International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 5853-5866
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Radiomics Meets Deep Learning: A Hybrid Approach for Breast Cancer Prediction from Mammographic Data

Rajendra Randa[1], Sanjeev Gour[2*], Priyanka Parmar[3], Chhaya Moghe[4], Priyanka Jain[5], Shiv Shakti Shrivastava[6]

[1]Research Scholar, Medicaps University Indore, MP, India.
[2]Research Supervisor & Assistant Professor, Medicaps University Indore, MP, India.
[3]Assistant Professor, Medicaps University Indore, MP, India.
[4]Assistant Professor, Medicaps University Indore, MP, India.
[5]Assistant Professor, Medicaps University Indore, MP, India.
[6]Professor, Parul University, Vadodara, Gujrat, India.
* Corresponding Author Email: sunj129@gmail.com - ORCID: https://orcid.org/0000-0002-2164-1590

Article Info:

DOI: 10.22399/ijcesen.3137
Received : 02 February 2025
Accepted : 30 March 2025

Keywords

Breast Cancer Diagnosis,
Mammography Image Classification,
Deep Neural Models,
CNN models,
Medical Image Analysis.

Abstract:

Breast cancer persists as the most commonly occurring malignancy in the female population around the world. Early and accurate detection is essential to enhance patient prognosis and the detection of pulmonary nodules using deep learning algorithms has recently been highlighted in medical image analysis. This work investigates hands-on deep neural learning approach for breast mammographic cancer prediction in early stages. Histopathology data are pre-processed (normalized and standard resized) to keeps the input size consistent and to maximize the features. A CNN architecture was constructed and trained using the Keras library, with TensorFlow as the backend. Data augmentation was used to regularization of the algo and to prevent high variance to the small amount of dataset, using rotation, flip, and scale manipulation. Models Five types of two-dimensional networks were developed: a custom CNN, VGG16 with frozen convolutional layers, and three networks (ResNet50, DenseNet121, and EfficientNetB3) trained from scratch.Of these, the custom CNN architecture showed potential: a high accuracy to discriminate between malignant and benign tumors. These results highlight the strength of deep convolutional models for medical images classification problems and suggest that the depth of the architecture and the right choice of regularization strategy are fundamental in achieving robust detection of cancer.This research further enriches the emerging area of AI-based diagnostics by providing comparative statistics for DL approaches in predicting the occurrence of breast cancer. Our results provide useful guidance for the future development of more accurate, scalable, and immediately practical predictors.

1. Introduction

In modern era, Breast Mammographic cancer disease is very common form of cancers and a leading cause of cancer-based deaths among women's lives worldwide. It represents a many high impact research area in medical image analysis[1] and demands accurate detection and classification approaches for screening mammograms[1]. Detection of breast cancer on early stage leads to better prognosis and increased survival rates, with the 5-year survival rate reaching 99% when the cancer is localized only in the breast[2]. Breast cancer diagnosis at an early stage using mammography is particularly important as it assists clinical specialists in treatment planning to increase survival rates[3].

At present, mammography stands as the most commonly employed and effective imaging method for the detection of this variance of the cancer. Nevertheless, the explication of mammograms can be challenging, often leading to diagnostic errors. This difficulty arises from the complexity of differentiating malignant lesions from surrounding normal tissues, which may result in false positives or negatives [3].

To support radiologists in accurately identifying breast abnormalities, computer-aided diagnosis (CAD) systems must effectively perform detection, segmentation, and classification of medical images [1]. This has led to significant interest in developing automated systems that can assist radiologists in the interpretation of mammograms. Advancements in deep learning—especially in convolutional neural models — have led to consequence improvements in the performance images recognition. Since 2012, deep CNNs have not only reached near-human accuracy but have also considerably outperformed conventional image analysis techniques, including those used in current CAD systems. These networks hold great promise for transforming medical image interpretation, particularly in the analysis of mammograms [4].

This literature review comprehensively explores the present state of research on breast cancer prediction using deep learning techniques applied to mammography data. It examines various methodologies, architectures, and evaluation approaches, with a particular focus on CNN-based models such as custom CNNs, VGG16, ResNet50, DenseNet121, and EfficientNetB3. The review also addresses data preprocessing techniques, data augmentation strategies, and performance evaluation methods.

2. Datasets and Data Preprocessing

2.1 Mammography Datasets

Multiple publicly available datasets are frequently utilized and evaluate deep neural models for training in breast cancer prediction. Notable examples include the CBIS-DDSM and the mini-MIAS mammography database, both of which are commonly employed in research studies [2]. The CBIS-DDSM serves as a refined and formalized version of the original DDSM, offering decompressed images, curated data selected by trained radiologists, enhanced segmentation for masses, bounding box annotations, and accompanying pathological diagnoses. It comprises 753 cases involving calcifications and 891 involving masses, providing a substantial dataset suitable for developing and assessing decision support systems in mammography [5].

In some cases, researchers utilize multiple datasets to improve model training and evaluation. For example, the TTCNN method has been applied across various datasets, including DDSM, INbreast, and MIAS, to enhance generalization and robustness of the classification outcomes [1]. The MIAS 2012 dataset has also been used, which contains seven classes with 322 image samples[6]. Some research conducts experimental processes on multiple publicly available datasets, such as CBIS-DDSM and INbreast, achieving average accuracies of 95.4% and 99.7% respectively[7].

2.2 Data Filtration Techniques

Data filtration is a crucial step in developing effective deep learning models for mammography analysis. Initial preprocessing typically involves removing unwanted regions of mammograms, enhancing quality, and highlighting cancerous lesions through various techniques including artifacts removal, noise reduction, and enhancement [3].

Robust image filtration techniques play a crucial role in minimizing diagnostic errors, particularly false positive and negative values, in mammographic breast cancer detection. These techniques encompass procedures such as background elimination, pectoral muscle removal, artificial noise insertion, and image enhancement to improve feature visibility and model accuracy [8]. Modified contrast enhancement methods have been particularly effective in accentuating edge details in mammogram images, aiding in more accurate interpretation [1].

For erasing the background, algorithms like Huang's Fuzzy Thresholding & Rolling Ball Algorithm have demonstrated complete success, effectively eliminating the background in 100% of cases. Similarly, Techniques like Canny Edge Detection and Hough Line Transform have demonstrated a high success rate—approximately 99%—in the removal of pectoral muscles from medical images. In addition, tools for image enhancement, including LUTs such as "Invert," "CTI_RAS," and "ISOCONTOUR," have been employed to highlight regions of interest (ROIs), thereby improving image clarity and aiding in more precise diagnostic assessments [8].

