Machine Learning for Automatic Detection of Breast Cancer

TANIA AOUDE, JIHAD DABA
Department of Electrical Engineering,
University of Balamand,
Kalhat, Al-Kurah P.O. box 100 Tripoli,
LEBANON

Abstract: - Breast cancer (BC) affects women with high incidence and death rates. Its diagnosis is achieved through medical imaging, requires the analysis of skilled radiologists, and is subject to human errors. This research utilizes Machine Learning (ML) for automatic BC classification with a focus on clinical context and standards for Artificial Intelligence (AI) in medicine. To limit the effort required by humans in analyzing medical images, Convolutional Neural Networks (CNN) are used to automatically extract features from preprocessed regions of interest (ROI). However, since open access medical data are scarce, Transfer Learning (TL) using pre trained VGG16 was selected and modified to classify the images as normal, benign or malignant. The model's performance is assessed using the appropriate performance metrics for imbalanced datasets. The results are reported based on Balanced Accuracy, Precision, Recall and F1 score. The model resulted in an Accuracy of 86% with a Balanced Accuracy of 73.33%.

Key-Words: - Breast cancer, multiclass classification, ROI, VGG16, image preprocessing, transfer learning.

Received: January 2, 2026. Revised: March 8, 2026. Accepted: May 9, 2026. Published: July 7, 2026.

1 Introduction

Cancer occurs when cells regenerate uncontrollably causing normal cells to be overshadowed. This disease can attack any part of the body and eventually spreads to neighboring areas if left untreated, [1]. According to a study by the World Health Organization (WHO) in 2022, BC, one of the three main types diagnosed, led to 670 000 deaths, [2]. Although BC affects both genders, it most commonly influences females and is one of the leading causes of death. Therefore, early detection is key as is plays a big role in the efficiency of the treatment.

While, there isn't an exact reason why it manifests, several risk factors can increase its incidence. Studies have shown that a family history of BC, mainly mutations in the DNA called BRCA1 and BRCA2, leads to a genetic predisposition to this disease. Additionally, factors related to age and body composition can boost the oncological risk. However, many of the patients that get diagnosed with BC do not have any of these mentioned predictors, [3].

Screening is the act of using tests to detect signs of cancer prior to the development of symptoms. Several imaging techniques are used such as mammography, magnetic resonance imaging (MRI), thermography, ultrasound and histopathological imaging, [4]. Out of these methods, mammograms are most used for early detection and are very sensitive to fatty tissues making them ideal for low density breasts, [4], [5]. Radiologists utilize a standardized scoring system to interpret and classify these imaging results. It is referred to as the breast imaging-reporting and data system (BI-RADS) and consists of marks in the range 0 to 6, [6].

Although various methods are used to diagnose BC, they are time consuming, costly and painful. Additionally, these methods require the analysis of skilled doctors and are frequently subject to human error leading to misdiagnosis, [7], [8]. Since early detection is extremely important for the successful treatment of the disease, reliable methods need to be used to detect its presence. These factors prompted researchers to look into automated systems for BC detection through the use of Computer Aided Diagnostic (CAD) systems and AI.

2 Literature Review

The analysis of medical images is a long tedious task that requires skilled professionals. Therefore, CAD systems were incorporated in the image interpretation process. Studies indicated that CAD software had low sensitivity which often led to the classification of images as false positives. Additionally, they recurrently overlooked certain cancers hence failing to mark the area as needing

further inspection. As AI became more popular, scientists began studying its potential in the health care field, [4].

