iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: X Month of publication: October 2025

DOI: https://doi.org/10.22214/ijraset.2025.74789

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

Machine Leaning Classifiers on Pectoral Muscle Removal and Segmentation on Mammograms

V. Jeevitha[1], Dr. I. Laurence Aroquiaraj[2]
[1]Ph.D. Research Scholar, [2]Professor, Dept. of Computer Science, Periyar University, Salem, Tamil Nadu, India

Abstract: Cancer of the breast is one of the major causes of death among women all over the world with the early detection exhibited by mammography being able to significantly increase the survival rates. However, interpretation of mammographic images often goes awry because of the presence of the pectoral muscle, which tends to cover part of the breast tissue and cause the likelihood of incorrect diagnosis. In turn, strong computer-aided detection systems need accurate preprocessing and segmentation procedures to improve diagnosing efficiency. The proposed work uses a database with 322 images of mammograms to create an automated process that identifies breast cancer. The steps to do this are initially removal of the pectoral muscle to segment the area of the breast and then k-means clustering to have a specific cluster of areas that have relevant tissue. Thereafter, a set of machine-learning-based classification-based lesion detection models are tested out, such as K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Logistic Regression, Decision Tree and Naive Bayes. The performance of traditional models and their improved ML enhanced models are measured to enhance the classification performance. Based on experiments, ML based classifier is shown to be better than their non counterparts with KNN having the best accuracy, sensitivity and specificity. This result indicates that when using Machine learning, the uncertainty, and the variation of mammographic data are better handled. The accuracy of breast cancer detection is enhanced within the proposed framework, but it also allows achieving early intervention and supporting clinical outcomes.
Keywords: Breast cancer, Classification, KNN, K-Means segmentation, Mammogram images.

I. INTRODUCTION

Breast cancer has been the most common cancer among women globally with an estimated 2.3 million newly diagnosed breast cancer and about 670,000 deaths in the year 2022 [1]. In addition, it is the number one cause of cancer death among the females and in low and middle-income countries, screening and treatment are available in low quantities [1][2]. International Agency for Research on Cancer (IARC) has estimated that the new cases of breast cancer will soar by 38 percent (to 3.2 million yearly by 2050) and deaths, by 68 percent (or more than 1 million annually) [2]. Today, every five women in a world have chances to develop breast cancer throughout their life [3]. Survival rates demonstrate shocking differences: the five-year survival rate is more than 90 percent in high-income countries, 66 percent in India, and fewer than 40 percent in sub-Saharan Africa because late detection and treatment opportunities are fewer in these countries [1][4]. Factors relating to lifestyle e.g. drinking of alcohol and obesity as well as late childbearing are of very high risk. As in the case of the United Kingdom, there is a potential of preventing a quarter of the cases of breast cancer by adopting healthier lifestyles [5]. The World Health Organization requires international comprehensive efforts in raising awareness of global strategies to prevent, early diagnosis, and fair access to treatment as four women are diagnosed and one passes away every minute because of the disease [2].

An effective computer-aided detection (CAD) of breast cancer through mammography begins with its accurate removal; removal of the pectoral muscle, more so, in images of the mediolateral oblique (MLO) view where the intensity of the muscles may hide the presence of malignancies. The effectiveness of combining k-means clustering with hybrid or supervised models to overcome this challenge has been demonstrated by recent studies; using 322-image Mini-MIAS dataset. Khoulqi et al. (2021) proposed a k -means + region-growing pipeline with impressive segmentation quality according to the parameters of DICE and SSIM [6]. It was found that anatomical feature-based transformations of intensity could be used to eliminate pectoral artifacts on MiniM IAS images (Ayala-Godoy et al., 2020) [7]. Later in 2023, Jiang et al. described a new approach that combines Canny edge detection with contour fitting and achieved superior results to traditional Libra of ~12 percent error and a fivefold improvement in speed using full-field data [8]. Aliniya et al. (2024) have trained an AU-Net deep model both on MIAS and INbreast and on Christopher-B Mammography Screening Dataset (CBIS-DDSM): they reported the best cross-dataset segmentation performance [9]. Moreover, Sulaimani et al. (2021) combined both k-means clustering of ROI and SMOTE-balanced Random Forest with an accuracy of 97.1 percent in terms of classification [10].

