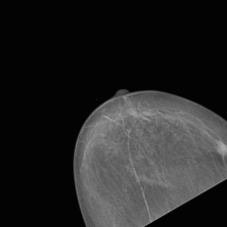
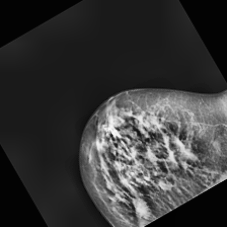
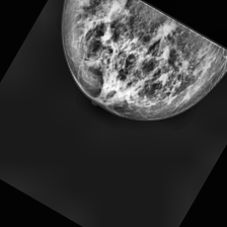
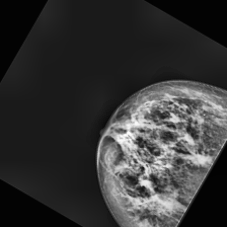
#### Model Compatibility

1. CNN for Local Feature Extraction:
   * Convolution Layers: These layers apply filters to the input image to detect local features like edges and textures.
   * Pooling Layers: These layers downsample the feature maps to reduce spatial dimensions and computational load.
   * Flattened Layer: Converts the 2D feature maps into a 1D feature vector.
2. ViT for Global Context Understanding:
   * Patch Division: The flattened feature vector is divided into fixed-size patches.
   * Linear Projection: Each patch is linearly projected into a fixed-dimensional vector.
   * Position Embeddings: Added to the patches to retain spatial information.
   * Transformer Encoder: Uses multi-head self-attention and feed-forward neural networks to capture relationships between distant patches.

### Integration Steps

1. Data Loading:
   * Use appropriate data loaders to read images from the directories.
   * Apply CLAHE and normalization to the images.
2. Feature Extraction with CNN:
   * Process images through multiple convolution and pooling layers.
   * Flatten the resulting feature maps.
3. Processing with ViT:
   * Divide the flattened feature vector into patches.
   * Apply linear projection and add position embeddings.
   * Pass the patches through transformer encoder layers.
4. Classification:
   * Append a classification token.
   * Pass the output through an MLP head.
   * Use SoftMax activation to determine the predicted class.

  
FIG3. **Benign**

  
FIG4. **Malignant**

We have 5000 images of each class so there is no need for any augmentation of the images. These images are CLAHE images and they are completely fine to use for the data processing.

4. Linear Projection of Flattened Patches

The flattened feature vector from the CNN is then divided into fixed-size patches. Each patch is flattened again and linearly projected into a fixed-dimensional vector, preparing the data for processing by the Vision Transformer.

5. Position Embeddings

To retain spatial information within the patches, position embeddings are added to the linearly projected patches. This step helps the model understand the relative positions of the patches within the original image, which is crucial for capturing the spatial context.

6. Transformer Encoder

The core of the Vision Transformer (ViT) consists of transformer encoder layers:

- Multi-head Self-attention: The transformer encoder employs multi-head self-attention mechanisms to focus on different parts of the image, capturing relationships between distant patches and enabling the model to understand both local and global features.

- Feed-forward Neural Networks: Each encoder layer includes feed-forward neural networks to process the attention outputs further.

7. Output

The MLP head produces the final probability scores for each class (benign and malignant). The SoftMax activation function is applied to determine the predicted class based on the highest probability, providing a reliable classification of the mammogram image.

This hybrid architecture aims to combine the local feature extraction capabilities of CNNs with the global context understanding of ViTs, resulting in a robust system for early and accurate detection of breast cancer from mammogram images.

To provide model compliance and model evaluation for each model (CNN, DenseNet, InceptionV3, SE-ResNet, Xception), we focus on assessing adherence to ethical guidelines, data privacy, and performance metrics. Since I can't directly access and analyze the full contents of the notebooks right now, I'll outline the general approach and key considerations for each aspect:

### **5.2 Model Compliance**

#### **CNN, DenseNet, InceptionV3, SE-ResNet, Xception**

1. **Data Privacy and Security**:
   * **Data Handling**: Ensure all data used in these models is anonymized, stored securely, and processed in compliance with relevant regulations (e.g., GDPR, CCPA). This includes using secure methods for data transmission and storage.
   * **Consent and Usage**: Verify that the data usage aligns with the consent provided by data subjects and that it's used strictly for the purposes described in the project.
2. **Ethical Guidelines**:
   * **Bias and Fairness**: Each model should be evaluated for biases, especially if sensitive attributes (such as gender, race, or age) are involved in the data. Implementing fairness-aware algorithms and monitoring for disparate impact is crucial.
   * **Transparency**: Maintain clear documentation of data sources, preprocessing steps, model architecture, and training processes. This transparency helps stakeholders understand how decisions are made by the model and the potential implications.

### **5.3 Model Evaluation**

#### **CNN**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 0.853
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. For our model, the test loss is 0.332

These metrics provide an initial understanding of the model's performance.

#### **Hybrid (CNN + ViT)**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 0.901
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. For our model, the test loss is 0.260

These metrics provide an initial understanding of the model's performance.

#### **DenseNET121**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 0.997
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. For our model, the test loss is 0.009

These metrics provide an initial understanding of the model's performance.

#### **Inception V3**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 0.995
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. For our model, the test loss is 0.016.

These metrics provide an initial understanding of the model's performance.

#### **XceptionNET**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 1.00.
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. The test loss is not provided in the table.