An exhaustive study in [9] examines several ML and Deep Learning (DL) algorithms in Scopus publications between 2010 and 2019. The authors structured the review based on the imaging techniques mentioned previously followed by the models used for each screening modality. The overall findings revealed mammography as the most used form of data with Support Vector Machine (SVM), Artificial Neural Networks (ANN) and CNN as the leading methods. Moreover, the performance of these models was assessed in [10] and [11] using Wisconsin Diagnostic Breast Cancer (WDBC), a dataset consisting of 30 features and 569 cases with 62.7% benign occurrences and 37.7% malignant. In summary, for ML, Linear Regression (LR) and Random Forest (RF) had the highest accuracy and F1 score of 98.25%. In contrast, for DL, ANN and CNN had accuracies 99.9% and 97.9% respectively with F1 scores of 99% and 97.1%.

Moreover, the research in [12] and [13] focused on TL by examining the performance of several pre trained CNNs both individually and as an ensemble. A study in [14] explored the use of Generative Adversarial Network (GAN) algorithms for synthetic image generation as a form of data augmentation using the Mammographic Image Analysis Society (MIAS) dataset.

However, several limitations need to be addressed, some of which are:

1. The datasets used are too small to efficiently train DL models which require large quantities of good quality data, [4].
2. Not having a proper representation of the distribution in the real world for the data related to the task affects the model's ability to classify important information, [15].

3 Methodology

The WHO offers a guideline on AI for healthcare intended use, [16]. It details things like the importance of clearly explaining the purpose of the developed system as well as documenting the development process in an honest and transparent way.

Hence, our study workflow is detailed in the flowchart in Figure 1. The first step consists of performing data cleaning on the dataset. Secondly, the preprocessing stage involves de-noising through morphological operations and median filtering (MF) followed by contrast enhancement and finally ROI extraction and resizing. The third step is splitting the images into train, validation and test sets. Consequently, the train set is augmented and balanced before being used along with the validation set during the training process. Finally, the isolated test set is used to assess the model's performance.

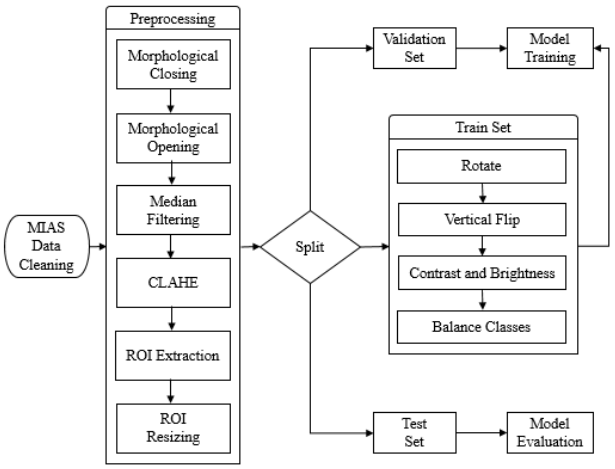

Fig. 1: Process flowchart
Source: created by the authors

3.1 Data Cleaning

The purpose of this classification task is to classify ROI patches extracted from mammograms. Out of the available open access datasets, MIAS was selected as it contained the needed information to perform this task, [17]. It consists of 322 portable gray map (PGM), grayscale images of three categories making it appropriate for multiclass classification problems. Before training any model, the dataset being used must be examined for missing values, duplicates and mistakes. As a result, cases mdb059 (B), mdb216 (M), mdb233 (M), mdb244 (M) were removed due to missing ROI coordinate and radius information. Additionally, mdb212 (B) and mdb214 (B) were also removed due to incorrect coordinates. This reduced the total number of images to 316 with a new distribution of 61 benign, 207 normal and 48 malignant.

3.2 Preprocessing

The preprocessing pipeline was inspired by Balve and Hendrix but adds MF and is explained in [18]. The result of these operations is shown in Figure 2 using case mdb104. Furthermore, Figure 3 details the effect of CLAHE by comparing the histogram of the ROI before and after enhancement.