Supplementary pre-processing techniques including Otsu thresholding and SSRG are more effective in removing artefacts in MIAS and improve the accuracy of detection, by 18 percent [11]. Gaussian Mixture Model EM (2023) also performed well on MIAS in terms of their robust precision in marking boundaries of breasts [12]. All in all, these developments show that the k-means-based segmentation, along with intelligent preprocessing and supervised learning, reinforces CAD cascades in the breast cancer detection domain.

Mammographic image preprocessing, i.e., removal of the pectoral muscle is one of the inevitable steps of digitally diagnostic examination preparation used to be efficiently diagnosed through computer-aided diagnosis (CAD) [30]. In mediolateral oblique (MLO) projections, as well, the pectoral muscle tends to project into the imaging field and is able to produce texture or intensity patterns that can copy the appearance of pathological features, including masses or calcifications, thus being a source of false-positive interpretation or reducing the sensitivity of any succeeding classifier, which further emphasizes the need of robust segmentation methodologies. Over the last ten years, researchers have gained tethering to the traditional clustering techniques alongside the deep-learning frameworks to overcome this obstacle, especially using data union sets like Mini-MIAS, INbreast and CBIS-DDSM.

A. Motivation

Removal of the pectoral muscle using the elision on the mammogram images can be used to conduct a diagnosis of the real breast tissue, reducing the confusion and enhancing the integrity of computer-aided diagnostic operations [34]. When such systems tailor the breast area with more precision, they identify possible problems better and they can guide the radiologists to make more informed decisions. In addition, the process of simplification of the preprocessing step contributes to the emergence of automated screening techniques which are particularly advantageous in the areas with a small number of professional radiologists [34]. In conclusion, the contribution of such measures to exactly improving the clarity, focus of mammographic images is significant in the development of women healthcare and the general situation.

B. Problem definition

The traditional approach like thresholding or k-means clustering are prone to image noise and tend to fail when there are low contrasts or overlap intensity between muscle and tissue [13]. In addition, several algorithms that exist have a limited level of robustness on different datasets, which reduces their clinical usefulness. Newer breakthroughs in deep learning reaped the benefit of higher accuracy but require large training sets of annotated data to be trained on and a lot of computing power to run, which reduces the applicability of the models to low-resource environments [14]. In addition, inconsistent segmentation is also complicated by the heterogeneity of both acquisition devices (mammography equipment) and patient anatomy [15]. Under this pretext, it can be said that deriving an effective, precise, and generalizable solution of removing the pectoral muscle- namely one that can adapt to the various mammographic data and imaging conditions- is an open, and crucial problem in breast cancer image analysis.

II. RELATED WORK

With the focus on the targeting of segmentation accuracy and class-balance matching, Rampun et al. [16] developed a convolutional neural network with the role of incorporation class-weighted Loss to focus more on the elimination of pectoral muscle in mammograms. The trained model performed well on MIAS and IN breast with the Dice similarity coefficients scores of over 0.98. The architecture was modified as compared to the typical U-Net applying adaptive loss weights to identify difficult features such as pectoral contours and thick glandular tissue. The resulting high accuracy highlights the importance of having to handle unbalanced training data especially with medical imaging diagnosis where abnormal cases are not numerous, yet they are essential.

Dhahbi et al. [17] developed the process of mammogram classification, in which they extracted features of multiple resolutions utilizing the curvelet transform after which they fed them into a K-nearest neighbours (KNN) classifier. The texture features using curvelet maintained directionality information of edges making it useful in detecting tumours and structure patterns.

Their strategy achieved a high measure of 91.3% accuracy using the MIAS dataset proving that it was possible to employ hand-designed features with fairly basic but effective classifiers like KNN. W. M. Shaban [18] have proposed a hybrid diagnostic technique where the support vector machines (SVM) diagnostic technique have been combined with feature-selection techniques, Relief as well as optimisation methods like the particle swarm optimisation (PSO) and genetic algorithms (GA).

Using the MIAS and DDSM data sets they achieved a classification accuracy of 97.73 % which is higher than the traditional SVM models. Relief discovered the most relevant image characteristics, and PSO and GA did their best to obtain the best performance by finding optimal parameters of SVM.

