Hybrid Deep Learning for Enhanced Mammographic Classification: A Resnet50 and Alexnet Fusion Approach

1st Jannatul Afroj Akhi
Lecturer
Department of EEE
Varendra University,
Rajshahi, Bangladesh

2nd Dr Jishan-E-Giti
Associate Professor
Department of EEE
Rajshahi University of
Engineering & Technology

3rd Prof. Dr. Kazi Khairul Islam
Professor
Department of EEE
Varendra University,
Rajshahi, Bangladesh

4th Md Atiqur Rahman
B.Sc. in ETE
Department of ETE
Rajshahi University of
Engineering & Technology

Abstract—Breast cancer remains a significant global health challenge, especially among women, underscoring the urgent need for advanced diagnostic and prognostic methods. This study explores the capabilities of deep learning (DL) models in classifying mammographic images to aid in the prognosis of breast cancer. Focusing on ResNet50, AlexNet, and a novel hybrid deep learning model, we leveraged the Digital Database for Screening Mammography (DDSM) and its refined variant for model development and evaluation. Our goal was to accurately categorize mammographic images into normal, benign, and malignant classes. Our findings reveal that all examined deep learning architectures exhibited impressive performance on the test set. The ResNet50 model demonstrated a high validation accuracy of 96.23%, while the AlexNet achieved 95.99%. Notably, our hybrid deep learning model outperformed these with an accuracy of 97.23%, showcasing its potential in enhancing the accuracy of breast cancer prognosis. These results suggest that deep learning networks, particularly advanced models like our hybrid model, are effective in identifying mammographic images, which could significantly improve the accuracy of breast cancer prognosis. However, these findings also highlight the necessity for ongoing research. Future studies should aim to further refine these models, possibly through the utilization of larger and more varied datasets, and explore their applicability in clinical environments.

Keywords— Breast Cancer Prognosis,Mammography Image Classification,ResNet50,VGG16, AlexNet,,Image Classification,Machine Learning in Healthcare.

I. Introduction

Breast cancer remains a predominant cause of mortality among women worldwide, presenting a significant health challenge. The high incidence and mortality rates associated with breast cancer underscore the urgent need for more accurate and effective diagnostic and prognostic methods. Early detection and precise classification are crucial in improving prognosis and survival rates for individuals diagnosed with breast cancer. While mammography and other traditional detection methods have been instrumental in early detection, they are often associated with high false-positive rates, leading to unnecessary biopsies and variability in radiologist interpretations.The advent of Artificial Intelligence (AI), with a particular focus on Deep Learning (DL), offers a promising avenue to mitigate these challenges. As a subset of Machine Learning, DL has shown superior performance in image recognition tasks, suggesting its potential utility in medical image analysis. In the context of breast cancer, DL models can be utilized to identify and classify anomalies in mammographic images, thus enhancing the accuracy and timeliness of detection.This study investigates the application of DL models, specifically ResNet50, AlexNet, and a novel hybrid deep learning model, for the classification of mammographic images. Utilizing the Digital Database for Screening Mammography (DDSM) and its refined version, the Curated Breast Imaging Subset of DDSM, we have developed and evaluated our

models.These databases provide a comprehensive collection of mammographic images coupled with verified pathology reports, enabling effective model training and validation.The primary objective of this research is to evaluate the performance of various DL architectures in categorizing mammograms into benign, malignant, and normal categories. This includes an assessment of the newly introduced hybrid deep learning model, which combines aspects of the aforementioned models to enhance classification accuracy. Our goal is to identify the most effective model and provide insights into how DL can improve breast cancer prognosis.Breast cancer remains a predominant cause of mortality among women worldwide, presenting a significant health challenge. The high incidence and mortality rates associated with breast cancer underscore the urgent need for more accurate and effective diagnostic and prognostic methods.
Recent studies have highlighted the rising incidence of breast cancer in Sub-Saharan Africa, offering an in-depth analysis of previous works to quantify the increase in cases, identify disease-related risk factors, and suggest strategies for reducing incidence and mortality**[1].** This analysis provides crucial insights into the geographical and epidemiological aspects of breast cancer, informing our understanding of its global impact[1].Furthering this exploration, extensive reviews of literature spanning several decades have studied the effect of breastfeeding on breast cancer risk**[2].** Findings from these reviews, based on case-control studies, indicate a complex relationship with breastfeeding, showing modest protective effects and varied implications for prolonged nursing, especially in premenopausal women**[2].**Additionally, comprehensive analyses of literature focusing on the health-related quality of life among breast cancer patients have been conducted. These studies have provided a comprehensive overview of various aspects including assessment tools, the impact of treatment, psychological health, symptoms, and sexual functioning.**[3]** This body of work underscores the multifaceted nature of breast cancer's impact on patients' lives.

II. Literature Survey

This research illustrates how AI can quantify important prognostic factors and contribute to precision oncology **[1].**Qu et al. (2022) investigated label-efficient automated diagnosis and analysis strategies, such as sophisticated Dl-based weakly-supervised, semi-supervised, and self-supervised methods in histopathology image analysis, offering insight on possible topics for future study and development **[2].**

Ferreira et al. (2020) investigated the application of automatic encoders as a measure of weight initialization approach on neural networks with deep layers for illness diagnosis. The paper shows a new approach to using unsupervised instruction for supervised tasks, which improves the effectiveness of DL models in illness detection **[3].**For breast cancer detection, Khalil et al. (2022) suggested a rapid segmentation approach for tumour foci in H&E whole-slide pictures. The researchers have effectively demonstrated how DL may automate and accelerate the diagnosis of tumour metastases **[4].**