2.3 Data Augmentation Strategies

Deep-CNN models typically need extensive datasets to perform effectively and avoid overfitting—a situation where the model becomes too tailored to the training data and fails to generalize to new, unseen inputs. This issue is particularly prevalent in domains like medical image analysis, where access to vast datasets is often restricted due to privacy concerns and limited availability. One widely adopted strategy to mitigate this challenge is data augmentation. This technique involves applying various transformations to existing images, thereby increasing both the volume and diversity of the training data. By doing so, it enables the development of more robust and generalizable deep learning models [9].To address issues of overfitting and underfitting, data augmentation techniques are commonly employed to increase the volume and diversity of mammogram datasets [3]. In certain studies, researchers applied seven distinct augmentation methods to enhance preprocessed mammograms, leading to a substantial expansion of the dataset. For instance, one dataset originally containing 1,442 images was augmented to a total of 11,536 mammograms [3].Data augmentation plays a vital role in medical image analysis, particularly in scenarios where datasets are small—an issue often encountered in biomedical imaging due to the limited number of patient cases. Techniques such as image rotation [10], scaling, and shifting are commonly used to artificially expand the dataset and introduce variability [11]. These transformations help reduce overfitting, enhance model generalization, and improve performance on previously unseen data. In mammography, both conventional augmentation methods and advanced deep learning–based augmentation strategies have been successfully applied to increase the robustness of diagnostic models [12].

3. Deep Learning Architectures for Breast Cancer Prediction

3.1 Custom CNN Architectures

Custom CNN models have been developed for the data classification problem of cancer from medical mammogram data. For instance, one study explored the MIAS database that including 332 digital mammogram images from various females, which were in two directories such as augmented and preprocessed, and provided into a custom CNN model with the aim of differentiating non-cancerous tissues from cancerous ones via classification mechanism. The custom CNN algo achieved testing accuracy, AUC, precision, recall, and F1 scores of 0.9362, 0.9407, 0.9200, 0.8025, and 0.8572 respectively, with minimum or less Overadaptation. This research proposed a new custom CNN model for better breast cancer classification using raw mammograms and focused on the importance of computer-aided detection models in the early stages of diagnosis of breast cancer[13].

Other research has presented novel ways of breast cancer detection using CNN and mammogram imaging systems to accurately classify mammogram images of tumors into benign (noncancerous) and malignant (cancerous). Some custom models are created in resemblance with established architectures like the VGG16 model[14].

3.2 Transfer Learning and Pre-trained Models

Several studies have evaluated adjusted convolution neural algo originally developed for other reasons. For example, one study evaluated six pre-trained CNNs such as VGG16, VGG19, MobileNetV2, ResNet50, DenseNet201, and InceptionV3 to determine which model yielded the better compatibilities. Based on the findings, a BreastNet18 algo. was proposed using VGG16 as the initial base since VGG16 performed with the maximum accuracy. An ablation study was conducted on BreastNet18 to assess its robustness and enhance accuracy. To improve image quality and reduce artifacts, different image preprocessing methods were applied using carefully selected parameter settings. The results suggested that accuracy improvement could be obtained through image pre-processing techniques, augmentation, and ablation study [3].

A comprehensive, two-phase experimental approach was adopted to evaluate the consistency and effectiveness of diagnostic classification in mammography analysis. The first phase involved five distinct deep convolutional neural networks (CNNs), each of which was both pre-trained on large-scale datasets and fine-tuned in an end-to-end manner using domain-specific mammogram images. These networks were assessed for their ability to accurately classify breast masses, demonstrating strong baseline performance.

In the second phase, deep features extracted from the aforementioned CNN models were utilized to train a Support Vector Machine (SVM) classifier. This hybrid approach—combining CNN-based feature extraction with the discriminative power of SVM—yielded exceptional classification results, outperforming the individual CNN models alone.

The process confirmed that integrating traditional machine learning algorithms with deep learning features can enhance diagnostic accuracy.

The experiments underscored the critical role of rigorous data preparation, including cleaning, preprocessing, and augmentation, in improving model reliability and generalization. The refined model achieved an impressive classification accuracy of 97.8%, surpassing contemporary leading-edge methods available at the time of the study.

These findings have significant implications for clinical practice. By integrating such AI-driven models into conventional pathological workflows, the diagnostic process can be streamlined, reducing the cognitive load and manual effort required by pathologists. Consequently, this approach supports more accurate and efficient analysis of mammographic images, aiding in early detection and improved patient outcomes [15].

Transfer learning has been used to increase the accuracy of early stage detection of breast cancer on medical images using the CNN approaches. Studies have shown that models like VGG-16, VGG-19, and ResNet-50 can achieve varying levels of accuracy, with ResNet-50 sometimes producing the maximum level of validation accuracy (71%) compared to other algo (VGG-16=64%, VGG-19=61%). It has been suggested that using more image datasets may create better accuracy[6].

3.3 Advanced Architectures

Advanced architectures have been developed to increase classification accuracy. One study presented a reliable and precise work for classifying this type of cancers using mammographic imaging data. Concatenated Convolutional-Neural-Models were designed based on following models: two by transfer learning and one entirely from basics. This approach can reduce misclassification of lesions from medical unstructured data. Bayesian based optimization method performs Parameter Sweeping of the neural networks, and synthetic data generation refines the model by more than one training data samples. Investigation of this model's accuracy uncovered that it can accurately predict disease with 97.26% accuracy in binary cases and 99.13% accuracy in multi-classification cases in healthcare. These results demonstrated a 16% increment in accuracies in multi-classification compared to various studies using the same dataset. Furthermore, an accuracy increment of 6.4% was achieved after Parameter Sweeping and adjustment from part data. Using a sequence of three distinct neural networks—both from scratch and via transfer learning—enables the extraction of diverse and meaningful features, ensuring that important information is retained and leading to more accurate diagnoses [16].

Another study proposed an enhanced digital neuron-based methods for breast cancer classification in mammogram data, leveraging multiple Convolutional Neural Networks (CNNs), Best-Worst Multi-Attribute Decision-Making (BWM-MCDM) feature selection, and feature concatenation. Eight pretrained CNN models (InceptionV3, ResNet50, MobileNet, VGG16, VGG19, Xception, DenseNet169, and EfficientNetB7) were fine-tuned on an augmented Mammographic Image Analysis Society (MIAS) dataset. Features for each image in each trained model were extracted and those features with zero variance were eliminated. These extracted features underwent two feature selection methods, BWM-MCDM and Mutual Information (MI) feature selection. All features from BWM-MCDM were concatenated, features from MI were concatenated, and the naïve (all) features—without feature selection but with zero variance removed—were also concatenated. A total of three sets of concatenated features were gathered and each used to train a deep neural model classifier. Results to proposed the framework with BWM-MCDM demonstrated a reduction in feature dimensionality while retaining critical information. From five cross-folds, the mean for each performance metric achieved an accuracy of 98.74%, F1-score of 98.73%, ROC-AUC of 99.80%, and MCC of 95.31%, all surpassing the eight individual CNN models' performance[17].