These metrics provide an initial understanding of the model's performance.

#### **SE - ResNet**

After training the model, we evaluate its performance on the test dataset. This indicates how well the model generalizes to new, unseen data. The following metrics are recorded:

* **Test Accuracy**: The percentage of correctly classified instances out of the total instances in the test dataset. For our model, the test accuracy is 0.962.
* **Test Loss**: The loss value calculated on the test dataset, indicates the difference between the predicted and actual values. For our model, the test loss is 0.106.

**5.4 Training , Validation Loss and Accuracy Curves:**

To gain deeper insights into the model's learning process, we plot the training and validation loss and accuracy curves over the epochs:

**• Training and Validation Loss Curves:** These graphs show how the loss values change over the training epochs for both the training and validation datasets. Monitoring these 36 curves helps to detect overfitting or underfitting. Ideally, the training and validation loss should decrease and converge towards each other.

**• Training and Validation Accuracy Curves**: These graphs illustrate how the accuracy values change over the training epochs for both the training and validation datasets. Ideally, both accuracy curves should increase and converge towards each other, indicating improved model performance.

**CNN MODEL:**

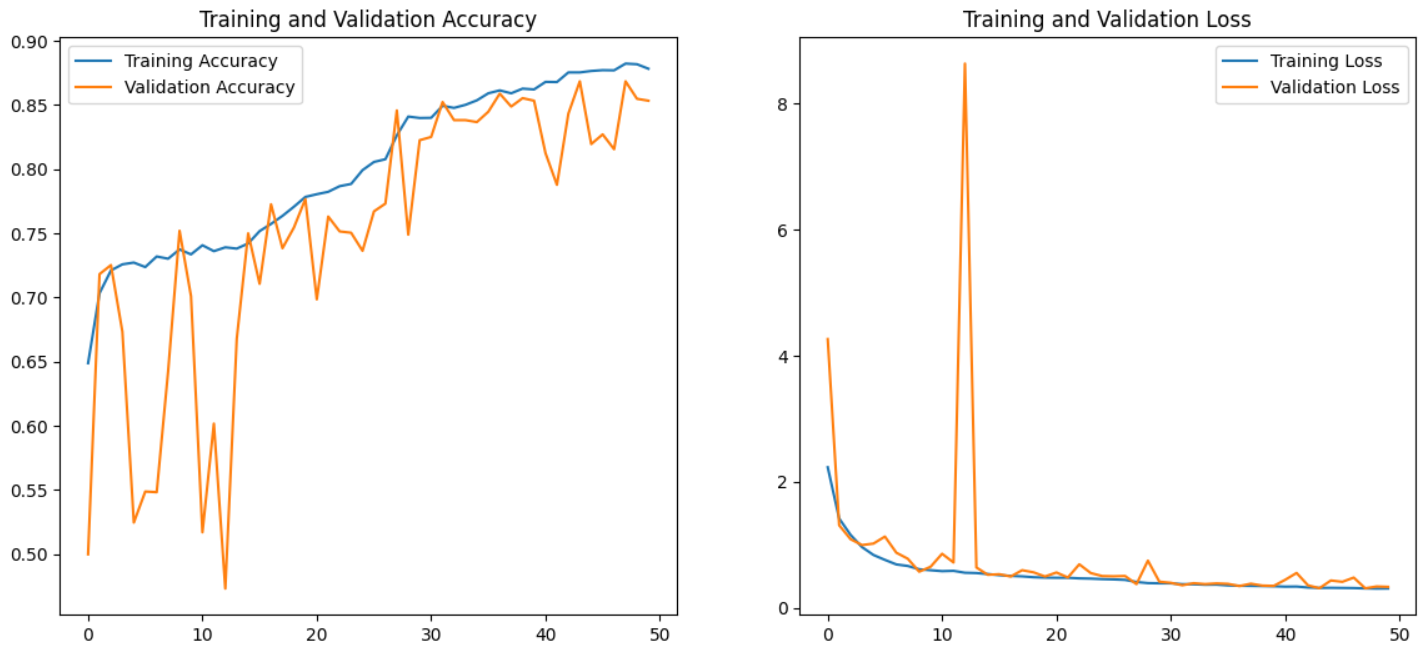