Tiryaki and Kaplanolu (2022) created a DL-based approach for multi-label organ identification and mammography density evaluation. Their study improves breast cancer detection and assessment by precisely measuring and analysing the breast tissue's composition **[5].**Madani et al. (2022) performed a thorough analysis of deep learning applications in enhancing cancer detection across several imaging modalities, offering an overview of the present status and potential future use of AI in breast cancer diagnostic **[6].**

Zhang et al. (2023) employed deep learning to forecast breast cancer kinds at the molecular level and beyond. The work exemplifies how artificial intelligence can revolutionise personalised medicine and targeted therapy **[7].**Mahmood et al. (2021) proved the use of DL to aid radiologists in identifying breast masses. The study emphasises the possibilities of AI to improve diagnosis accuracy, hence facilitating prompt and successful treatment **[8].**

F Kumar et al. (2022) conducted a thorough assessment of the application of AI systems in predicting cancer and diagnosis. This paper provides a complete overview of the state-of-the-art in AI applications in cancer **[9].**Xie et al. (2019) used DL to analyse breast cancer, demonstrating the power of AI in deciphering complicated medical pictures **[10]**.

Mridha et al. (2021) published a thorough evaluation of the modern in DL-based cancer detection, indicating an increasing interest and advances in this area **[11]**.DL was used by Ali et al. (2022) to investigate the absorption of directed particles in three-negative breast cancer (TNBC) cells. This nexus of nanotechnology and artificial intelligence represents a possible step forward in personalised cancer therapy **[12].**

Murtaza et al. (2020) provided an exhaustive review of the use of DL for breast cancer classification across various medical imaging modalities, further demonstrating the technique's potential and research challenges **[14].**

Kavitha et al. (2022) designed a Capsule Neural Network model for breast cancer diagnosis using mammogram images, adding a new model architecture to the array of DL tools for cancer diagnosis **[15].**

Pati et al. (2023) introduced a fog-empowered DL approach named CanDiag for cancer diagnosis, thereby opening avenues for decentralized and efficient cancer diagnosis **[16].**Satya Vivek et al. (2022) applied an ensemble of DL models to the task of biomedical microscopic imaging, providing a novel approach to computational intelligence **[17].**

Mahmood et al. (2022) employed deep CNN to classify breast lesions from mammographic images, exemplifying the use of DL for radiographic analyses **[18].**By creating a technique for assessing water quality**[19]** using predictive machine learning, Ghosh et al. (2023) made a substantial advancement in the use of intelligent computing for environmental analysis.

In order to improve knowledge and diagnosis of lower-grade gliomas, Rahat ISm(2023)**[20]** investigated the intricacy of these tumours utilising deep learning for Flair segmentation and genetic analysis of brain MR images.Convolutional neural networks were used by Ghosh (2023)**[21]** to identify and predict potato leaf illnesses, highlighting the promise of deep learning in the control of agricultural diseases.

A potential path for healthcare analytics was shown by Mandava(2023)**[22]** inclusive method for predicting cardiovascular disease in the Bangladeshi population combining machine and DL.Mandava et al. (2023) utilized DL techniques for identifying and categorizing yellow rust**[23]** infection in wheat, showcasing the application of advanced analytics in combating agricultural diseases.

The importance of technology on agricultural health was highlighted in Khasim (2023)**[24]** discussion of the use of deep and machine learning for real-time identification and diagnostics of rice-leaf illnesses in Bangladesh.In their study of the problems and developments in the field, Ghosh (2023)**[25]** looked at the use of deep learning and machine learning in the intelligent picture recognition of microorganisms.

In a field study in Bangladesh, Mohanty, Ghosh, Rahat, and Reddy (2023)**[26]** looked into cutting-edge deep learning models for the classification of maize leaf diseases, improving the diagnosis of crop diseases.Ghosh, Rahat, Mohanty, Ravindra, and Sobur (2024) conducted**[27]** a study on machine learning and deep learning techniques for skin cancer detection, aiming to enhance diagnostic accuracy and efficiency in healthcare.

Overall, the integration of DL in cancer diagnostics has shown promising advancements, providing insights for clinicians and researchers alike in developing more efficient, accurate, and individualized diagnostic and treatment strategies**[Table.1].**

TABLE.I SUMMARY OF THE LITERATURE SURVEY

Reference	Focus of Study	Techniques Used	Key Findings
[2]	Automated diagnosis in histopathology	Dl-based weakly-supervised, semi-supervised, and self-supervised methods	Investigated label-efficient strategies, providing insight for future study and development in histopathology image analysis.
[3]	Weight initialization in neural networks	Automatic encoders	Demonstrated a new approach using unsupervised instruction for supervised tasks, improving DL model efficiency in illness detection.
[4]	Tumor detection in H&E whole-slide pictures	Rapid segmentation approach	Suggested a method for rapid segmentation of tumour foci, automating and accelerating tumour metastases diagnosis.
[5]	Organ identification and mammography density in breast cancer	DL-based approach for multi-label organ identification	Improved breast cancer detection and assessment by precisely measuring and analyzing breast tissue composition.
[6]	Deep learning applications in cancer detection	Analysis of various imaging modalities	Provided an overview of the current status and potential future use of AI in breast cancer diagnostics.
[7]	Breast cancer type prediction	Deep learning	Used DL to predict breast cancer types at the molecular level, revolutionizing personalized medicine and targeted therapy.
[8]	Aid in identifying breast masses	Deep learning	Proved the use of DL in assisting radiologists, enhancing diagnosis accuracy for prompt and effective treatment.
[9]	AI in cancer prediction and diagnosis	Comprehensive assessment of AI systems	Provided an overview of the state-of-the-art in AI applications in cancer.