3.2.1 De Noising and Contrast Enhancement

The image was made binary by applying an Otsu threshold in which the minimum and maximum pixel values of 0 and 255 are specified corresponding to the range of grayscale. The later was essential before performing the morphological operations that were built on binary images. A closing operation consisting of dilation followed by erosion was administered to that binary image to seal minor holes within the breast region through a 5 by 5 ellipse shaped kernel. Afterwards, an opening operation with the reverse order of the closing operation was used on the image to reduce noise using a similar kernel. The outermost contour with the largest area was extracted and drawn onto a black mask with the same size as the image effectively removing the tag in the upper right corner. Finally, this mask was dilated and put on the original image with a kernel of ellipse shape and size 100 by 100 to reduce the clipping of the breast edges. Afterwards, MF with a kernel size of 3 by 3 was implemented to remove salt and pepper noise which is commonly found in mammograms, [19]. Lastly, Contrast Limited Adaptive Histogram Equalization (CLAHE) was employed to enhance the contrast of the image by computing the histogram of each 8 by 8 tile and then equalizing it by evenly redistributing the surplus of the bins that exceed the clip limit of 2.0 to the remaining bins, [20]. The effect of these operations is shown in Figure 2.

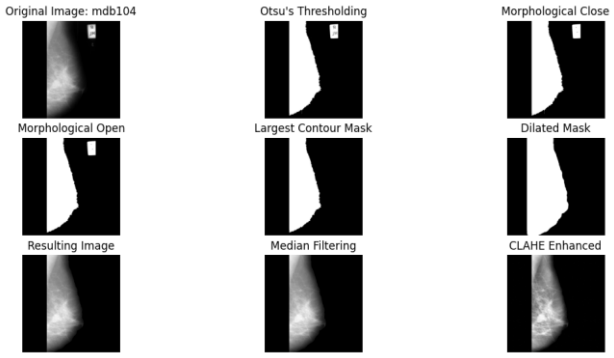

Fig. 2: Results on mdb104
Source: created by the authors

As revealed in Figure 3, the de-noising reduced the appearance of artifacts on the raw ROI and the contrast enhancement redistributed the pixels to a wider range. The raw ROI pixels were originally clustered in a range between nearly 165 and 215. Following the CLAHE enhancement, these pixels became more dispersed to a range roughly between 125 and 235. The appearance of the region became easier to distinguish from the background compared to before. Additionally, the ROI before and after preprocessing showed a reduction in the intensity of the noisy line on the right side of the image.

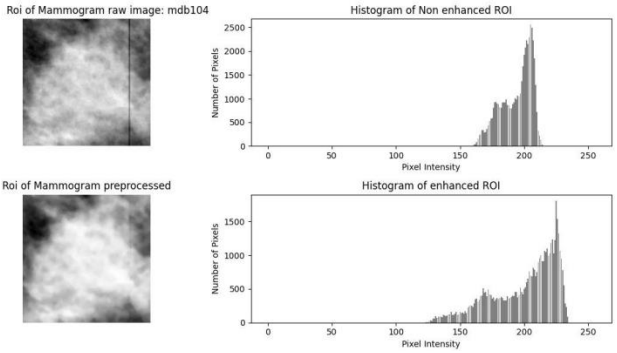

Fig. 3: CLAHE image enhancement result
Source: created by the authors

Although the selected clip limit appeared to work well for most cases that were examined through their histograms, it was not ideal for every image in the dataset. However, this value was selected based on recommendations in the literature to enhance the images without injecting too much noise. Furthermore, experimenting with other values such as 1.0, 3.0 and 4.0 did not show improvement in the redistribution of the surplus to the other bins.