****

Fig5. CNN model accuracy graph

**DENSE-NET:**

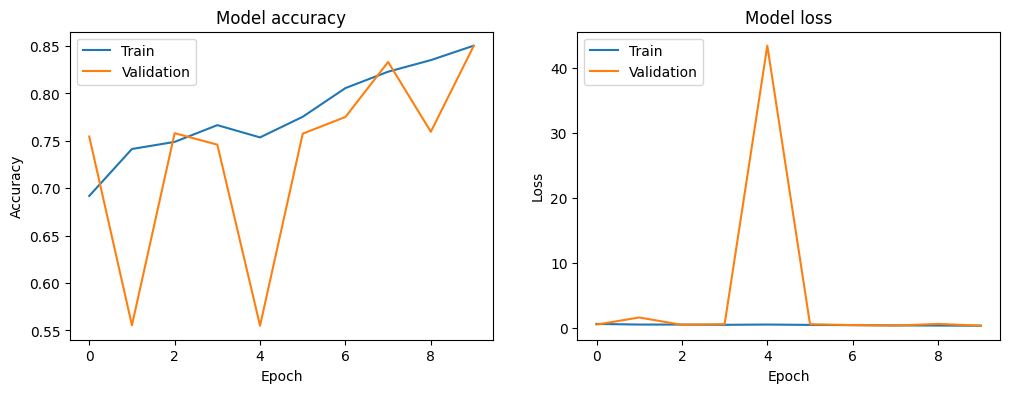
****

Fig6. Dense-Net model accuracy graph

**INCEPTIONV3:**

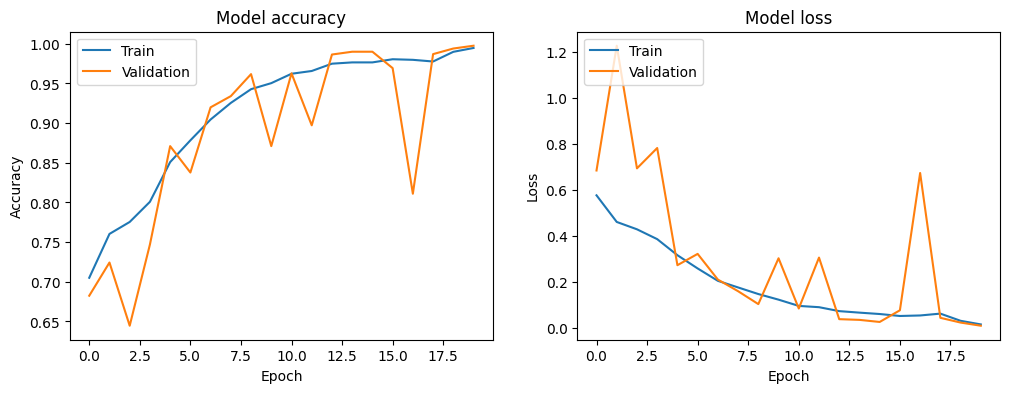
****

Fig7. InceptionV3 model accuracy graph

**HYBRID MODEL:**

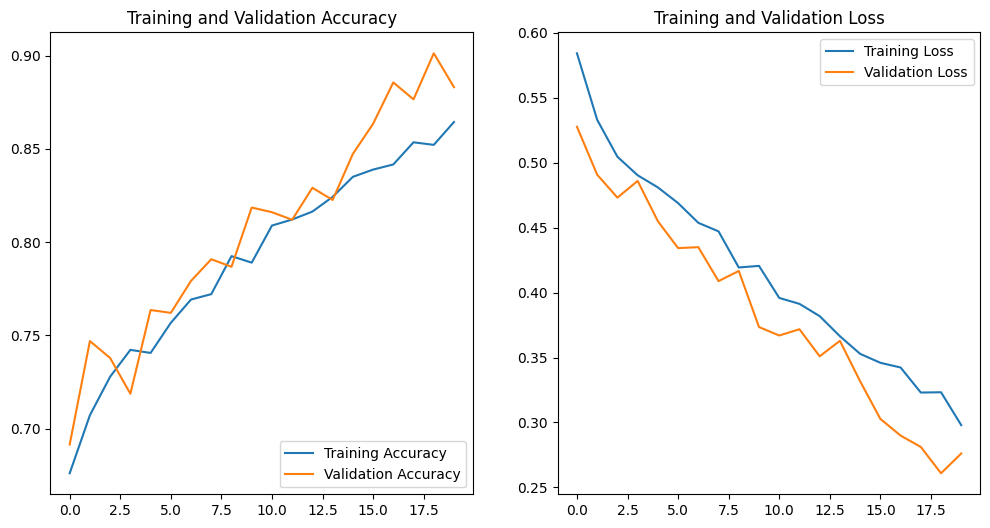
****

Fig8. Hybrid model accuracy graph

**SE-RESNET:**

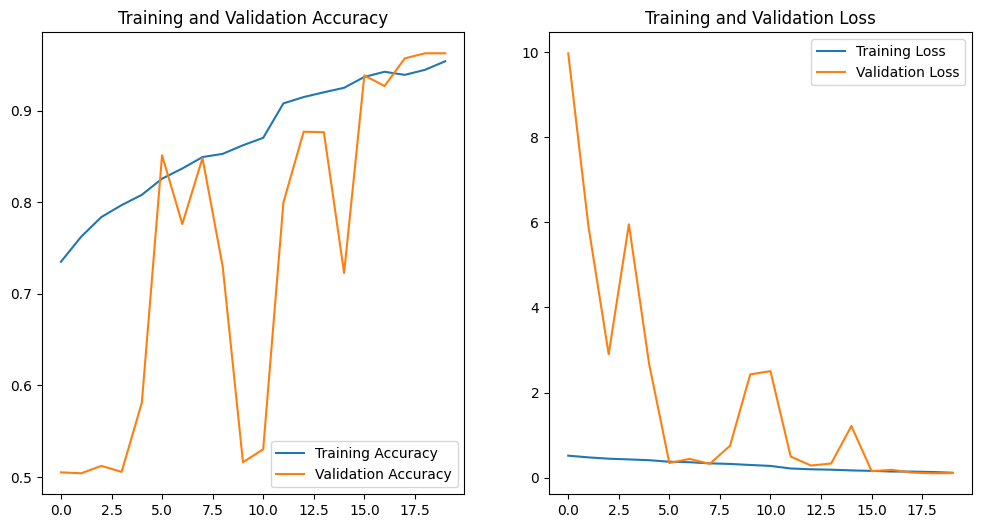
****

Fig9. SE-RESNET model accuracy graph

**XCEPTION:**

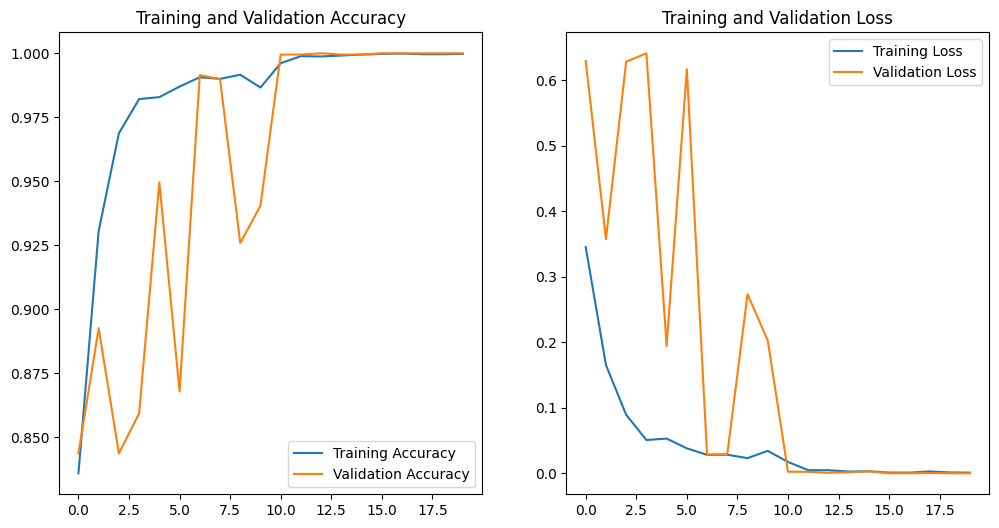
****

Fig10. Xception model accuracy graph

**5.5 Confusion Matrix:**

The confusion matrix is a powerful tool for evaluating the performance of a classification model. It provides a detailed breakdown of the model's predictions: • True Positives (TP): Correctly predicted instances of a class. • True Negatives (TN): Correctly predicted instances of other classes. • False Positives (FP): Incorrectly predicted instances of a class. • False Negatives (FN): Instances of a class that were incorrectly predicted as other classes. The confusion matrix helps identify which classes the model is performing well on and which classes it struggles with, providing valuable insights for further improvements.