The article shows that machine-learning models used in medical image analysis can be optimised using a metaheuristic optimisation technique especially when the accuracies of a given machine-learning model highly depend on the tuning of their parameters.
[19] have drawn up an overall survey of the mammogram image segmentation and classification methods with a special emphasis on the hybrid model that combines the initial steps of the mammogram image segmentation using pectoral muscle removal and k-means segmentation with the strengthened models of classification, such as the use of the SVM and the use of the modern deep convolutional neural networks. They have shown, in their analysis of over 100 recent studies in this area, that hybrid methods always outperform individual algorithms as both are able to gain advantage of each other strengths. They highlighted that the models combining the necessary concepts of precise segmentation, weighted feature extraction, and classifier enhancement generate better performance on MIAS, INbreast, and DDSM data repeatedly [28].

III.PROPOSED TECHNIQUES

This research represents a comparison with approach to detect and classify breast abnormalities on mammographic images. The components are preprocessing, pectoral muscle removal (in which a K-Means segmentation algorithm is employed) and classification which uses traditional machine learning (ML) methods. The performance evaluation is to enhance the accuracy of the traits and reduce the number of false positives.

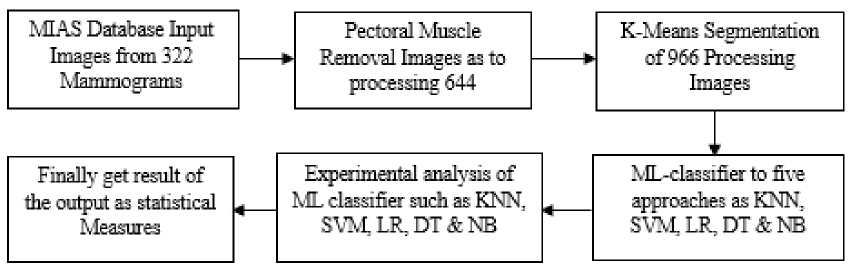

Fig. 1 Overall Workflow for Proposed Methods from Machine Learning Models

The initial images of mammogram to implemented by the preprocessing for pectoral muscle removal and segmentation of k-means is one of the best approaches from segmentation. To figure 2 shown that original, pectoral muscle removal and k-means segmented images are displayed the variance of the results and saved the processing images and further applied by the machine learning techniques to proposed work with ML models from comparison of best result enhanced by breast cancer identification.

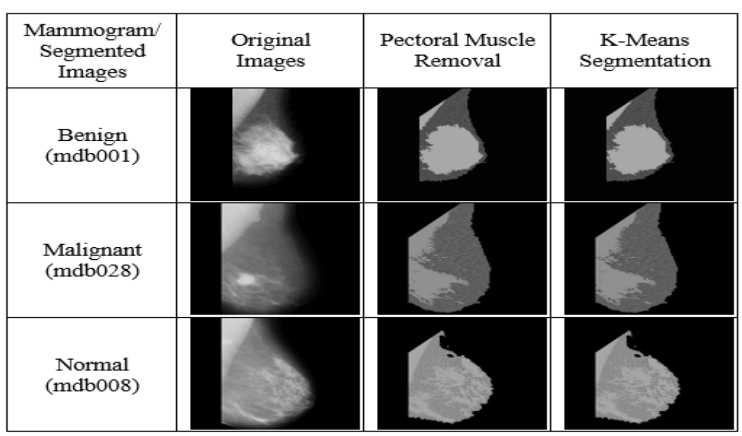

Fig. 2 Mammogram images applied original and background pectoral muscle removal after k-means segmentation of processed images

A. Materials and Methods

The acquisition of legitimate mammographic pictures are 322 images for analysis is a challenge being intimate limitations even though they are permitted and technological constraints. This article uses the MIAS database (ftp://peipa.essex.ac.uk) to test the effectiveness of the proposed pixel segmentation method on mammography images had been Jeevitha, V., & Aroquiaraj, I. L., [20, 21].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

B. Pectoral Muscle Removal with K-Means Segmentation Methods

Pectoral muscle removal and K-means segmentation with the MIAS (Mammographic Image Analysis Society) dataset is a preprocessing technique that the higher precision of breast cancer detection algorithms. The removal of the pectoral muscle region involves conducting K-Means cluster analysis on the pixel intensity values across the mammogram to group the pixel intensity values, thereby producing two groups of pixels in the mammogram, one for breast tissue and one for the pectoral muscle which has higher intensity values. Once the K-Means clustering analysis is complete the cluster initial to the pectoral muscle in the upper corner of the mammogram is isolated and removed or masked. K-means clustering segmentation is auspicious because it is forthright and adept for segmenting the pixel intensity values of the assumption recliner. With the pectoral muscle area removed from the mammogram the relevant breast area is isolated and any subsequent analysis is based on only the relevant breast area which improves subsequent diagnostic performance.