3.4 Multi-View and Ensemble Approaches

Conventional-CADx systems for this type of cancer typically use single-view information of data to assist cancer professionals. Previous work has based on more robustly than one view. Existing multi-view based CADx systems casually employ only two views, such as Medio-Lateral-Oblique and Cranio-Caudal methods. The data aggregation of the double-views has proven effective for mammogram medical classification, which cannot be analyzed by single-view information. However, combination of the information of multiple views of medical images can further increase the compatibility of classification-model. Some studies have proposed Multi-View parameter combination (MVFF) based CADx systems using representation merging techniques of IV-views for mammographic imaging classification [2].

The Systems based on Comprehensive-computer-aided-diagnosis for mammography typically involve three sequential stages. The first stage distinguishes between normal and abnormal mammograms, the second detects specific anomalies such as masses or calcifications, and the final stage classifies the abnormalities as benign or malignant. Convolutional-Neural-based models are often used for dimensionality reduction from each mammographic view independently, after which the features are integrated at a later stage to form a final prediction.

Studies have demonstrated that using all four standard mammographic views with multi-view feature fusion (MVFF), especially when combined with data augmentation, leads to better classification performance compared to models based on a single view. For example, one multi-view system achieved 0.932 value for classifying masses and calcifications in the AUC-ROC curve, 0.84 for distinguishing benign from cancerous lesions, and 0.93 for identifying abnormal versus normal findings—outperforming all single-view approaches [2].

4. Development of Prediction Model:

This research provides practical approaches to make prediction from scratch using mammography imaging data for breast cancer prediction on their early stages. Deep learning provides the effective services to classified images effectively. For predicting breast cancer, we are using medical imaging data and various deep learning models such as custom CNN, pre-training transfer learning models such VGG16, DenseNet121, EfficientNetB3 and so on. We can customize layer and dense value to improve efficiency based on dataset in custom-CNN and on other pre-trained CNNs models.

4.1 Python-Based Development Environment

The model was developed entirely in Python due to its extensive support for machine learning and image processing tasks. Several well-established modules were utilized to facilitate various components of the research:

- **TensorFlow and Keras** were employed as the primary frameworks for building and training convolutional neural networks (CNNs). These libraries offer high-level APIs for rapid prototyping, as well as low-level control when fine-tuning model architecture and hyperparameters.
- **OpenCV and NumPy** played a central role in image manipulation and preprocessing operations. OpenCV was used for tasks such as image reading, resizing, and basic transformations, while NumPy enabled efficient numerical computations and array handling.
- **Matplotlib and Seaborn** were utilized to create visual representations of the data and model performance metrics. These tools were especially helpful in generating plots for loss curves, accuracy trends, and class distribution.
- **Scikit-learn** provided utilities for splitting the dataset, applying stratified sampling to maintain class balance, and computing standard evaluation metrics like precision, recall, F1-score, and confusion matrices.

This software stack provided an adaptable and robust environment for handling mammographic data and implementing advanced deep learning architectures suited for medical image classification.

4.2 Dataset Description and Experimental Setup

The dataset used in this study was the Mammogram Mastery dataset, which includes a total of 745 grayscale mammographic images. Each image is annotated with one of two possible class labels:

- Cancer: indicating the presence of malignant tissue.
- Non-Cancer: indicating the absence of malignancy.

This dataset is particularly suitable for binary classification tasks aimed at distinguishing pathological samples from healthy ones.

The practical setup for the experiment included:

- Hardware: The development and training were performed on a workstation equipped with a GPU-enabled environment to accelerate deep learning computations.
- Software: All processing was conducted within a Python environment, using Jupyter Notebook as the primary interface for running experiments and visualizing results.

4.3 Image Preprocessing and Data Handling

Previously feeding the data into the neural models, several data wrangling steps were implemented to standardize input formats and enhance learning efficiency:

1. **Image Resizing:**All data were reshaped to a fixed resolution of 128×128 pal to ensure reliability in input shape across the dataset. This dimension strikes a balance between retaining meaningful features and reducing computational complexity.

2. **Pixel Normalization:**Each image was normalized by scaling values of pal to a range in the form of [0, 1]. This was achieving by dividing pal intensities by 255, the maximum possible grayscale value. Normalization helps advance and hasten the training procedures by preventing large numerical gradients.

3. **Label Encoding:**The categorical class labels (Cancer, Non-Cancer) were converted into a binary one-hot encoded format. This transformation is required for binary classification with categorical cross-entropy loss.

4. **Dataset Splitting:**The datas was divided into 2 sections such as training and rest is testing subsets using an 80:20 via train_test_split from Scikit-learn. Stratified-sampling-data was applied to maintain the originality of class distribution in both sets, ensuring that the model receives balanced representation during training and evaluation.

5. **Class Distribution Validation:**A visual inspection was conducted to verify the balance between the two classes. Bar plots and pie charts were used to confirm that no significant class imbalance was present, which could otherwise bias the learning process.

4.4 Models Used

The following deep learning models were implemented and evaluated for performance comparison:

1. **Custom CNN:** This model served as the baseline for the experiment. It included two network layers, each go along with by dropout and max pooling network layers to reduce the risk of overfitting. After feature extraction, the architecture used a flattening layer, followed by more than one fully connected neural-layers. The last network layer work with softmax activation for binary classification. This model was kept lightweight to ensure faster training and lower calculational workload demands, making it compatible for resource-constrained environments.

2. **VGG16:** Implemented through transfer learning, this model used pretrained weights from the ImageNet dataset. The original top layers were replaced with an average global pooling network layer and custom dense layers tailored for binary prediction. VGG16's structured design and rich pretrained features enabled strong generalization, even when trained on limited mammogram data.

3. **ResNet50:** This architecture was trained from scratch to examine its standalone performance. It incorporates residual blocks that facilitate efficient gradient flow via skip connections, helping mitigate the vanishing gradient issue common in deeper networks. Although its design supports deeper architectures, the model's performance was limited due to the small dataset and the lack of pretraining.

4. **DenseNet121:** Also trained from the ground up, DenseNet121 links each layer to every other preceding layer, enabling high feature reuse and effective gradient propagation. Despite its efficient architecture for image classification, the absence of pretrained weights and the limited dataset size restricted its performance in this experiment.

5. **EfficientNetB3:** Chosen for its optimal trade-off between accuracy and computational cost, EfficientNetB3 was trained without transfer learning. It applies fundamental scaling to balance network depth and width, with resolution. However, its ability to perform well was constrained by the small dataset and lack of pretrained weights, which hindered full utilization of its architecture.

All models were trained under the same conditions: with 0.0001 learning rate for five epochs in the Adam optimizer. Identical preprocessing procedures were applied to ensure fair comparison across architectures.