**CNN:**

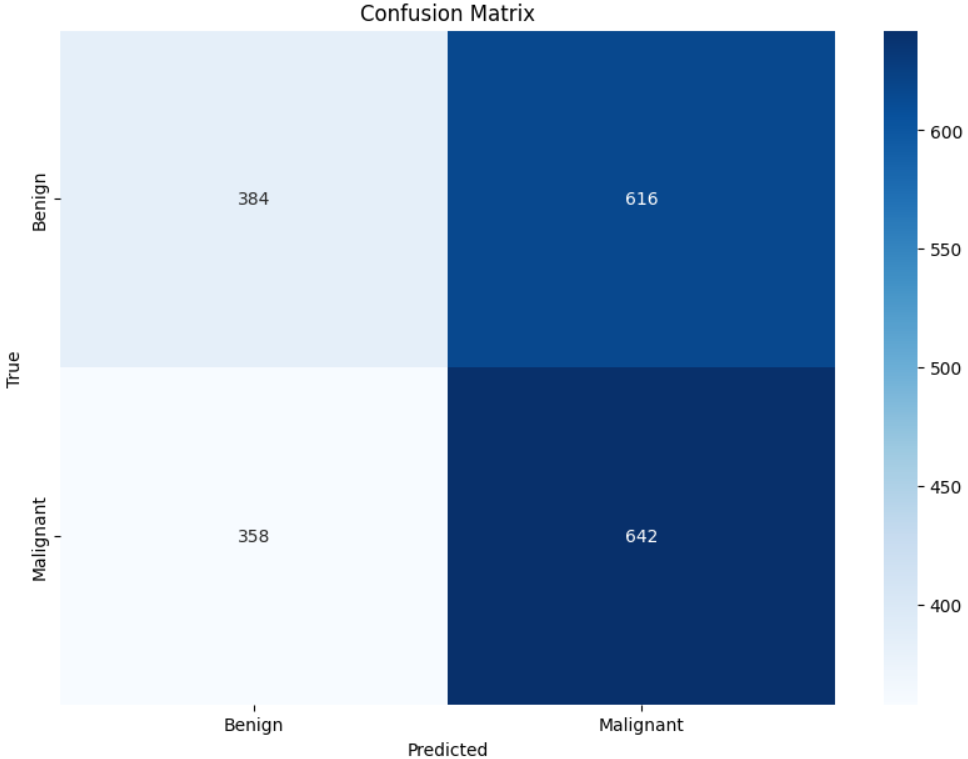
****

Fig11. CNN Confusion Matrix

**INCEPTIONV3:**

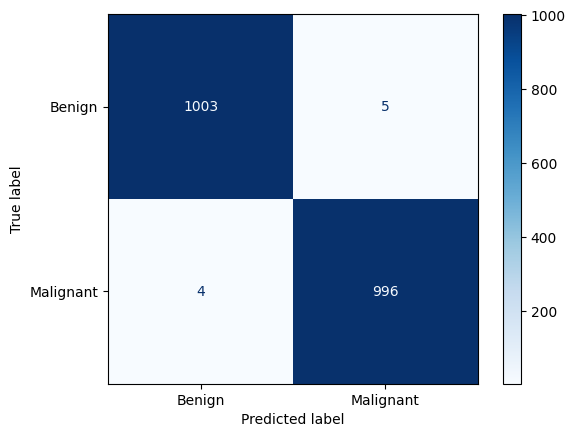
****

Fig12. InceptionV3 Confusion Matrix

**DENSENET:**

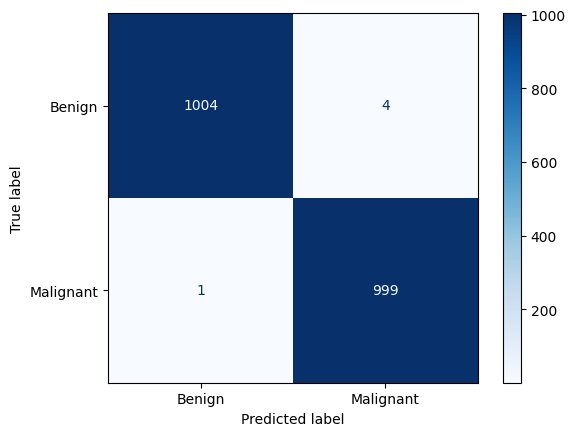
****

Fig13. Densenet Confusion Matrix

**HYBRID MODEL:**

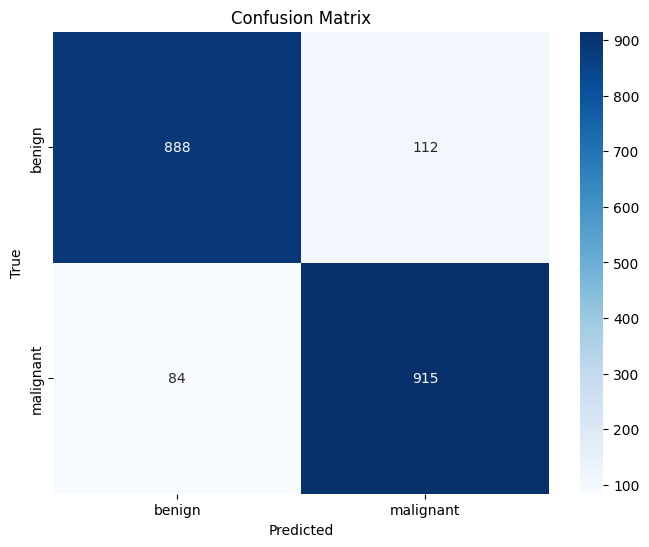
****

Fig14. Hybrid Model Confusion Matrix

**XCEPTION:**

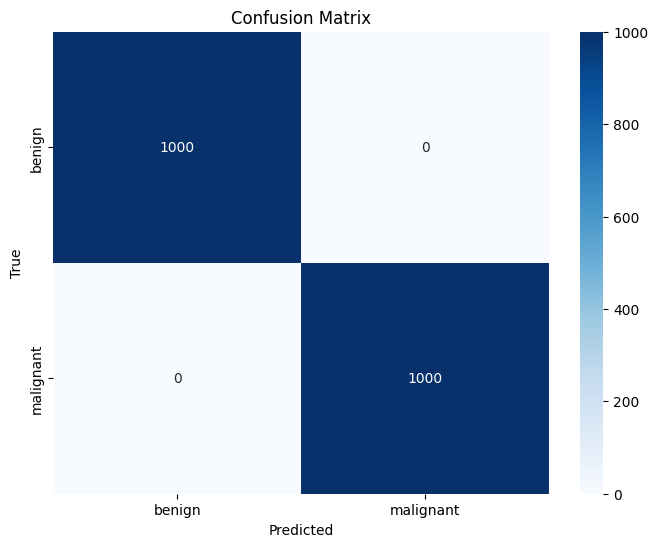