3.2.2 ROI Extraction and Resizing

For normal cases, a central breast area was used to extract an ROI of radius 40 by determining the moments m10 and m01 related to the x, y coordinates of the center and $m00$ which represents the contour area of the largest contour (breast area) extracted earlier. The x and y coordinates were obtained according to the following:

$$x = int\left({}^{m10}/_{m00}\right), \quad (1)$$

$$y = int\left({}^{m01}/_{m00}\right). \quad (2)$$

For benign and malignant cases, the ROI information of the abnormality was extracted from the provided annotations. Since OpenCV uses a coordinate system with a top left origin instead of the bottom left, the y coordinate was modified by subtracting it from the height of the image minus one. Afterwards the top left corner was obtained by subtracting the radius and the bottom right corner by adding the radius as follows:

$$x_{tl} = x - radius, \quad (3)$$

$$y_{tl} = y - radius, \quad (4)$$

$$x_{br} = x + radius, \quad (5)$$

$$y_{br} = y + radius. \quad (6)$$

Then a slice of the de-noised and enhanced image was extracted using this information according to the formula below:

$$ROI = Img\left[int(y_{tl}) : ceil(y_{br}), int(x_{tl}) : ceil(x_{br})\right] \quad (7)$$

Some of the ROIs were located at the edge of the breast which led to the cropping of the solid black region as well. Binary thresholding was implemented with a threshold of 3 to obtain the maximum external contour in these images, effectively removing the black rectangular edge as featured in Figure 4 for case mdb110.

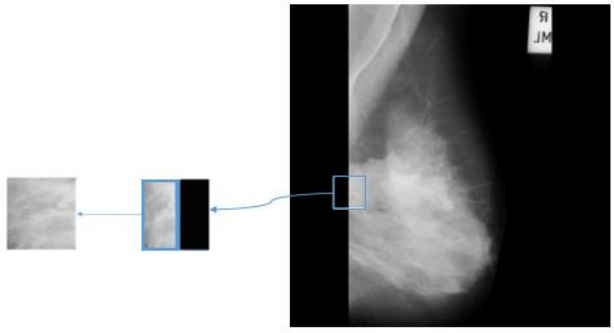

Fig. 4: Cropping the black region of the ROI
Source: created by the authors

These ROIs were resized to 224 by 224 and stored in a directory with subfolders labelled by class.

3.3 Data Splitting

The splitting was done before the augmentation to ensure no leakage between the sets which could give overly optimistic and hyper inflated results, [21]. Therefore, a stratified split (used to preserve the initial ratio of classes) of the directory of ROIs of 70% train, 15% validation and 15% test was implemented with a seed of 42 for reproducibility. This ensures that the model hyperparameters are adjusted using a validation set that is representative of the real world distribution of BC.

3.4 Train Set Augmentation and Balancing

To train a CNN a sufficient number of images is needed, data augmentation through manipulations of the original images are the most common way to expand the train set especially when it is originally small [22]. In this work, the focus was on transformations that do not distort the images at all in order to preserve the clinical context. Chain augmentation was executed on the images which were then saved with the original. First right angle rotations of 90, 180 and 270 degrees were implemented. Second, the original and rotated images were vertically flipped. Last, the original, rotated and vertically flipped images were subjected to brightness and contrast changes within the ranges (-15,15) and (0.5,1.5) respectively. Figure 5, shows an example of these augmentations on a benign case.

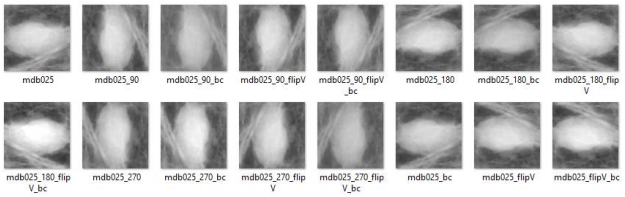

Fig. 5: Augmentation result
Source: created by the authors

This augmentation strategy expanded the dataset by a factor of 16 from 219 training images to 3504 of which 672 are benign, 528 are malignant and 2304 are normal.