4.5 result and discussion:

To evaluate the classification capabilities of different deep learning architectures, model accuracy was assessed using the designated test dataset. The comparison included five distinct models: a custom-built Convolutional Neural Network (CNN), along with four widely recognized pre-trained architectures—VGG16, ResNet50, DenseNet121, and EfficientNetB3.Among these, VGG16 demonstrated the highest test accuracy, reaching 92.11%. This notable performance is attributed to the application of transfer learning, where the algo. launched with pre-trained weights originally optimized on the dataset with ImageNet. By leveraging feature representations learned from millions of natural images, the model was able to effectively generalize to mammographic data, even with a limited number of training examples.Transfer learning offers several advantages in medical imaging contexts: it reduces the need for large labeled datasets, shortens training time, and improves convergence stability. The success of VGG16 in this case highlights the effectiveness of using established architectures fine-tuned for domain-specific tasks such as mammogram classification.The custom CNN model performed competitively with an accuracy of 89.14%, which is notable given its simpler architecture and fewer parameters compared to the large-scale pre-trained networks. This suggests that the custom CNN was sufficiently capable of learning meaningful features specific to the mammogram images, despite its relative simplicity.On the other hand, the ResNet50, DenseNet121, and EfficientNetB3 models all showed similar, lower accuracies around 82.38%. These models trained from scratch without pre-trained weights, which could be a significant factor in their underperformance. Training very deep architectures from scratch typically requires large numbers of data with labels to avoid overfitting and to achieve effective feature extraction. The relatively small size of the mammogram dataset (745 images) likely limited the ability of these complex models to generalize well, resulting in comparatively lower accuracy.Loss function comparison: The validation loss trends reflected the accuracy outcomes, with the VGG16 and custom CNN models exhibiting consistently lower validation losses by the end of training. Lower loss values indicate that the models were better at minimizing the discrepancy between predicted outputs and true labels, thereby confirming their better fit to the data.In contrast, the higher validation losses for ResNet50, DenseNet121, and EfficientNetB3 suggest that these models struggled to optimize effectively, potentially due to overfitting or insufficient data coverage for the large number of parameters they contain. These findings emphasize the importance of considering dataset size and model complexity when selecting architectures for medical image classification tasks.The AUC-ROC curves were plotted for all models. VGG16 showed the highest AUC, indicating robust discriminative ability between classes. AUC scores reinforced the superiority of VGG16 in this task.

5. Evaluation Metrics and Performance

5.1 Common Evaluation Metrics

Accuracy and F1-score, derived from confusion matrices, are commonly utilized performance metrics in binary classification problems. Nonetheless, these measures can present misleadingly high evaluations, when we deal with imbalanced datasets. In contrast, the Correlation Coefficient is grant as a dependable metric. It yields a high value only when the classifier plays well across all parameters of the evolution matrix such as true positives and negatives with false positives and negatives—while also accounting for the relative parameters of both instances like positives and negatives in the dataset [18].Multiple performance metrics are employed for validation, such as accuracy, precision, sensitivity, specificity, Curve of AUC-ROC, and F1-score. In one study, the reported performance outcomes were: accuracy of 86.21%, precision of 85.50%, sensitivity of 85.60%, specificity of 84.71%, F1-score of 88%, and an AUC value of 0.89 [19].Various metrics are employed to assess medical image segmentation performance, including the Dice coefficient, Jaccard index, sensitivity, specificity, Rand index, ROC curves, Cohen's Kappa, and Hausdorff distance. It is also important to consider challenges such as class imbalance and potential biases in both statistical analysis and result interpretation. Guidelines for standardized image segmentation of medical assessment have been proposed to improve assessment quality, comparability with reproducibility in the research field[20].

5.2 Comparative Analysis of Model Performances

To investigate the diagnostic effectiveness of various deep learning approaches in mammogram classification, we implemented and evaluated five models: a custom Convolutional Network, VGG16 with transfer learning, ResNet50, DenseNet121, and EfficientNetB3. All models were trained under identical preprocessing conditions on a curated dataset consisting of 745 mammographic images, equally representing cancerous and non-cancerous cases.Among the tested models, VGG16 found the highest classification compatibility, accuracy during test 92.11%. This can be attributed to its use of transfer machine learning, where pretrained weights on ImageNet were fine-tuned to the mammogram dataset. The model showed consistent learning progression and low validation loss, indicating good generalization despite the relatively limited dataset size.The custom CNN model also performed well, reaching a test accuracy of 89.14%. Designed with a simpler architecture, it demonstrated that lightweight models can still yield competitive performance when appropriately regularized and trained on well-processed data. This makes it an attractive option for resource-constrained environments where inference time and computational efficiency are critical.On the other hand, the deeper models—ResNet50, DenseNet121, and EfficientNetB3—which were trained from scratch, exhibited slightly lower test accuracies, each around 82.38%. These models, while architecturally advanced, did not benefit from pretraining in this experiment. Their performance may have been limited by the relatively small dataset size, leading to underfitting or slow convergence, which is a common challenge when training deep networks without transfer learning in medical imaging tasks.These results are consistent with previous studies. For instance, Montaha et al. [3] reported superior performance using a fine-tuned VGG-based model (BreastNet18) with a test accuracy of 98.02%, highlighting the strength of pretrained architectures in breast cancer detection. Similarly, Mahmood et al. [15] demonstrated that combining ConvNet features with support vector machines (ConvNet+SVM) significantly outperformed standard deep models such as ResNet50 and DenseNet121, especially on small datasets. Moreover, Fulton et al. [21] emphasized the effectiveness of deep vision approaches in achieving 97% classification accuracy and leveraging gradient-based attention mechanisms for clinical decision support.Although EfficientNetB3 and DenseNet121 have shown strong potential in other studies with larger datasets or advanced augmentation pipelines, their performance in our setup remained constrained. This further reinforces the notion that model complexity must be matched with dataset scale and quality to avoid diminishing returns.In conclusion, the VGG16 model emerged as the most effective architecture in our comparative analysis, offering a strong balance between accuracy and training stability. The custom CNN proved to be a viable lightweight alternative, while deeper models require either larger datasets or transfer learning to realize their full potential in breast cancer image classification.

6. Current Challenges and Limitations

6.1 Dataset Limitations

Training deep learning models on large datasets generally enhances their predictive accuracy. However, in the biomedical field, datasets are often limited in size due to small patient cohorts and privacy-related restrictions [10]. The conclusiveness of supervised deep learning models is largely influenced by the volume of labeled training data, which typically requires manual annotation by expert radiologists—a process that is time-consuming with labor-intensive. As a result, publicly available biomedical image datasets are often small, and strict legal frameworks further restrict access to extensive medical imaging data [12].

This data scarcity increases the likelihood of overfitting, thereby reducing the model's compatibilities to generalize effectiveness on unseen cases [12]. Although breakthroughs in image recognition have largely stemmed from the accessibility of expansive, annotated datasets and neural architectures, replicating such progress in medical imaging remains difficult due to the absence of datasets comparable to ImageNet in terms of size and quality [25]. In mammography-focused decision support systems, a major barrier to reproducibility is the lack of standardized evaluation datasets. Many CADx and CADe systems for breast cancer detection rely on private datasets or undefined portions of public databases, hindering performance benchmarking and replication of previous research findings [5].