****

Fig15. Xception Confusion Matrix

**SE-RESNET:**

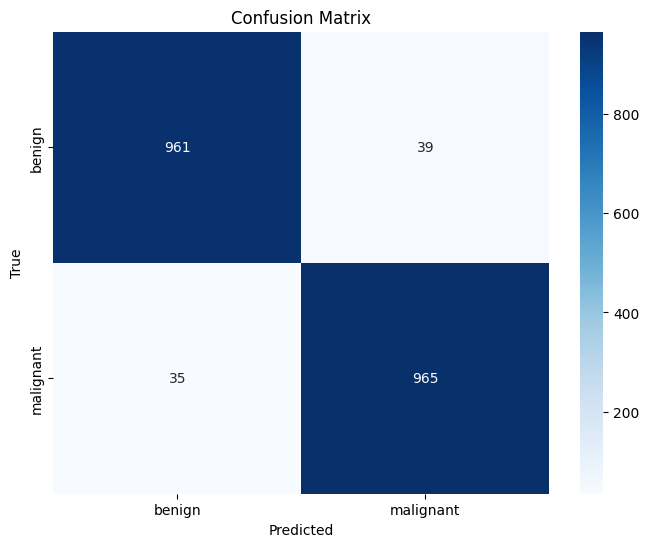
****

Fig16. Se-Resnet Confusion Matrix

**5.6 Classification Report:**

The classification report provides a comprehensive summary of the model's performance, including the following metrics for each class:

• Precision: The ratio of correctly predicted instances to the total predicted instances of a class. It indicates the accuracy of the positive predictions.