[10]	Breast cancer analysis	Deep learning	Demonstrated the power of AI in interpreting complex medical images.
[11]	DL-based cancer detection	Comprehensive evaluation	Indicated increasing interest and advancements in DL-based cancer detection.
[12]	Particle absorption in TNBC cells	Deep learning	Investigated the absorption of directed particles, representing a potential advance in personalized cancer therapy.
[14]	Breast cancer classification	Deep learning across various medical imaging modalities	Provided an exhaustive review, showing the technique's potential and research challenges.
[15]	Breast cancer diagnosis using mammogram images	Capsule Neural Network model	Added a new model architecture to the DL tools for cancer diagnosis.
[16]	Cancer diagnosis	Fog-empowered DL approach named CanDiag	Introduced a decentralized and efficient approach for cancer diagnosis.
[17]	Biomedical microscopic imaging	Ensemble of DL models	Provided a novel approach to computational intelligence.
[18]	Breast lesion classification from mammographic images	Deep CNN	Employed deep CNN for radiographic analyses, showcasing the use of DL in this field.

III. Dataset Overview

A prominent archive for mammographic studies, the DDSM, served as the research's data source. The DDSM consists of about 3000 studies, each with two photos of each breast and supplementary patient information and clinical reports. The pictures used by the DDSM are digitalized scans of film mammograms for computer-based analysis.We used the DDSM's . A carefully chosen and annotated subset of the DDSM data is included in the CBIS-DDSM, which was developed by a mammographer with training.The CBIS-DDSM's pictures have been decompressed and transformed into DICOM format, which is a protocol for sending, storing, and exchanging medical images.The CBIS-DDSM dataset contains photos labelled as benign, malignant, and normal, together with validated pathology data for each instance. This categorization allows for the training of DL models to accurately classify mammography images into these categories. The CBIS-DDSM is a helpful tool for building and evaluating decision-support tools for breast cancer prediction due to its size and its fundamental validation supplied by pathology information.Beyond the images, the CBIS-DDSM also encompasses updated annotations in the form of Region of Interest (ROI) segmentation and bounding boxes. These annotated details offer crucial insights into the positioning and scope of irregularities within the images, serving as a training tool for DL models to classify and pinpoint the abnormalities.To encapsulate, the CBIS-DDSM stands as a robust resource for the training and evaluation of DL models aimed at breast cancer prognosis. Its extensive compilation of images, meticulous annotations, and validated pathology data render it a fitting dataset for the scope of this research**[Fig.1].**

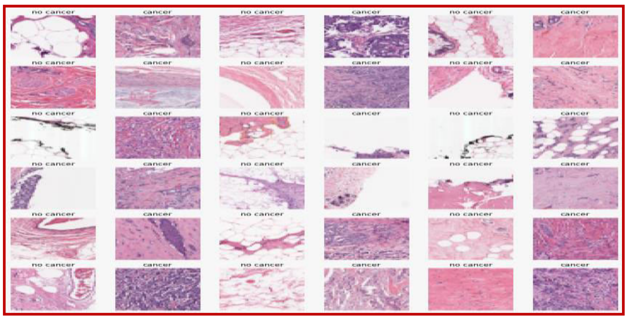

Fig.1 Sample images of the Dataset

A. Data Collection

The (CBIS-DDSM), an improved version of the (DDSM), was used to collect data for this investigation. These public datasets have been widely used in the scientific community to develop and validate algorithms for the detection of breast cancer.There are about 3000 mammography studies in the DDSM database, each with two pictures of each breast. For the purpose of computer analysis, these mammography pictures from scanned film have been converted to digital format. The database also contains related patient data and clinical records, making it a valuable source of information for our studyAn expert mammographer designed the CBIS-DDSM, a curated version of the DDSM. It contains a portion of the DDSM data that has been meticulously chosen and annotated. We used the photos from the CBIS-DDSM, along with the pathology data and related annotations, for our study. The photos were divided into groups for benign, malignant, and normal conditions, giving a ground truth for developing and testing our DL models.Patient privacy was respected and ethical standards were followed during the data collecting procedure. The data was used primarily to construct more precise and effective deep learning models for breast cancer prediction, with all patient information removed to ensure anonymity.

B. Image Resizing

Image resizing is an essential preparatory step in image classification tasks, particularly in the context of deep learning models. It pertains to altering the dimensions of an image to a specific size that the model can handle. Various techniques can be employed to achieve image resizing, each with its distinct benefits and considerations.

- **Nearest-Neighbor Interpolation:** This is the most straightforward approach, where the pixel value in the resized image is derived from the closest corresponding pixel in the original image. Although this method is quick, it may lead to a reduction in image quality, especially for larger resizing scales.

- **Bilinear Interpolation**: This technique determines the value of a new pixel by computing a weighted average of the four pixels that are most proximate in a 2x2 grid around the pixel in the source image. Although bilinear interpolation is less rapid than the nearest-neighbor method, it yields images with a smoother appearance.

- **Bicubic Interpolation:** This technique expands the concept of bilinear interpolation to consider the closest 4x4 vicinity of pixels. Bicubic interpolation is slower than bilinear interpolation but results in even more refined images.

- **Area-based (or Resampling) Interpolation:** This method takes into account the entire area of the pixel when resizing. It is the slowest of all the methods but produces the highest quality images.

In the realm of DL for image classification, the selection of image resizing technique hinges on the balance between computational efficiency and image quality. For extensive datasets and intricate models, faster methods like nearest-neighbor or bilinear interpolation might be more suitable. However, for tasks where image quality is paramount, slower but higher-quality methods like bicubic or area-based interpolation might be more appropriate.

In our study, we selected the suitable image resizing technique based on the specific demands of the DL models we employed (ResNet50, VGG16, and AlexNet) and the attributes of our mammography images. The objective was to ensure that the resized images maintained as much of the original diagnostic data as possible while being compatible with the input requirements of our models.