One major challenge in breast cancer prediction using CNNs is the limited availability of large, annotated mammography datasets. Most medical datasets are small because of privacy constraints and the requirement of expert annotation, which can hinder model generalization and increase the risk of overfitting.

6.2 Technical and Methodological Challenges

The shortcomings of the systems for aided-detection based on traditional architecture in images—especially their limited accuracy in early breast cancer detection and the serious implications of false diagnoses—have led researchers to explore more robust alternatives. Neural learning, particularly Neural models, a class of deep learning models, have emerged as a powerful and promising approach for solving complex image analysis tasks. Techniques such as transfer adaptation learning, data expansion, batch simplification, and dropout have been widely adopted to minimize overfitting and enhance the generalization capacity of CNN models. Despite these advancements, there remain significant research challenges that the scientific community must continue to address [26].Mammographic breast cancer diagnosis (MBCD) is a persistent and complex problem, and numerous CNN-based models have been developed to support clinical decision-making. These models typically fall into three main categories: (1) shallow or adapted CNN architectures designed to reduce computational complexity and training time; (2) pre-trained networks leveraged through transfer learning and fine-tuning for domain-specific tasks; and (3) hybrid approaches where CNNs serve as feature extractors, while classification is performed using traditional machine learning algorithms to distinguish between benign and malignant cases [27].A lot of researchers have proposed approaches, techniques, and tools based on ML-driven for assessing medical images, but these approaches have produced detection and interpretation errors resulting in false-positive with negative cases when used in the real-time scenarios. These challenges can be effectively addressed by applying robust image preprocessing techniques, which help generate high-quality training data for deep convolutional neural networks (Deep-CNNs)[8].From a technical perspective, training deep models on limited data often requires careful use of transfer learning, regularization, and augmentation techniques. However, performance still varies significantly across architectures, and reproducibility remains a concern due to inconsistent evaluation protocols and dataset partitioning.

6.3 Clinical Integration Challenges

Despite substantial advances in medical science, this type of cancer remains the one of the most casual cause of cancer-related deaths globally. While early stage detection plays a vital role in lowering mortality rates, identifying breast abnormalities at an early stage remains challenging. To address this, various diagnostic modalities—such as mammography, ultrasound, and thermography—have been developed to aid in breast cancer screening. Image processing techniques combined with artificial intelligence have the potential to support radiologists in detecting abnormalities in breast imaging more efficiently and accurately, thereby enhancing diagnostic workflows [28].Computer-Aided Diagnosis (CAD) systems applied to mammographic imaging are regarded as among the most efficient and accessible methods for detecting breast cancer. Accurate identification of abnormalities through CAD can significantly contribute to lowering breast cancer-related mortality. Key indicators such as masses and clusters of microcalcifications serve as critical early signs, aiding in the identification and diagnosis of breast cancer during its initial stages [29].However, neural network-based mammographic images breast cancer diagnosis (MBCD) is still in its developmental phase, and significant progress is required before deep learning technologies can be fully integrated into clinical workflows. Achieving this goal will demand that these tools provide tangible value not only to medical researchers but also to engineers and specialists working on intelligent diagnostic systems [27]. Clinically, integrating AI-based diagnostic tools into existing radiology workflows is non-trivial. Issues like explainability, clinician trust, and regulatory compliance must be addressed before widespread adoption is feasible.

Table 1. Models and Accuracy

Model	Accuracy (%)
CNN	89.14
VGG16	92.11
ResNet50	82.38
DenseNet121	82.38
EfficientNetB3	82.38