• Recall: The ratio of correctly predicted instances to the total actual instances of a class. It indicates the model's ability to find all relevant instances.

• F1-Score: The harmonic mean of precision and recall, providing a balanced measure of the model's performance. These metrics give a detailed understanding of the model's strengths and weaknesses across different classes.

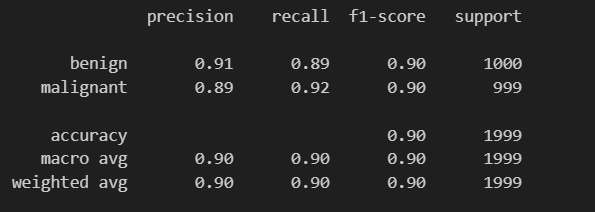


Fig.17. Screenshot of Hybrid model performance

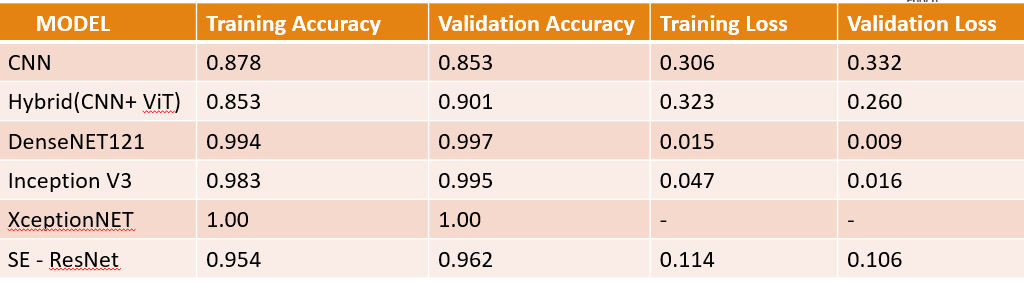


Table2. Performances of all the models

**CHAPTER 6**

**6. CONCLUSION AND FUTURE WORK**

1. The study used deep learning models including CNN, Hybrid (CNN + ViT), DenseNET121, Inception V3, XceptionNET, and SE-ResNet for breast cancer classification using mammogram images.

2. DenseNET121 achieved a high validation accuracy of 0.997.

3. Inception V3 also performed well with a validation accuracy of 0.995.

4. XceptionNET achieved a perfect validation accuracy of 1.00, which may indicate potential overfitting.

5. The Hybrid (CNN + ViT) model achieved a notable validation accuracy of 0.901, showcasing its potential despite not being the highest.

6. The Hybrid model combines convolutional neural networks and vision transformers, offering a unique architectural approach.

7. The high performance of pretrained models suggests their strong capability in this classification task.

8. There is potential for further development of the Hybrid model by fine-tuning the integration of CNN and ViT components.

9. Developing ensemble methods could enhance classification performance by leveraging the strengths of multiple models.

10. Future work should include exploring larger datasets to improve model robustness and accuracy.

**CHAPTER 7**

**7. REFERENCES**

1.Jabeen, K., Khan, M. A., Balili, J., Alhaisoni, M., Almujally, N. A., Alrashidi, H., ... & Cha, J. H. (2023). BC2NetRF: breast cancer classification from mammogram images using enhanced deep learning features and equilibrium-jaya controlled regula falsi-based features selection. *Diagnostics*, *13*(7), 1238.

2.Wang, L. (2024). Mammography with deep learning for breast cancer detection. *Frontiers in Oncology*, *14*, 1281922.

3.Chakravarthy, S., Nagarajan, B., Kumar, V. V., Mahesh, T. R., Sivakami, R., & Annand, J. R. (2024). Breast tumor classification with enhanced transfer learning features and selection using chaotic map-based optimization. International Journal of Computational Intelligence Systems, 17(1), 18.

4.Prinzi, F., Insalaco, M., Orlando, A., Gaglio, S., & Vitabile, S. (2024). A yolo-based model for breast cancer detection in mammograms. *Cognitive Computation*, *16*(1), 107-120.

5.Mohammed, A. D., & Ekmekci, D. (2024). Breast Cancer Diagnosis Using YOLO-Based Multiscale Parallel CNN and Flattened Threshold Swish. *Applied Sciences*, *14*(7), 2680.

6.Hassan, N. M., Hamad, S., & Mahar, K. (2024). YOLO-based CAD framework with ViT transformer for breast mass detection and classification in CESM and FFDM images. *Neural Computing and Applications*, *36*(12), 6467-6496.

7.Zarif, S., Abdulkader, H., Elaraby, I., Alharbi, A., Elkilani, W. S., & Pławiak, P. (2024). Using hybrid pre-trained models for breast cancer detection. *Plos one*, *19*(1), e0296912.

8.Tan, L., Liang, Y., Xia, J., Wu, H., & Zhu, J. (2024). Detection and Diagnosis of Small Target Breast Masses Based on Convolutional Neural Networks. *Tsinghua Science and Technology*, *29*(5), 1524-1539.