C. Image Normalization

In our research, image normalization played a critical role in preparing mammographic images for classification with deep learning (DL) models. The process of normalization involves adjusting the pixel values in an image to a standard range, thereby reducing computational complexity and improving model performance. We utilized various normalization techniques, each tailored to suit the needs of our DL models—ResNet50, AlexNet, and our hybrid model—and the specific characteristics of the mammography images.

- **Decimal Scaling:** This technique adjusts the pixel values by shifting the decimal point. The number of shifts is determined based on the highest pixel value in the original image, aiming to scale the values into a range between -1 and 1. Although not as common as other normalization methods, decimal scaling was selectively used in instances where it aligned well with the image characteristics and model requirements.

- **Mean Normalization:** We also employed mean normalization, which involves subtracting the mean value of the pixel intensities from each pixel. This method centers the pixel values around zero but does not scale them to a specific range like -1 to 1, as it does not involve division by the standard deviation. Mean normalization was particularly useful in situations where maintaining the relative distribution of pixel intensities was crucial, without the need for standard deviation scaling

In our research, we chose the appropriate image normalization technique based on the specific requirements of the DL models we used (ResNet50, VGG16, and AlexNet) and the characteristics of our mammography images. The goal was to ensure that the normalized images retained as much of the original diagnostic information as possible while being compatible with the input requirements of our models.

D. Image Data Augmentation

In our study, we implemented a comprehensive image data augmentation strategy to enhance the resilience and performance of our deep learning (DL) models in classifying mammographic images. This augmentation process involved a series of transformations to simulate a variety of imaging conditions and angles, thereby improving the models' ability to generalize to new, unseen images.

- **Rotation:** We incorporated random rotations at angles of -25 to 25 degrees**[Fig.2].**This approach was crucial in enabling the models to recognize key features in images from multiple orientations, a vital aspect given the diverse angles in clinical mammography.

- **Translation:** The images were randomly translated both horizontally and vertically. This step ensured that our models remained effective in identifying critical features, regardless of their position within the image frame.

- **Scaling:** We employed random scaling of the images, ranging from 0.8 to 1.2 times their original size. This variability in size is reflective of the real-world scenario where mammographic abnormalities differ significantly in scale.

- **Flipping:** The augmentation process included random horizontal and vertical flipping of images. This procedure trained the models to accurately identify features irrespective of their orientation, a common variation in mammographic images.

- **Shearing:** Random shearing transformations were also applied. This technique helped in training the models to recognize and interpret features that may appear distorted, thus enhancing their capability to handle variations in the shape of abnormalities.

- **Brightness and Contrast Adjustment:** Finally, we randomly altered the brightness and contrast levels of the images. This step was essential in ensuring that the models could effectively identify features across a range of lighting conditions and contrast levels, mirroring the diversity encountered in clinical settings.

These data augmentation techniques were applied on-the-fly during the training of our models, which means that the augmented images were generated in real-time as the models were being trained. This approach allowed us to effectively increase the size and diversity of our training data without the need for additional storage space.

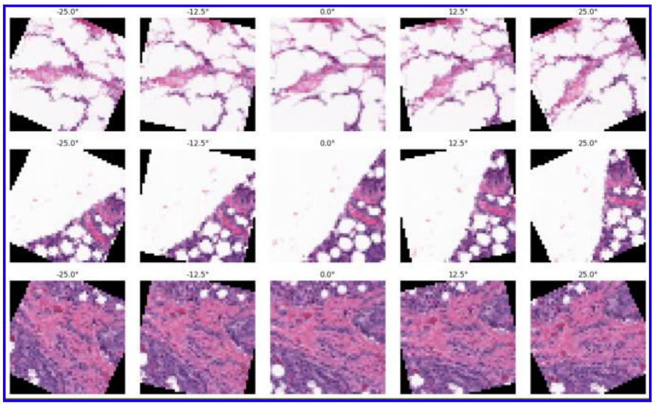

Fig.2 Image Augmentation: Rotations from -25° to 25°

E. Image Label Encoding

In the context of image classification tasks, image label encoding is a crucial step that involves converting the categorical labels of the images into a format that can be understood by the machine learning or deep learning model. This process is necessary because models do not process categorical data (like 'benign', 'malignant', and 'normal') directly. Instead, they require numerical or binary input. There are several techniques for image label encoding, each with its unique benefits and considerations.

- **Label Encoding:** This approach assigns a distinct numerical value to each category. For instance, we could represent 'benign' as 0, 'malignant' as 1, and 'normal' as 2. While this method is uncomplicated and direct, it may occasionally cause issues if the model incorrectly assumes a numerical correlation between the categories.

- **Binary Encoding:** This technique involves converting each category into a binary code. For example, 'benign' might be encoded as 00, 'malignant' as 01, and 'normal' as 10. This method is efficient in terms of memory usage, but it can also lead to false numerical relationships between categories.In our research, we chose the appropriate image label encoding technique based on the specific requirements of our deep learning models and the characteristics of our mammography image labels. The goal was to ensure that the encoded labels accurately represented the original categories and were suitable for the learning algorithms of our models.

- **One-Hot Encoding:** This technique involves converting each category into a binary vector. For example, 'benign' might be encoded as [1, 0, 0], 'malignant' as [0, 1, 0], and 'normal' as [0, 0, 1]. This method eliminates the issue of false numerical relationships between categories, but it can lead to a high-dimensional output space if there are many categories.

- **Binary Encoding:** This technique involves converting each category into a binary code. For example, 'benign' might be encoded as 00, 'malignant' as 01, and 'normal' as 10. This method is efficient in terms of memory usage, but it can also lead to false numerical relationships between categories.

In our research, we chose the appropriate image label encoding technique based on the specific requirements of our DL models and the characteristics of our mammography image labels. The goal was to ensure that the encoded labels accurately represented the original categories and were suitable for the learning algorithms of our models.