It is generally not recommended to train a CNN on an unbalanced training set since this teaches the model to recognize the majority class more than the minority classes. This causes the model to incorrectly predict minority classes when tested on the holdout test set even if it had high training accuracy, [22]. Additionally, since the same augmentations were applied to all the classes in the training set, the imbalance between classes was amplified even further. This is why this training set needed to be balanced to obtain a number of samples in each class that is equal to the malignant minority class. The latter was achieved by randomly selecting images from the normal and benign classes and moving them to a new directory for the balanced and under sampled data until the needed count is reached. This brought the number of samples in each class down to 528 with a total of 1584 images in the new training set.

3.5 Model Training

For TL a model must be loaded in a specific way and then adjusted to fit the task at hand. Additionally, finding the best hyperparameters for the model is a slightly iterative process based on trial and error. VGG16 consists of five feature extracting convolutional blocks which are followed by two dense layers with 4096 units each and a final dense layer with 1000 units to classify the input into one of the 1000 classes of ImageNet, [23]. In the case of TL VGG16 was loaded from Keras applications without its top layer which is the original classification head used when pre trained on

ImageNet. Furthermore, the weights were set to ImageNet weights and the input shape specified as three channel with a size of 224 by 224. Additionally, the literature recommends rescaling to a range of [0,1] by dividing the pixels in the image by 255, [24]. Afterwards, the model is set to non-trainable to ensure that the pre trained weights will not be updated while training the custom classification head later on. Finally, Keras's sequential was used to stack this base model with new custom layers that consisted of global average pooling which reduces the size of the feature map, followed by three dense layers with 'relu' activation function and 512, 256 and 128 units respectively. A dropout layer was added afterwards with a value of 0.5 to minimize overfitting. The final classification layer consisted of three units representing the number of classes and a 'softmax' activation function. The model structure is summarized in Figure 6.

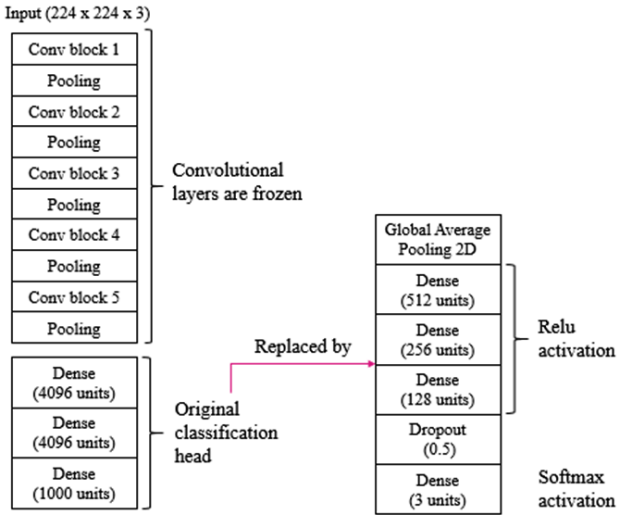

Fig. 6: VGG16 model structure
Source: created by the authors

The model was compiled using categorical cross entropy loss since this is a multiclass problem along with Adam optimizer, [25]. The evaluation metrics were set to accuracy, precision, recall, F1 score and balanced accuracy. The model was trained over a 100 epochs using a batch size of 16 and weight penalty {'B': 2 'M': 2 'N': 1}. An early stopping callback with a patience of 5 was used to monitor the validation loss in order to stop training and restore the best weights when the model starts to over fit. Additionally, a learning rate scheduler was used to reduce the starting learning rate of 1e-3 to a minimum of 1e-7 by a factor of 0.5 with a patience of 5 epochs without improvement of the validation loss.

3.6 Model Evaluation

A set of performance metrics are commonly used to assess the model's ability to complete the task using data such as: true positive (TP), false positive (FP), true negative (TN) and false negative (FN).