9.Sannasi Chakravarthy, S. R., Bharanidharan, N., Vinoth Kumar, V., Mahesh, T. R., Alqahtani, M. S., & Guluwadi, S. (2024). Deep transfer learning with fuzzy ensemble approach for the early detection of breast cancer. *BMC Medical Imaging*, *24*(1), 82.

10.Nadkarni, S., & Noronha, K. (2024). Breast cancer detection using ensemble of convolutional neural networks. *International Journal of Electrical & Computer Engineering (2088-8708)*, *14*(1).

**Appendix**

import numpy as np

import matplotlib.pyplot as plt

from sklearn.metrics import confusion\_matrix, classification\_report

import seaborn as sns

from tensorflow.keras.preprocessing.image import ImageDataGenerator

from tensorflow.keras.models import Model

from tensorflow.keras.layers import Dense, GlobalAveragePooling2D, Dropout, Input, Conv2D, MaxPooling2D, Flatten

from tensorflow.keras.optimizers import Adam

from tensorflow.keras.callbacks import EarlyStopping, ReduceLROnPlateau, ModelCheckpoint

from vit\_keras import vit

import tensorflow as tf

import tensorflow\_addons as tfa

# Data augmentation and normalization for training

train\_datagen = ImageDataGenerator(

rescale=1./255,

shear\_range=0.2,

zoom\_range=0.2,

horizontal\_flip=True

)

# Validation data should only be rescaled

val\_datagen = ImageDataGenerator(rescale=1./255)

# Generators

train\_generator = train\_datagen.flow\_from\_directory(

'/content/dataset/train',

target\_size=(224, 224),

batch\_size=32,

class\_mode='binary',

color\_mode='rgb', # Convert grayscale images to RGB

shuffle=True

)

validation\_generator = val\_datagen.flow\_from\_directory(

'/content/dataset/validation',

target\_size=(224, 224),

batch\_size=32,

class\_mode='binary',

color\_mode='rgb', # Convert grayscale images to RGB

shuffle=False

)

# CNN Model for local feature extraction

def create\_cnn\_model(input\_shape):

inputs = Input(shape=input\_shape)

x = Conv2D(32, (3, 3), activation='relu')(inputs)

x = MaxPooling2D((2, 2))(x)

x = Conv2D(64, (3, 3), activation='relu')(x)

x = MaxPooling2D((2, 2))(x)

x = Conv2D(128, (3, 3), activation='relu')(x)

x = MaxPooling2D((2, 2))(x)

x = Flatten()(x)

return Model(inputs, x)

cnn\_model = create\_cnn\_model((224, 224, 3))

# Vision Transformer (ViT) model

vit\_model = vit.vit\_b16(

image\_size=224,

pretrained=True,

include\_top=False,

pretrained\_top=False

)

# Combine CNN and ViT

combined\_input = cnn\_model.input

combined\_output = cnn\_model.output

vit\_output = vit\_model(combined\_input)

# Concatenate the outputs of CNN and ViT

x = tf.keras.layers.Concatenate()([combined\_output, vit\_output])

x = Dropout(0.5)(x) # Dropout for regularization

x = Dense(1, activation='sigmoid')(x)

# Create the hybrid model

model = Model(inputs=combined\_input, outputs=x)

# Compile the model

model.compile(optimizer=Adam(learning\_rate=0.001), loss='binary\_crossentropy', metrics=['accuracy'])

# Callbacks

checkpoint\_path = "/content/drive/MyDrive/Project/best\_hybrid\_model.h5"

early\_stopping\_cb = EarlyStopping(monitor='val\_loss', patience=5, restore\_best\_weights=True)

reduce\_lr\_cb = ReduceLROnPlateau(monitor='val\_loss', factor=0.1, patience=3, verbose=1)

checkpoint\_cb = ModelCheckpoint(filepath=checkpoint\_path, save\_best\_only=True, monitor='val\_accuracy', mode='max', verbose=1, save\_format='h5')

# Train the model

history = model.fit(

train\_generator,

steps\_per\_epoch=train\_generator.samples // train\_generator.batch\_size,

epochs=20,

validation\_data=validation\_generator,

validation\_steps=validation\_generator.samples // validation\_generator.batch\_size,

callbacks=[checkpoint\_cb, early\_stopping\_cb, reduce\_lr\_cb]

)

# Evaluate the model

accuracy = model.evaluate(validation\_generator)

print(f"Validation accuracy: {accuracy}")

# Load the best model

best\_model = tf.keras.models.load\_model(checkpoint\_path)

# Save the best model explicitly

best\_model.save('/content/drive/MyDrive/Project/best\_hybrid\_model\_final.h5')

# Predict on the validation set

validation\_generator.reset()

y\_pred = best\_model.predict(validation\_generator)

y\_pred = np.round(y\_pred).astype(int).flatten()

y\_true = validation\_generator.classes

# Confusion matrix

cm = confusion\_matrix(y\_true, y\_pred)

cm\_plot\_labels = ['benign', 'malignant']

plt.figure(figsize=(8, 6))

sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=cm\_plot\_labels, yticklabels=cm\_plot\_labels)

plt.xlabel('Predicted')

plt.ylabel('True')

plt.title('Confusion Matrix')

plt.show()