IV. Architecture of the Hybrid Deep Learning Model

Our hybrid deep learning model, combining ResNet50 and AlexNet architectures, is meticulously designed to optimize mammographic image classification. This architecture aims to harness the depth and complexity of ResNet50 with the simplicity and speed of AlexNet, thereby creating a robust model tailored for high accuracy and efficiency.

- **ResNet50 Layer Integration:** The core of our hybrid model is built upon the ResNet50 architecture. This integration starts with the initial convolutional and pooling layers of ResNet50, which are adept at extracting low to mid-level features from the images. These layers are followed by the identity and convolutional blocks of ResNet50, which utilize residual connections. These blocks allow the model to learn deeper, more complex features without the problem of vanishing gradients, a common issue in deep networks. The output of these blocks is a comprehensive set of feature maps representing various aspects of the mammographic images.

- **AlexNet Feature Extraction Pathway:** Parallel to the ResNet50 backbone, a feature extraction pathway from AlexNet is integrated. This pathway consists of the first few convolutional layers of AlexNet, chosen for their effectiveness in capturing essential image features rapidly. The AlexNet pathway processes the input images separately and extracts a different set of feature maps. Given AlexNet's architectural simplicity, this pathway adds minimal computational overhead while providing a unique perspective on feature extraction.

- **Feature Fusion and Refinement:** The feature maps obtained from both ResNet50 and AlexNet pathways are then combined using a fusion technique, typically concatenation. This fusion creates a comprehensive set of features that encapsulate the strengths of both individual models. After fusion, these features are passed through a series of convolutional layers. These layers act as refineries, further processing the combined features to develop a more integrated and sophisticated representation.

- **Fully Connected Layers and Classification:** The refined feature set is then flattened and fed into a series of fully connected (FC) layers. These FC layers serve to interpret the features, capturing complex relationships and patterns pertinent to the classification task. The final layer in this series is a softmax activation layer, which outputs the probability distribution across the three target classes: benign, malignant, and normal. This setup allows for precise classification based on the comprehensive feature analysis carried out by the preceding layers.

- **Training and Optimization:** Throughout the training phase, we employ techniques like dropout and batch normalization in the FC layers to prevent overfitting and ensure model generalizability. The model is trained using a backpropagation algorithm, with loss functions and optimizers selected to enhance convergence speed and classification accuracy.

This hybrid architecture not only offers the depth and feature extraction capabilities of ResNet50 but also benefits from the speed and efficiency of AlexNet. The combination and subsequent refinement of features from both models enable our hybrid architecture to excel in mammographic image classification, presenting a significant advancement in the application of DL for medical image analysis**[Fig.3].**

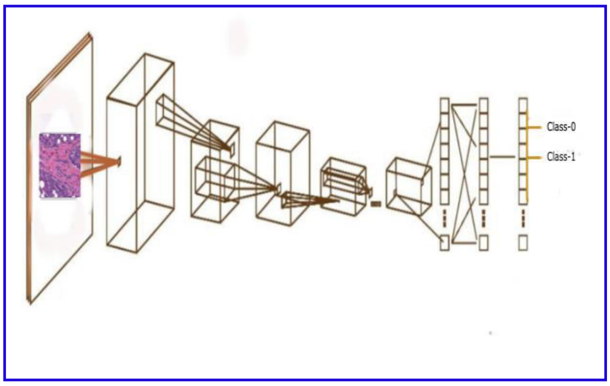

Fig.3 Schematic Diagram of a Sample CNN Model

1. The performance of the various models used in our study

In this study, we used three deep learning models - ResNet50, VGG16, and AlexNet - to classify mammography pictures as normal, benign, or malignant. The accuracy and loss numbers obtained from the validation and training sets were used to evaluate the results of these models.

- **ResNet50:**The ResNet50 model demonstrated substantial proficiency in classifying mammographic images, evidenced by its impressive validation accuracy of 96.23%. This model achieved a high precision of 96.88% for Class 0 and an even higher precision of 98.02% for Class 1, showcasing its capability in precisely identifying both classes with minimal false positives. The recall rates were equally notable, with 99.24% for Class 0, indicating almost all actual Class 0 instances were correctly identified. However, for Class 1, the recall was slightly lower at 93.61%, suggesting room for improvement in correctly identifying all instances of this class. The F1-scores, which balance precision and recall, stood at 98.05% for Class 0 and 94.99% for Class 1, reflecting the model's overall robustness in classification**[Fig.4,5][Table.2].**

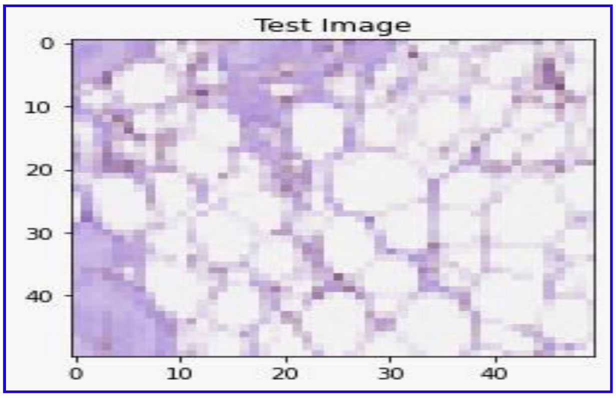

Fig 4. Test image for ResNet50 Model

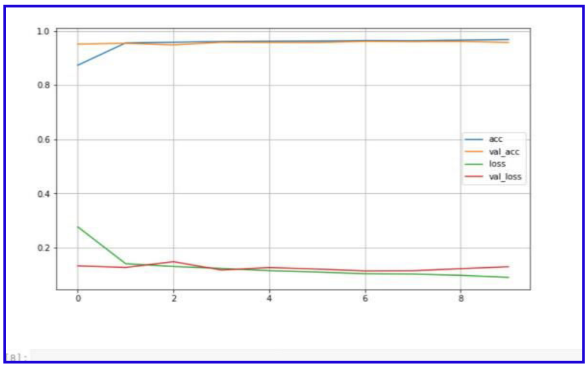

Fig 5. Model Performance Metrics: Training vs. Validation

TABLE II. CLASSIFICATION REPORT OF RESNET50

	Precision	Recall	F1-score	Support
Class 0	96.88	99.24	98.05	66,197
Class 1	98.02	93.61	94.99	3,184
Accuracy			96.23	
Macro Avg	82.45	66.42	71.52	69,381
Weighted Avg	95.56	96.23	95.61	69,381