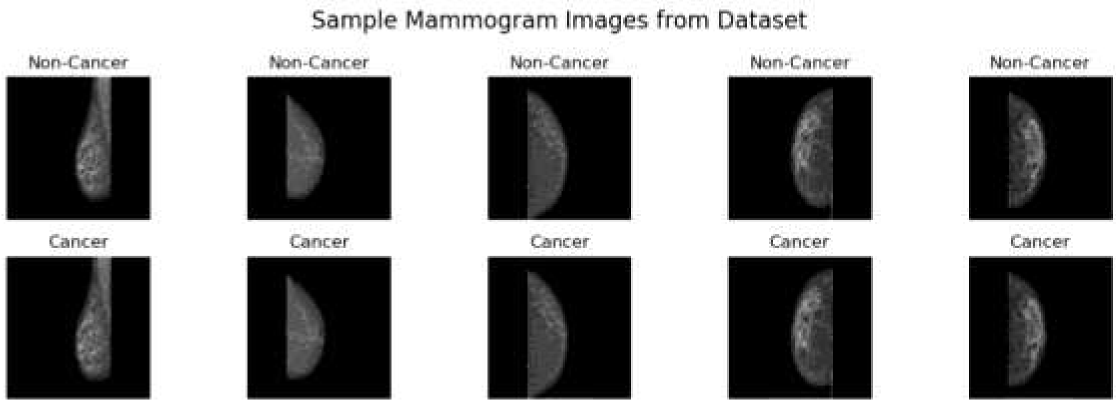

***Figure 1**: Sample of Mammogram Datatset*

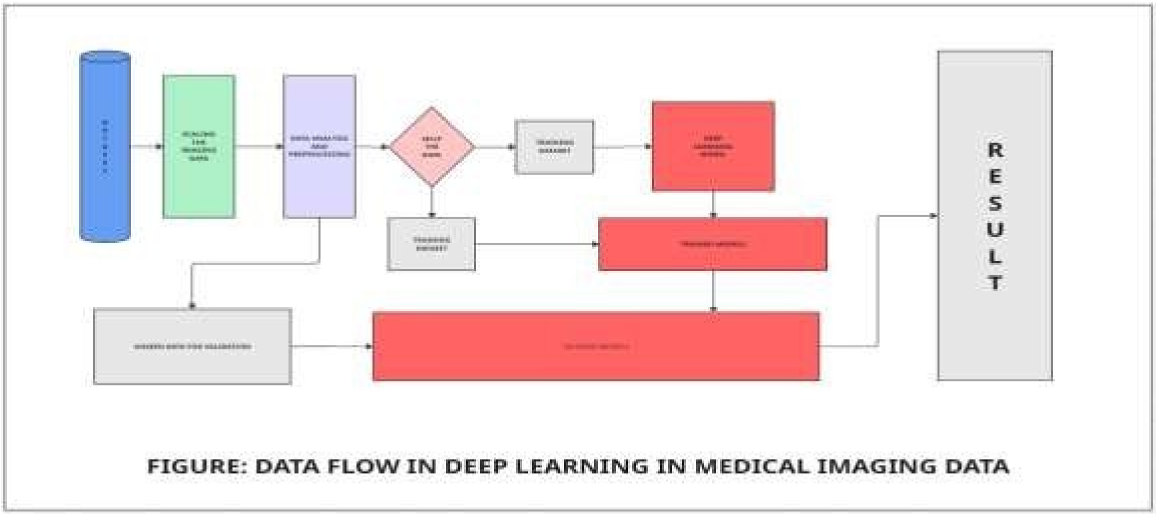

Figure 2: *workflow of deep learning in heart CVD prediction*

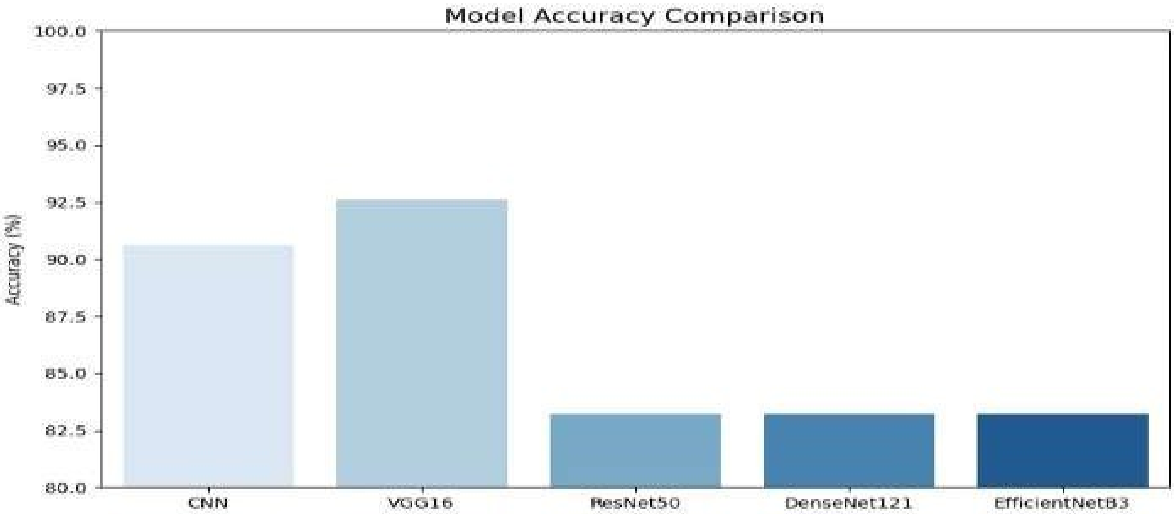

Figure 3*: accuracy comparison of models*

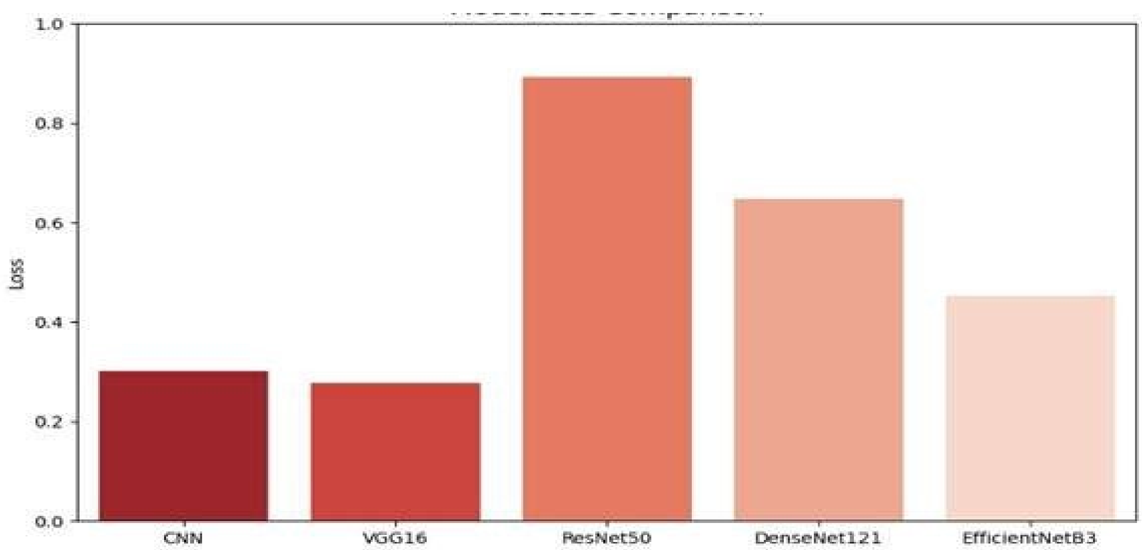

Figure 4: *Loss function comparison*

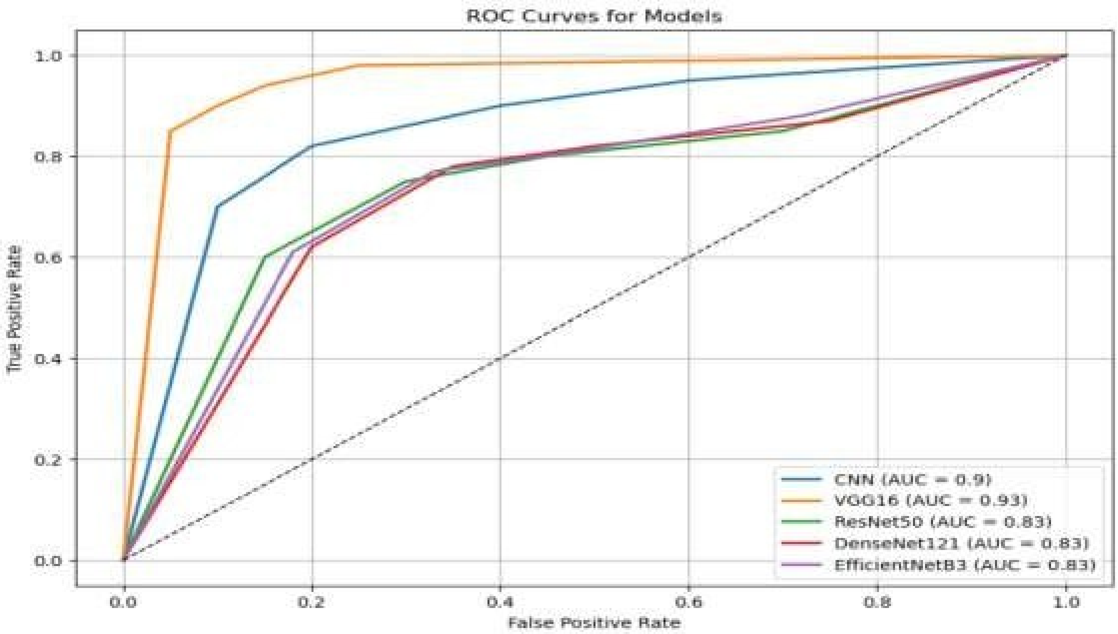

Figure 5: *AUC-ROC curve for models*

7. Future Directions and Conclusion

7.1 Emerging Trends and Research Gaps

Deep learning models employing end-to-end training strategies have shown effectiveness in early breast cancer detection by efficiently utilizing mammography images within computer-aided diagnostic systems. Future advancements are expected to focus on enhanced preprocessing techniques—such as refined contrast enhancement—to better define edge details in mammographic scans. The design of specialized architectures, including the transferable texture convolutional neural network (TTCNN), offers potential for improving classification outcomes. One promising approach involves replacing traditional pooling layers with energy layers to extract richer texture features from convolutional outputs. Continued evaluation of deep features across different CNN models and identifying optimal layers for improved classification accuracy will remain crucial. Moreover, the adoption of convolutional sparse image analysis for feature extract and entropy-guided techniques for feature selection appears promising in advancing diagnostic performance. These approaches have demonstrated superior performance compared to prior methods and offer maximum potential to increase medical tools to minimize false positive with negative results via screening imaging mammography data [1].Future investigations may prioritize incorporating attention mechanisms and feature fusion strategies to enrich the semantic representation of features extracted by CNNs. Innovative data augmentation techniques, such as Random Center Cropping (RCC), offer potential by expanding dataset diversity while maintaining image resolution and central content. Furthermore, adapting the downsampling scale within network architectures to better accommodate low-resolution images presents another avenue for optimization. These technical refinements have demonstrated potential in enhancing baseline model performance, yielding high accuracy and AUC scores in breast cancer classification tasks [30].Future research could enhance the effectiveness of advanced architectures such as ResNet50, DenseNet121, and EfficientNetB3 by employing techniques like data augmentation, fine-tuning of pre-trained models, and expanding dataset size.

7.2 Conclusion

This comprehensive literature review has examined the current study of research on breast cancer prediction based on mammography data using deep learning. The review has highlighted various methodologies, architectures, and evaluation approaches, with a particular focus on CNN-based models.Numerous studies in this domain focus on developing accurate breast image classification techniques that minimize error rates. Preprocessing techniques such as unwanted region removal, quality enhancement, and cancerous lesion highlighting through various methods have proven beneficial. Data augmentation has been widely applied to increase dataset diversity and mitigate overfitting issues. The evaluation of various CNN architectures, including fine-tuned pre-trained models like VGG16, has demonstrated the effectiveness of these approaches. The development of specialized models based on established architectures, such as BreastNet18, along with ablation studies and parameter optimization, has led to significant improvements in classification accuracy. Recent progress indicates that classification accuracy can be significantly enhanced by integrating image pre-processing, data augmentation, and optimized neural network architectures [3].Convolutional deep learning techniques have proven effective in identifying features across different tissue densities and distinguishing between normal and potentially abnormal regions in