# Classification report

print(classification\_report(y\_true, y\_pred, target\_names=cm\_plot\_labels))

# Plot training history

acc = history.history['accuracy']

val\_acc = history.history['val\_accuracy']

loss = history.history['loss']

val\_loss = history.history['val\_loss']

epochs\_range = range(len(acc))

plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)

plt.plot(epochs\_range, acc, label='Training Accuracy')

plt.plot(epochs\_range, val\_acc, label='Validation Accuracy')

plt.legend(loc='lower right')

plt.title('Training and Validation Accuracy')

plt.subplot(1, 2, 2)

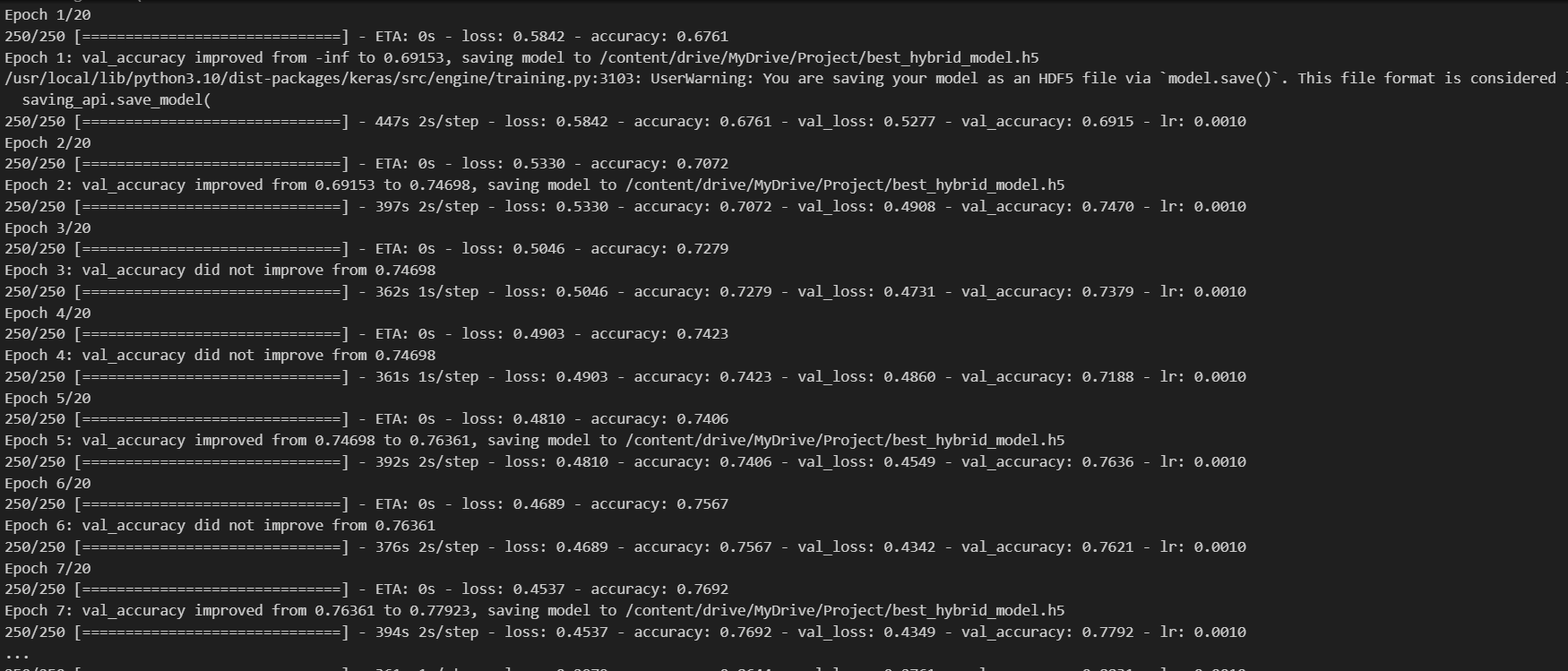
plt.plot(epochs\_range, loss, label='Training Loss')

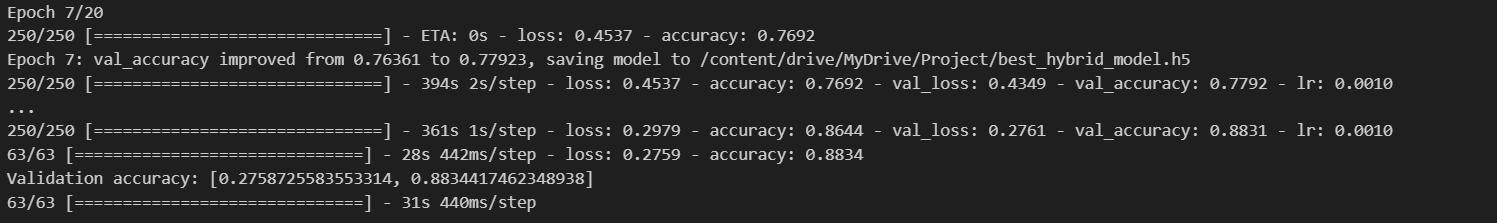
plt.plot(epochs\_range, val\_loss, label='Validation Loss')

plt.legend(loc='upper right')

plt.title('Training and Validation Loss')

plt.show()

**Screenshots:**

****

**Fig.18** Screenshot of epoches