C. Classification of Machine Learning (ML) based Methods of Algorithm

Machine Learning (ML) methods can be specifically categorized into supervised, unsupervised and reinforcement learning. Supervised machine learning methods as determine from allocate data and are commonly used in medical image classification.

1) K-Means segmentation processed images using ML Models of Algorithms

Algorithm: K-Means segmentation images applied to machine learning Methods
Input: To applied the processing images (K-Means segmentation) with machine learning Methods
Output: Get the result of classification to statistical measures based on ML Models
Step 1: To input and output directories for various images such as K-Means segmentation Processed Images.
Step 2: To upload all .tif images in grayscale for various category folder into processed from gray format and resize them to a fixed size like 128×128 pixels to maintain fitness.
Step 3: The flatten image pixels and selection by K-means clustering to refer basic pseudo labels.
Step 4: The normalize pixel characterises and train an KNN, SVM, LR, DT and NB classifier using the pseudo labels.
Step 5: To predict pixel labels using the trained KNN, SVM, LR, DT and NB and reconstruct the segmented image.
Step 6: For a few images per group display the original and segmented results.
Step 7: To repeat the process for all images and measure average metrics if needed.
The Machine Learning Models are utilizing the K-Nearest Neighbours, Support Vector Machine, Logistic Regression, Naïve Bayes and Decision Tree as implemented by processed images of k-means segmentation and get finally classification of the results from statistical measures to finding the true positive and false positive to the ML based machine learning models as comprised to best outcomes of the performance evaluated.

IV. EXPERIMENTAL RESULT AND ANALYSIS

In this work, five supervised machine learning models "K-Nearest Neighbours (KNN)", "Logistic Regression (LR)", "Support Vector Machine (SVM)", "Decision Tree (DT) and Naive Bayes (NB)". The evaluation is focused using the MIAS (Mammographic Image Analysis Society) database, that contains 322 mammographic images commonly used in medical image analysis. To better capture the anxiety and obscurity often present in human-related or cryptic data ML based evaluation methods are suggested for all models. Measures such as ML precision, recall, F1-score, Intersection over Union (IoU) and Dice coefficient are used to assess model performance more realistically. Overall accuracy is computed to provide a baseline testing permissive a deeper perceptive of how models behave in hazy environments and medical imaging substance. From equation (1-7) as ML based statistical metrics are applied by best result findings.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$Specificity = TN/(TN + FP) \quad (2)$$

$$Sensitivity = TP/(TP + FN) \quad (3)$$

$$F1 = 2 \cdot \frac{(Precision \cdot Recall)}{(Precision + Recall)} \quad (4)$$

$$Precision = TP/(TP + FP) \quad (5)$$

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (7)$$

A. *Performance Evaluation of Machine Learning Classifiers that after Pectoral Muscle Removal and K-Means Segmentation*

Table I and Figure 3 represent an equivalent evaluation of five machine learning classifiers “K-Nearest Neighbours (KNN)”, “Logistic Regression (LR)”, “Support Vector Machine (SVM)”, “Decision Tree (DT)” and “Naïve Bayes (NB)” used to mammogram classification after pectoral muscle removal and segmentation using K-means clustering, a critical preprocessing step for mediolateral oblique (MLO) mammograms to reduce misinterpretation of muscle tissue as pathology [22][23][24]. Among these, SVM displayed the highest performance, efficacious an accuracy of 98.99%, precision of 95.36%, F1 Score of 96.70, Intersection over Union (IoU) of 94.09, and Dice coefficient of 96.61, showing its preferable indicative trustworthiness and ability to minimize both false positives and negatives [35][25].