mammographic images. Various experimental strategies, such as end-to-end deep CNNs utilizing pre-trained and fine-tuned models, along with hybrid systems that integrate deep feature extraction with conventional classifiers like SVM, have yielded promising results. The consistent role of data cleaning, preprocessing, and augmentation in enhancing the accuracy of mass detection has been well-documented. These techniques are increasingly being integrated into traditional pathology workflows, offering the potential to ease the workload on pathologists by supporting more reliable clinical outcome predictions based on mammographic image analysis [15].Custom CNN models, when properly designed and trained on augmented and preprocessed mammogram data, have achieved high accuracy, AUC, precision, recall, and F1 scores while minimizing overfitting. These models, along with computer-aided detection systems more broadly, represent important tools for the early stage detection to make prevent of breast cancer [13].In conclusion, deep neural approaches for breast mammographic cancer identification based on mammography data have shown remarkable progress and hold great promise for improving early detection and diagnosis. Continued research in this field, addressing current limitations and exploring emerging trends, is likely to lead to even more effective and clinically valuable systems in the future. The results demonstrate the advantage of transfer domain learning for medical imaging activities, especially when working with limited datasets. Utilizing pre-trained weights, the VGG16 model offered a robust base for effective feature extraction, contributing to improved accuracy and reduced loss. In contrast, the custom CNN demonstrated that carefully crafted, compact architectures can also achieve high performance, without the complexity associated with deeper networks.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Maqsood, Sarmad, Damaševičius, R., and Maskeliūnas, R.. 2022. "TTCNN: A Breast Cancer Detection and Classification towards Computer-Aided Diagnosis Using Digital Mammography in Early Stages". Applied Sciences. https://doi.org/10.3390/app12073273

[2] Khan, Hasan Nasir, Shahid, A. R., Raza, B., Dar, A., and Alquhayz, H.. 2019. "Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network". IEEE Access. https://doi.org/10.1109/ACCESS.2019.2953318

[3] Montaha, Sidratul, Azam, S., Rafid, A. R. H., Ghosh, Pronab, Hasan, Md Zahid, Jonkman, M., and Boer, F. De. 2021. "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images". Biology. https://doi.org/10.3390/biology10121347

[4] Ribli, D., Horváth, A., Unger, Z., Pollner, P., and Csabai, I.. 2017. "Detecting and classifying lesions in mammograms with Deep Learning". Scientific Reports. https://doi.org/10.1038/s41598-018-22437-z

[5] Lee, R. S., Gimenez, Francisco, Hoogi, A., Miyake, K., Gorovoy, Mia, and Rubin, D.. 2017. "A curated mammography data set for use in computer-aided detection and diagnosis research". Scientific Data. https://doi.org/10.1038/sdata.2017.177

[6] Susilo, Arief Broto and Sugiharti, E.. 2021. "Accuracy Enhancement in Early Detection of Breast Cancer on Mammogram Images with Convolutional Neural Network (CNN) Methods using Data Augmentation and Transfer Learning". Journal of Advances in Information Systems and Technology. https://doi.org/10.15294/jaist.v3i1.49012

[7] Jabeen, Kiran, Khan, Muhammad Attique, Balili, Jamel, Alhaisoni, Majed, Almujally, N., Alrashidi, Huda, Tariq, U., and Cha, Jaehyuk. 2023. "BC2NetRF: Breast Cancer Classification from Mammogram Images Using Enhanced Deep Learning Features and Equilibrium-Jaya Controlled Regula Falsi-Based Features Selection". Diagnostics. https://doi.org/10.3390/diagnostics13071238

[8] Beeravolu, A. R., Azam, S., Jonkman, M., Shanmugam, Bharanidharan, Kannoorpatti, K., and Anwar, A.. 2021. "Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN". IEEE Access. https://doi.org/10.1109/ACCESS.2021.3058773

[9] Shorten, Connor and Khoshgoftaar, T.. 2019. "A survey on Image Data Augmentation for Deep Learning". Journal of Big Data. https://doi.org/10.1186/s40537-019-0197-0