3.6.1 Precision

Precision is the amount of positive predictions which are truly positive, [26]. It is calculated using,

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

3.6.2 Recall

Recall is the number of positive predictions that are correct out of the entire correct positive data in the set, [26]. It is calculated below,

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

3.6.3 F1 score

The F1 score is obtained from the weighted average of the recall and precision, [27]. Its formula is:

$$F1\ score = \frac{2 \times (Recall \times Precision)}{Recall + Precision} \quad (10)$$

3.6.4 Accuracy

Accuracy indicates the amount of objects that are classified correctly out of the total, [27]. It is obtained using:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

4 Results and Discussion

The trained model was tested using 50 images out of which 10 are benign, 8 are malignant and 32 are normal. An overall accuracy of 86% was obtained with an actual balanced accuracy of 73.33%. A summary of the performance metrics per class is shown in Table 1.

Table 1. Summary of the results

Class	Precision	Recall	F1 score
Benign	0.78	0.70	0.74
Malignant	0.57	0.50	0.53
Normal	0.94	1.00	0.97

Source: created by the authors

The recall rates indicate that the model is not capable of correctly recognizing benign and malignant cases as opposed to normal cases. Although the obtained balanced accuracy was not very low, it was still insufficient for this type of application. Furthermore, it could be assumed that expanding the number of samples in the test set

could lead to a decrease in performance since the model might make even more mistakes. Several limitations imposed on this work affected the results. For starters, the dataset used in this study was too small for the model to effectively learn from, even with the use of augmentation techniques it was not possible to modify the splitting ratio to better represent the minority classes in the validation and test sets as this would mean sacrificing some of the already limited training data. Additionally, careful considerations had to be taken when performing the augmentation on the training set to ensure no distortion occurs. Specifically, since the images were already cropped to the ROIs, any augmentation technique that distorts the image could lead to a change in the medical context of that image. Moreover, the preprocessing steps used were extracted from the literature without much change to the combinations and order in which the techniques are applied. Furthermore, working on this project as an engineer without guidance from medical professionals affects the quality of this work as the clinical context is considered using knowledge gathered from medical sources on the internet which cannot be fully understood without a proper medical background.

5 Future Work

Several factors can be worked on to improve the findings of this study, these suggestions are split into the following categories.

5.1 Preprocessing Techniques

Different combinations of preprocessing techniques can be used to explore their effect on the overall performance of the model. Additionally, a certain typical order is followed in the literature when using these methods. Breaking the order of these operations could lead to better results. Furthermore, proper data handling needs to be used to ensure that the final result of the model is truly accurate rather than over inflated.

5.2 Dataset

Combining several datasets to obtain one big diverse and uniform set can help the model train more effectively, certain researchers have already worked on this but require peer reviewing and ground truth validation to ensure that they are medically accurate, [28]. Additionally, GAN can be used to augment the dataset but would require validation from radiologists to ensure that the generated images truly represent the classes. This would also require finding the correct balance between synthetic and real data since GANs can inject features into their images that would not be typically found in the scans.

5.3 Collaboration

A collaboration between AI professionals and health care professionals would aid in the design of a system that is truly fit for clinical use. The health care professionals would serve as feedback for the effectiveness of this system all while learning how to properly manage and use this system on their own. Once launched for use, constant maintenance would be performed based on general maintenance protocols as well as suggestions based on the weak points pinpointed through the radiologists.

References:

[1] American Cancer Society, “What is cancer? Cancer basics”, 2022, [Online]. https://www.cancer.org/cancer/understanding-cancer/what-is-cancer.html (Accessed Date: December 3, 2025).

[2] “Global cancer burden growing, amidst mounting need for services”, *Saudi medical journal*, vol. 45, no 3, pp. 326–327, March 2024, [Online]. https://pmc.ncbi.nlm.nih.gov/articles/PMC11115397/ (Accessed Date: December 31, 2025).

[3] Mayo Clinic, “Breast cancer”, 2025, [Online]. https://www.mayoclinic.org/diseasesconditions/breast-cancer/symptoms-causes/syc-20352470(Accessed Date: January 3, 2025).