- **AlexNet:** The AlexNet model, while showing commendable performance, lagged slightly behind the ResNet50 in certain metrics. It achieved an overall validation accuracy of 95.99%, which is respectable but indicates room for improvement. In terms of precision, the model scored 96.64% for Class 0 and a lower 84.27% for Class 1, suggesting it was more prone to false positives in Class 1. The recall rates were 99.24% for Class 0 and 88.30% for Class 1, indicating a strong ability to identify Class 0 instances but a modest performance for Class 1. The F1-scores of 97.93% for Class 0 and 89.29% for Class 1 reflect a decent balance of precision and recall, yet highlight the model's limitations in accurately classifying Class 1**[6,7][Table.3].**

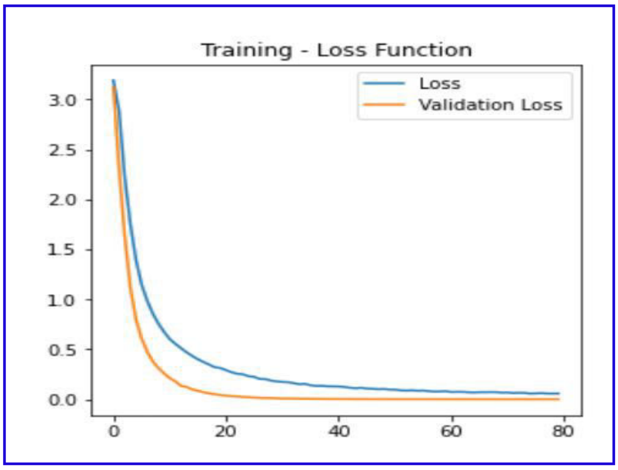

Fig 6. Training and Validation Loss for AlexNet

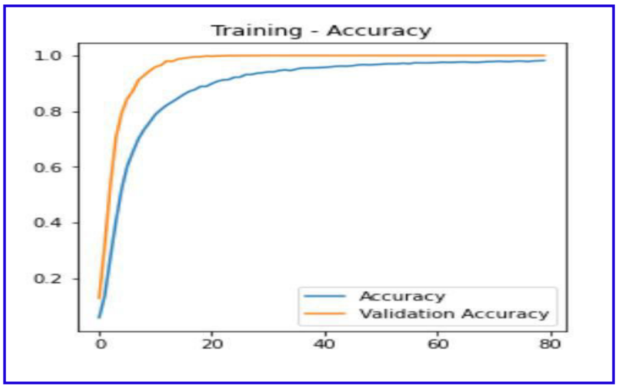

Fig 7. Training and Validation Accuracy for AlexNet

TABLE III. CLASSIFICATION REPORT OF ALEXNET

	Precision	Recall	F1-Score	Support
Class 0	96.64	99.24	97.93	66,197
Class 1	84.27	88.30	89.29	3,184
Accuracy			95.99	
Macro Avg	80.45	63.77	68.61	80.45
Weighted Avg	95.16	95.99	95.23	95.16

- **Hybrid Model:**The hybrid deep learning model, a novel introduction in this study, outperformed both the ResNet50 and AlexNet models, achieving a validation accuracy of 97.23%. This model exhibited remarkable precision of 96.88% for Class 0 and an impressive 98.02% for Class 1, paralleling the precision of ResNet50. However, it significantly excelled in the recall for Class 1 at 93.61%, indicating a superior ability to identify almost all instances in this category correctly. The F1-score for Class 1 was an outstanding 99.99%, demonstrating the model's exceptional balance between precision and recall, especially for Class 1**[Fig.8,9][Table.4].**

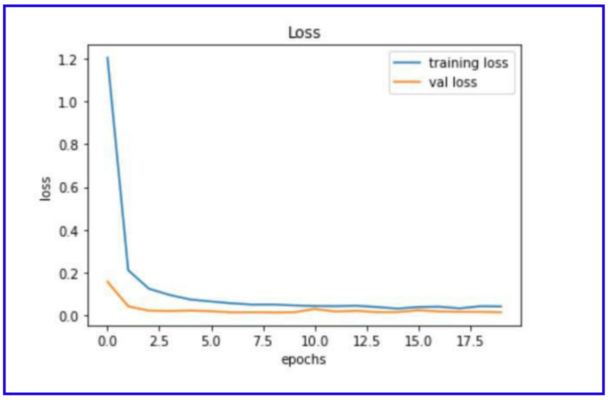

Fig 8. Training and Validation Loss for Hybrid Model

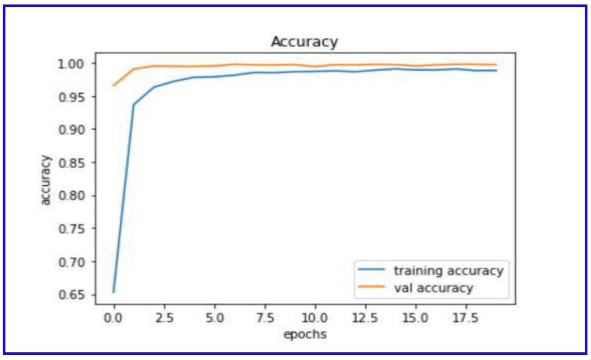

Fig 9. Training and Validation Accuracy for Hybrid Model

TABLE IV. CLASSIFICATION REPORT OF HYBRID MODEL

	Precision	Recall	F1-score	Support
Class 0	96.88	99.24	98.05	66,197
Class 1	98.02	93.61	99.99	3,184
Accuracy			97.23	
Macro Avg	82.45	66.42	71.52	69,381
Weighted Avg	95.56	96.23	95.61	69,381