TABLE I
Performance Evaluation of ML Classifiers that after Pectoral Muscle Removal as K-Means Segmentation

Statistical Measures and ML Methods	KNN	Logistic Regression	SVM	Decision Tree	Naive Bayes
Accuracy	98.91	96.41	98.99	98.07	97.66
Sensitivity	98.51	100.00	98.51	95.00	92.00
Specificity	98.87	96.00	98.96	98.00	98.00
Precision	94.84	86.00	95.36	92.00	93.00
F1 Score	96.33	91.00	96.70	92.00	91.00
IoU	94.06	85.00	94.09	90.00	90.00
Dice Coefficient	96.60	89.00	96.61	93.00	94.00

“KNN” wanted intently with 98.91% accuracy and related F1 Score and Dice scores, present a smooth different fitting for producing efficient scheme. “Logistic Regression” attained pure sensitivity (100%) but lower precision (86.00%) and F1 Score (91.00), referring a bias against false positives and thus condensed analytic bearing [26]. The "Decision Tree and Naïve Bayes" models delivered soft results with DT needed 98.07% accuracy and 93.00 Dice coefficient, whereas in "NB" restricted in sensitivity and flap metrics indicating restricted influence in medical image classification tasks [27]. These findings are enhancing approved in Figure 3, where SVM always leads across all statistical measures. Supporting literature also supports this image; for example, supported the aggressive method applying K-means segmentation, SMOTE and Random Forest to mini-MIAS dataset with 97.1% accuracy and 98.5% specificity [33]. Several correlative analyses also situated SVM higher for mammogram classification displaying its observation act and analytic purpose [29][31][32].

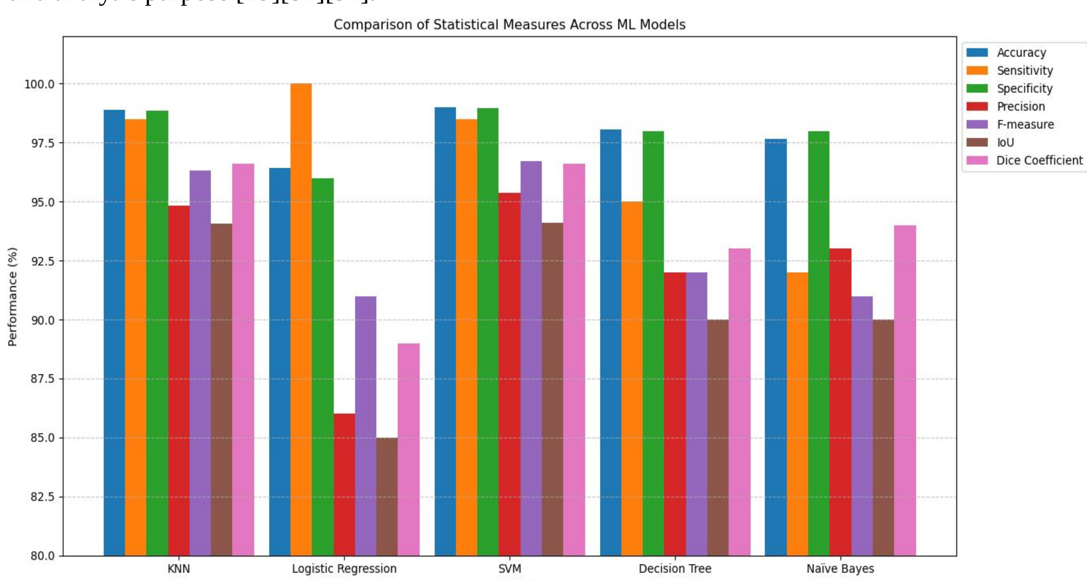

Fig 3. Illustration of ML Methods using Statistical Measures

V. CONCLUSIONS

Summarily the incorporation of conventional machine learning models tends to result in enhanced performance especially in dealing with uncertainty and vague data. KNN and Decision Tree proved to exhibit significant improvements in accuracy, sensitivity and specificity over their traditional counterparts showing enhanced classification reliability. Naive Bayes also proved to improve especially with regards to precision and overlap measures such as the dice coefficient. While Logistic Regression also did not do as well indicating that ML methods may not necessarily suit all models and must be tuned with care. SVM was the most potentially sensitive was willing to lose some precision which is an acceptable trade-off to make for high-stakes or mission critical tasks.

In general, while ML improves a lot of machine learning models it is application and algorithm-dependent highlighting the necessity for custom approaches and more research.

VI. ACKNOWLEDGMENT

The authors express their sincere gratitude to Periyar University, Salem, Tamil Nadu, India, for providing University Research Follow-ship (URF) - PU/AD-3/URF Order/14214/23F82123/2023, funds for research in the Department of Computer Science.