[4] N. M. Ud Din, R. A. Dar, M. Rasool, A. Assad, “Breast cancer detection using deep learning: Datasets, methods, and challenges ahead”, *Computers in Biology and Medicine*, vol. 149, October 2022, https://doi.org/10.1016/j.compbiomed.2022.106073.

[5] PDQ Screening and Prevention Editorial Board, “Breast cancer screening (PDQ): Patient version “, *National Cancer Institute*, February 2025, [Online]. https://www.ncbi.nlm.nih.gov/books/NBK65715/ (Accessed Date: December 31, 2025).

[6] S. J. Magny, R. Shikhman, A. L. Keppke, “Breast Imaging Reporting and Data System”, *Treasure Island (FL): StatPearls publishing*, August 2023, [Online]. https://www.ncbi.nlm.nih.gov/books/NBK459169/ (Accessed Date: December 31, 2025).

[7] L. Berlin, “The missed breast cancer: perceptions and realities”, *AJR. American*

journal of roentgenology, vol. 173, no 5, pp. 1161–1167, November 1999, https://doi.org/10.2214/ajr.173.5.10541081.

[8] O. Onder, Y. Yarasir, A. Azizova, G. Durhan, M. R. Onur, O. M. Ariyurek, "Errors, discrepancies and underlying bias in radiology with case examples: a pictorial review", *Insights into imaging*, vol. 12, no 51, April 2021, https://doi.org/10.1186/s13244-021-00986-8.

[9] E. H. Hossein, M. M. Emam, A. A. Ali, N. S. Ponnuthurai, "Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review", *Expert Systems with Applications*, vol. 167, April 2021, https://doi.org/10.1016/j.eswa.2020.114161.

[10] M. Chen, W. Fan, W. Tang, T. Liu, D. Li, O. Dib, "Review of machine learning algorithms for breast cancer diagnosis*" In International Conference on Data Mining and Big Data,* pp. 229-243, December 2023, https://doi.org/10.1007/978-981-97-0844-4_17.

[11] M. N. Haslinah, A. B. Noor Mohd, S. Zainuddin, S. M. Mohd, M. I. Ida Syafiza, M. N. A. H. Sha'abani, "The comparative study of deep learning neural network approaches for breast cancer diagnosis" *International Journal ofmOnline and Biomedical Engineering*, vol. 19, no 6, pp. 127-140, May 2023, https://doi.org/10.3991/ijoe.v19i06.34905.

[12] Y. Khourdifi, A. E. Alami, M. Zaydi, Y. Maleh, O. Er-Remyly, "Early breast cancer detection based on deep learning: An ensemble approach applied to mammograms", *BioMedInformatics*, vol. 4, no 4, pp. 2338-2373, December 2024, https://doi.org/10.3390/biomedinformatics4040127.

[13] E. AkbarnezhadSany, H. EntezariZarch, M. AlipoorKermani, B. Shahin, M. Cheki, A. Karami, S. Zahedi, Z. AhmadPour, S. A. Ahmadi-Mazhin, A. Rahimnezhad, S. Sayfollahi, S. Bijari, M. Shojaee, M. R. Seyed, "YOLOv8 outperforms traditional CNN models in mammography classification: Insights from a multiinstitutional dataset", *International Journal of Imaging Systems and Technology*, vol. 35, no 1, December 2024, https://doi.org/10.1002/ima.70008.

[14] D. Kumari, S. K. Vyshnavi, R. Dhar, B. S. A. S. Rajita, S. Panda, J. Christopher, "Smart GAN: A smart generative adversarial network for limited imbalanced dataset", *The Journal of Supercomputing*, vol. 80, pp. 20640–20681, June 2024, https://doi.org/10.1007/s11227-024-06198-3.

[15] J. M. Net, F. Collado-Mesa, "Bias in artificial intelligence: Impact on breast imaging", *Journal of Breast Imaging*, vol. 7, no 4, pp. 408–419, May 2025, https://doi.org/10.1093/jbi/wbaf027.