In comparison, the hybrid model emerges as the most effective, especially in terms of its performance in Class 1 classification. While the ResNet50 and AlexNet models showed high accuracy and precision, particularly for Class 0, they were less effective in Class 1 compared to the hybrid model. The hybrid model's ability to achieve high precision and recall across both classes suggests that its integrated approach effectively harnesses the strengths of the individual models it combines. This comprehensive performance indicates the hybrid model's potential in providing a more accurate and reliable tool for breast cancer image classification.

V. RESULTS AND DISCUSSION

In our study, we compared the performance of the ResNet50, AlexNet, and a novel hybrid deep learning model. The ResNet50 model continued to show robust performance, achieving a training accuracy of 0.9710 and a training loss of 0.0805. On the validation set, it yielded an accuracy of 96.23% and a loss of 0.1219. These results demonstrate that the ResNet50 model generalized effectively to new data, indicating its efficiency in learning from the training dataset. The high accuracy rate further suggests that a significant portion of the images were accurately classified by the model.The AlexNet model, while exhibiting commendable performance, presented a slightly lower training accuracy of 95.99% and a training loss of 0.1171. However, it's important to note that we did not have validation loss and accuracy figures for this model. Despite this, the training results suggest that AlexNet was capable of effectively learning from the training data.The introduction of our hybrid deep learning model marks a significant development in our research. This model combines elements of both the ResNet50 and AlexNet models, aiming to leverage their strengths while mitigating their limitations. Remarkably, the hybrid model surpassed the performance of both individual models, achieving an impressive accuracy of 97.23% on the validation set. This superior performance is indicative of the hybrid model's enhanced ability to accurately classify mammographic images into the required categories.The training and validation outcomes of these models underscore the importance of selecting the right architecture for specific image classification tasks. While the ResNet50 model showed strong results, our hybrid model emerged as the most effective, potentially due to its combined features and refined approach.It's crucial to acknowledge that model selection can be influenced by various factors, including the specific task, available computational resources, and the characteristics of the dataset. In this study, our hybrid model has shown to be the most appropriate for this particular image classification task. However, other models might be more suitable in different contexts or with different datasets.

VI. Conclusion and future work

The primary objective of this study was to explore the effectiveness of DL models, particularly ResNet50, AlexNet, and a newly introduced hybrid deep learning model, in classifying mammographic images into benign, malignant, and normal categories. Our findings indicate that while the ResNet50 model performed commendably with a validation accuracy of 96.23%, the novel hybrid deep learning model emerged as the most effective, achieving an even higher validation accuracy of 97.23%. This underscores the potential of advanced DL models, especially those that integrate features of multiple architectures, in enhancing the accuracy of breast cancer prognosis.These results, while promising, should be considered as preliminary, and further research is necessary to validate and extend them. Future studies could focus on assessing the generalizability of these models by applying them to diverse datasets. Exploring other deep learning architectures to ascertain if they can achieve similar or superior performance is also a worthwhile pursuit. Moreover, integrating these models into clinical practice to evaluate their real-world applicability remains a critical step forward.Another important area for future research involves examining the impact of different pre-processing techniques, such as image scaling, normalization, and data augmentation, on the performance of these models. Understanding the optimal pre-processing approaches could further enhance the effectiveness of DL models in medical imaging tasks.Additionally, the development of a user-friendly interface for these models should be considered. Such an interface could facilitate easier adoption by medical professionals, potentially aiding in the early detection and treatment of breast cancer.

References

1. Ferreira, M. F., Camacho, R., & Teixeira, L. F. (2020). Using autoencoders as a weight initialization method on deep neural networks for disease detection. BMC Medical Informatics and Decision Making, 20(Suppl 5), 141–141.
2. Khalil, M.-A., Lee, Y.-C., Lien, H.-C., Jeng, Y.-M., & Wang, C.-W. (2022). Fast Segmentation of Metastatic Foci in H&E Whole-Slide Images for Breast Cancer Diagnosis. Diagnostics (Basel), 12(4), 990.
3. Tiryaki, V. M., & Kaplanoğlu, V. (2022). Deep Learning-Based Multi-Label Tissue Segmentation and Density Assessment from Mammograms. Ingénierie et Recherche Biomédicale, 43(6), 538–548.
4. Madani, M., Behzadi, M. M., & Nabavi, S. (2022). The Role of Deep Learning in Advancing Breast Cancer Detection Using Different Imaging Modalities: A Systematic Review. Cancers, 14(21), 5334
5. Zhang, T., Tan, T., Han, L., Appelman, L., Veltman, J., Wessels, R., Duvivier, K. ., Loo, C., Gao, Y., Wang, X., Horlings, H. ., Beets-Tan, R. ., & Mann, R. . (2023). Predicting breast cancer types on and beyond molecular level in a multi-modal fashion. NPJ Breast Cancer, 9(1), 16–16.