[10] Ragab, D., Sharkas, M., Marshall, S., and Ren, Jinchang. 2019. "Breast cancer detection using deep convolutional neural networks and support vector machines". PeerJ. https://doi.org/10.7717/peerj.6201

[11] Alruwaili, Madallah and Gouda, Walaa. 2022. "Automated Breast Cancer Detection Models Based on Transfer Learning". Italian National Conference on Sensors. https://doi.org/10.3390/s22030876

[12] Oza, P., Sharma, Paawan, Patel, Samir B., Adedoyin, F., and Bruno, Alessandro. 2022. "Image Augmentation Techniques for Mammogram Analysis". Journal of Imaging. https://doi.org/10.3390/jimaging8050141

[13] Mamun, Rafsan Al, Rafin, Gazi Abu, Alam, Adnan, and Sefat, Md. Al Imran. 2021. "Application of Deep Convolution Neural Network in Breast Cancer Prediction using Digital Mammograms". 2021 2nd International Informatics and Software Engineering Conference (IISEC). https://doi.org/10.1109/iisec54230.2021.9672368

[14] Angane, Ronil, Bhogale, Gaurij, Lanjekar, Sejal, Gholkar, Aditya, and Chaudhari, R.. 2022. "Breast Cancer Analysis using Convolutional Neural Network". 2022 International Conference on Breakthrough in Heuristics And Reciprocation of Advanced Technologies (BHARAT). https://doi.org/10.1109/bharat53139.2022.00037

[15] Mahmood, T., Li, Jianqiang, Pei, Yan, and Akhtar, F.. 2021. "An Automated In-Depth Feature Learning Algorithm for Breast Abnormality Prognosis and Robust Characterization from Mammography Images Using Deep Transfer Learning". Biology. https://doi.org/10.3390/biology10090859

[16] Alshayeji, M. and Al-Buloushi, Jassim. 2023. "Breast Cancer Classification Using Concatenated Triple Convolutional Neural Networks Model". Big Data and Cognitive Computing. https://doi.org/10.3390/bdcc7030142

[17] Pulvera, Eusib Vincent J. and Lao, D.. 2024. "Enhancing Deep Learning-Based Breast Cancer Classification in Mammograms: A Multi-Convolutional Neural Network with Feature Concatenation, and an Applied Comparison of Best-Worst Multi-Attribute Decision-Making and Mutual Information Feature Selections". 2024 9th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS). https://doi.org/10.1109/ICIIBMS62405.2024.10792816

[18] Chicco, D. and Jurman, Giuseppe. 2020. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation". BMC Genomics. https://doi.org/10.1186/s12864-019-6413-7

[19] Wang, Xiaomei, Ahmad, Ijaz, Javeed, D., Zaidi, Syeda Armana, Alotaibi, F. M., Ghoneim, Mohamed E., Daradkeh, Y., Asghar, Junaid, and Eldin, E. T.. 2022. "Intelligent Hybrid Deep Learning Model for Breast Cancer Detection". Electronics. https://doi.org/10.3390/electronics11172767

[20] Müller, D., Soto-Rey, Iñaki, and Kramer, F.. 2022. "Towards a guideline for evaluation metrics in medical image segmentation". BMC Research Notes. https://doi.org/10.1186/s13104-022-06096-y

[21] Fulton, Lawrence V., McLeod, A., Dolezel, Diane, Bastian, Nathaniel D., and Fulton, Christopher P.. 2021. "Deep Vision for Breast Cancer Classification and Segmentation". Cancers. https://doi.org/10.3390/cancers13215384

[22] Agarwal, Richa, Díaz, Oliver, Lladó, X., Yap, Moi Hoon, and Martí, R.. 2019. "Automatic mass detection in mammograms using deep convolutional neural networks". Journal of Medical Imaging. https://doi.org/10.1117/1.JMI.6.3.031409

[23] Poojary, Ramaprasad, Raina, Roma, and Mondal, A.. 2021. "Effect of data-augmentation on fine-tuned CNN model performance". IAES International Journal of Artificial Intelligence (IJ-AI). https://doi.org/10.11591/IJAI.V10.I1.PP84-92

[24] Wagner, S., Khalili, Nadieh, Sharma, Raghav, Boxberg, M., Marr, C., Back, Walter De, and Peng, Tingying. 2021. "Structure-Preserving Multi-Domain Stain Color Augmentation using Style-Transfer with Disentangled Representations". International Conference on Medical Image Computing and Computer-Assisted Intervention. https://doi.org/10.1007/978-3-030-87237-3_25

[25] Shin, Hoo-Chang, Roth, H., Gao, Mingchen, Lu, Le, Xu, Ziyue, Nogues, Isabella, Yao, Jianhua, Mollura, D., and Summers, R.. 2016. "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning". IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2016.2528162

[26] Abdelhafiz, Dina, Yang, Clifford, Ammar, Reda, and Nabavi, S.. 2019. "Deep convolutional neural networks for mammography: advances, challenges and applications". BMC Bioinformatics. https://doi.org/10.1186/s12859-019-2823-4

[27] Zou, L., Yu, Shaode, Meng, T., Zhang, Zhicheng, Liang, Xiaokun, and Xie, Yaoqin. 2019. "A Technical Review of Convolutional Neural Network-Based Mammographic Breast Cancer

Diagnosis". Computational and Mathematical Methods in Medicine. https://doi.org/10.1155/2019/6509357

[28] sadoughi, F, Kazemy, Zahra, Hamedan, Farahnaz, Owji, Leila, Rahmanikatigari, Meysam, and Azadboni, T. T.. 2018. "Artificial intelligence methods for the diagnosis of breast cancer by image processing: a review". Breast Cancer. https://doi.org/10.2147/BCTT.S175311

[29] Giri, Prannoy and Saravanakumar, K.. 2017. "Breast Cancer Detection using Image Processing Techniques". Oriental journal of computer science and technology. https://doi.org/10.13005/OJCST/10.02.19

[30] Wang, Jun, Liu, Qianying, Xie, Haotian, Yang, Zhaogang, and Zhou, Hefeng. 2020. "Boosted EfficientNet: Detection of Lymph Node Metastases in Breast Cancer Using Convolutional Neural Networks". Cancers. https://doi.org/10.3390/cancers13040661.

[31] Gour, S., & Randa, R. (2023). Digital health revolution: Navigating the growing landscape of machine learning in medical application. International Journal of Artificial Intelligence, Internet of Things and Cloud Computing (IJAIC), 2, 7–14.

[32] Gour, S., Randa, R. (2024). Applications of Emerging Machine Learning Models in Healthcare Industry: A Comprehensive Review. In: Goar, V., Sharma, A., Shin, J., Mridha, M.F. (eds) Deep Learning and Visual Artificial Intelligence. ICDLAI 2024. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-97-4533-3_22