[16] World Health Organization, "Regulatory considerations on artificial intelligence for health", October 2023, [Online]. https://www.who.int/publications/i/item/9789240078871 (Accessed Date: December 31, 2025).

[17] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, J. Savage, "Mammographic Image Analysis Society (MIAS) database v1.21", *Apollo - University of Cambridge Repository*, August 2015, https://doi.org/10.17863/CAM.105113.

[18] A. Balve, P. Hendrix, "Interpretable breast cancer classification using CNNs on mammographic images", *Proceedings of Machine Learning Research*, vol. 248, August 2024, https://doi.org/10.48550/arXiv.2408.13154.

[19] E. M. F. El Houby, N. I. R. Yassin, "Malignant and nonmalignant classification of breast lesions in mammograms using convolutional neural networks", *Biomedical Signal Processing and Control*, vol. 70, September 2021, https://doi.org/10.1016/j.bspc.2021.102954.

[20] L. M. M. Saravia, A. R. M. Saravia, A. J. G. Vargas, M. M. Silva, Y. M. C. Canaviri, M. A. M. Conde, …, M. J. M. Sanchez, "Optimized mammography preprocessing for tumor detection with YOLO11-seg in young women from the bolivian altiplano", *EEE 8th Portuguese Meeting on Bioengineering (ENBENG)*, pp. 109–112, September 2025, https://doi.org/10.1109/ENBENG67130.2025.11199643.

[21] P. Ghafariasl, M. Zeinalnezhad, S. Chang, "Fine-tuning pre-trained networks with emphasis on image segmentation: A multi-network approach for enhanced breast cancer detection", *Engineering Applications of Artificial Intelligence*, vol. 139, January 2025, https://doi.org/10.1016/j.engappai.2024.109666.

[22] L. Abokaff, M. Saii, "Hybrid model for breast cancer diagnosis on mammograms using

transfer learning", *ResearchSquare*, December 2022, https://doi.org/10.21203/rs.3.rs-2356955/v1.

[23] K. Simonyan, A. Zisserman," Very deep convolutional networks for largescale image recognition" *ArXiv preprint arXiv,* April 2015, https://doi.org/10.48550/arXiv.1409.1556.

[24] S. Akhter, A. H. Sarower, F. Kamal, T. Imam, M. S. Ahammad, J. Liu, "A comprehensive review of datasets and preprocessing techniques in deep learning based breast cancer classification", *2025 IEEE world AI IOT congress (AIIoT)*, pp. 0994-1002, 2025, https://doi.org/10.1109/AIIoT65859.2025.11105273.

[25] H. AlSalman, T. Alfakih, M. A. Al-Rakhami, M. H. Mohammad, A. Alabrah, "A genetic algorithm-based optimized transfer learning approach for breast cancer diagnosis" *CMES - Computer Modeling in Engineering and Sciences*, vol. 141, no 3, pp. 2575–2608, October 2024, https://doi.org/10.32604/cmes.2024.055011.

[26] V. Nemade, V. Fegade, "Machine learning techniques for breast cancer prediction", *Procedia Computer Science*, vol. 218, pp. 1314–1320, 2023, https://doi.org/10.1016/j.procs.2023.01.110.

[27] M. H. Md, M. S. F. M. Javed, R. B. Md, A. H. Rabea, T. Khan, S. Molla, "Breast cancer detection: An effective comparison of different machine learning algorithms on the wisconsin dataset", *Bulletin of Electrical Engineering and Informatics*, vol. 12, no 4, pp. 2446–2456, August 2023, https://doi.org/10.11591/eei.v12i4.4448.

[28] Bhole, G., Suba, S., Parekh, N., "Mammo-bench: A large-scale benchmark dataset of mammography images", *Radiology and Imaging*, October 2025, https://doi.org/10.1101/2025.01.31.25321510.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.