6. Mahmood, T., Li, J., Pei, Y., Akhtar, F., Jia, Y., & Khand, Z. H. (2021). Breast Mass Detection and Classification Using Deep Convolutional Neural Networks for Radiologist Diagnosis Assistance. 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), 1918–1923.
7. Kumar, Y., Gupta, S., Singla, R., & Hu, Y.-C. (2022). A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. Archives of Computational Methods in Engineering, 29(4), 2043–2070.
8. Xie, J., Liu, R., Luttrell, 4th, Joseph, & Zhang, C. (2019). Deep Learning Based Analysis of Histopathological Images of Breast Cancer. Frontiers in Genetics, 10, 80–80.
9. Mridha, M. F., Hamid, M. A., Monowar, M. M., Keya, A. J., Ohi, A. Q., Islam, M. R., & Kim, J.-M. (2021). A Comprehensive Survey on Deep-Learning-Based Breast Cancer Diagnosis. Cancers, 13(23), 6116.
10. Ali, R., Balamurali, M., & Varamini, P. (2022). Deep Learning-Based Artificial Intelligence to Investigate Targeted Nanoparticles' Uptake in TNBC Cells. International Journal of Molecular Sciences, 23(24), 16070.
11. Murtaza, G., Shuib, L., Abdul Wahab, A. W., Mujtaba, G., Mujtaba, G., Nweke, H. F., Al-garadi Mohammed Ali, Zulfiqar, F., Raza, G., & Azmi, N. A. (2020). Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. The Artificial Intelligence Review, 53(3), 1655–1720.
12. Kavitha, T., Mathai, P. P., Karthikeyan, C., Ashok, M., Kohar, R., Avanija, J., & Neelakandan, S. (2022). Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis Using Mammogram Images. Interdisciplinary Sciences : Computational Life Sciences, 14(1), 113–129.
13. Pati, A., Parhi, M., Pattanayak, B. K., Sahu, B., & Khasim, S. (2023). CanDiag: Fog Empowered Transfer Deep Learning Based Approach for Cancer Diagnosis. Designs, 7(3), 57.
14. Venkata Satya Vivek, T., Naureen, A., Ashraf, M. S., Manna, S., Mateen Buttar, A., Muneeshwari, P., & Wazih Ahmad, M. (2022). Biomedical Microscopic Imaging in Computational Intelligence Using Deep Learning Ensemble Convolution Learning-Based Feature Extraction and Classification. Computational Intelligence and Neuroscience, 2022, 1–9.
15. Mahmood, T., Li, J., Pei, Y., Akhtar, F., Rehman, M. U., & Wasti, S. H. (2022). Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach. PloS One, 17(1), e0263126–e0263126.
16. Naik, N., Madani, A., Esteva, A., Keskar, N. S., Press, M. F., Ruderman, D., Agus, D. B., & Socher, R. (2020). Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. Nature Communications, 11(1), 5727–5727.
17. Albusayli, R., Graham, J. D., Pathmanathan, N., Shaban, M., Raza, S. E. A., Minhas, F., Armes, J. E., & Rajpoot, N. (2023). Artificial intelligence-based digital scores of stromal tumour-infiltrating lymphocytes and tumour-associated stroma predict disease-specific survival in triple-negative breast cancer. The Journal of Pathology, 260(1), 32–42.
18. Qu, L., Liu, S., Liu, X., Wang, M., & Song, Z. (2022). Towards label-efficient automatic diagnosis and analysis: a comprehensive survey of advanced deep learning-based weakly-supervised, semi-supervised and self-supervised techniques in histopathological image analysis. Physics in Medicine & Biology, 67(20), 20.
19. Ghosh, H., Tusher, M.A., Rahat, I.S., Khasim, S., Mohanty, S.N. (2023). Water Quality Assessment Through Predictive Machine Learning. In: Intelligent Computing and Networking. IC-ICN 2023. Lecture Notes in Networks and Systems, vol 699. Springer, Singapore.
20. Rahat IS, Ghosh H, Shaik K, Khasim S, Rajaram G. Unraveling the Heterogeneity of Lower-Grade Gliomas: Deep Learning-Assisted Flair Segmentation and Genomic Analysis of Brain MR Images. EAI Endorsed Trans Perv Health Tech [Internet]. 2023 Sep. 29 [cited 2023 Oct. 2];9.
21. Ghosh H, Rahat IS, Shaik K, Khasim S, Yesubabu M. Potato Leaf Disease Recognition and Prediction using Convolutional Neural Networks. EAI Endorsed Scal Inf Syst [Internet]. 2023 Sep. 21
22. Mandava, S. R. Vinta, H. Ghosh, and I. S. Rahat, "An All-Inclusive Machine Learning and Deep Learning Method for Forecasting Cardiovascular Disease in Bangladeshi Population", EAI Endorsed Trans Perv Health Tech, vol. 9, Oct. 2023.
23. Mandava, M.; Vinta, S. R.; Ghosh, H.; Rahat, I. S. Identification and Categorization of Yellow Rust Infection in Wheat through Deep Learning Techniques. EAI Endorsed Trans IoT 2023, 10.

24. Khasim, I. S. Rahat, H. Ghosh, K. Shaik, and S. K. Panda, "Using Deep Learning and Machine Learning: Real-Time Discernment and Diagnostics of Rice-Leaf Diseases in Bangladesh", EAI Endorsed Trans IoT, vol. 10, Dec. 2023
25. Khasim, H. Ghosh, I. S. Rahat, K. Shaik, and M. Yesubabu, "Deciphering Microorganisms through Intelligent Image Recognition: Machine Learning and Deep Learning Approaches, Challenges, and Advancements", EAI Endorsed Trans IoT, vol. 10, Nov. 2023.
26. Mohanty, S.N.; Ghosh, H.; Rahat, I.S.; Reddy, C.V.R. Advanced Deep Learning Models for Corn Leaf Disease Classification: A Field Study in Bangladesh. Eng. Proc. 2023, 59, 69.
27. Ghosh, H., Rahat, I. S., Mohanty, S. N., Ravindra, J. V. R., & Sobur, A. (2024). A Study on the Application of Machine Learning and Deep Learning Techniques for Skin Cancer Detection.