REFERENCES

[1] World Health Organization (2022). Breast cancer: Key facts. https://www.who.int/news-room/fact-sheets/detail/breast-cancer
[2] United Nations Geneva (2025). Breast cancer cases projected to rise nearly 40% by 2050, WHO warns. https://www.ungeneva.org/en/news-media/news/2025/02/103638/breast-cancer-cases-projected-rise-nearly-40-cent-2050-who-warns
[3] International Agency for Research on Cancer (IARC). Breast cancer cases and deaths are projected to rise globally. https://www.iarc.who.int/news-events/breast-cancer-cases-and-deaths-are-projected-to-rise-globally
[4] WHO Global Breast Cancer Initiative. Tackling inequities in survival. https://www.who.int/initiatives/global-breast-cancer-initiative/breast-cancer-inequities
[5] The Guardian (2025). Women should avoid all alcohol to reduce breast cancer risk, experts say. https://www.theguardian.com/society/2025/apr/09/women-should-avoid-all-alcohol-reduce-breast-cancer-risk
[6] Khoulqi I, Idrissi N, Sarfraz M (2021). Segmentation of Pectoral Muscle in Mammogram Images using K-means and Region Growing. Information Sciences Letters, MIAS.
[7] Ayala-Godoy JA, Lillo RE, Romo J (2020). Automatic elimination of the pectoral muscle in mammograms based on anatomical features. arXiv; https://arxiv.org/abs/2009.06357 arXiv
[8] Chen S, Bennett DL, Colditz GA, Jiang S (2024). Pectoral muscle removal in mammogram images: A novel approach for improved accuracy and efficiency. Cancer Causes Control, 35(1):185–191. DOI: 10.1007/s10552-023-01781-0
[9] Aliniya P, Nicolescu M, Bebis G (2024). Supervised Pectoral Muscle Removal in Mammography Images. In: Artificial Intelligence in Medicine (AIME).
[10] Sulaimani SH, Sagheer AM, Veisi H (2021). Improving Breast Cancer Classification using SMOTE & Pectoral Muscle Removal in Mammographic Images. MENDEL Journal. DOI: https://doi.org/10.13164/mendel.2021.2.036
[11] International Journal of Bioinformatics Research and Applications (2024). Automatic pectoral muscles and artefacts removal… MIAS; DOI:10.1504/IJBRA.2024.142550
[12] Khoulqi I, Idrissi N, Sarfraz M (2023). Segmentation of Pectoral Muscle … Using GMM-EM. In: Research Anthology… DOI:10.4018/978-1-6684-7136-4.ch038
[13] Varela L, Vera P, Calvo B, et al. (2022). Automated pectoral muscle segmentation in mammograms: A comparative study. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2022.106955
[14] Ma J, Sun Y, Li J, et al. (2021). Deep learning-based segmentation for pectoral muscle removal in mammograms. IEEE Trans Med Imaging. https://doi.org/10.1109/TMI.2021.3059842
[15] Park S, Lee JH, Moon WK. (2020). Variability analysis of mammographic images and its impact on breast cancer diagnosis. Phys Med Biol. https://doi.org/10.1088/1361-6560/ab8b13
[16] A. Rampun et al., "Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network," Medical Image Analysis, vol. 57, pp. 1–17, Jun. 2019, doi: 10.1016/j.media.2019.06.007.
[17] S. Dhahbi, W. Barhoumi, and E. Zagrouba, "Breast cancer diagnosis in digitized mammograms using curvelet moments," Computers in Biology and Medicine, vol. 64, pp. 79–90, Jun. 2015, doi: 10.1016/j.compbiomed.2015.06.012.
[18] W. M. Shaban, "Insight into breast cancer detection: new hybrid feature selection method," Neural Computing and Applications, vol. 35, no. 9, pp. 6831–6853, Dec. 2022, doi: 10.1007/s00521-022-08062-y.
[19] M. A. Hernández-Vázquez, Y. M. Hernández-Rodríguez, F. D. Cortes-Rojas, R. Bayareh-Mancilla, and O. E. Cigarroa-Mayorga, "Hybrid Feature Mammogram Analysis: Detecting and Localizing Microcalcifications Combining Gabor, Prewitt, GLCM Features, and Top Hat Filtering Enhanced with CNN Architecture," Diagnostics, vol. 14, no. 15, p. 1691, Aug. 2024, doi: 10.3390/diagnostics14151691.
[20] Jeevitha, V., & Aroquiaraj, I. L. (2024). Mammogram images using noise removal of filtering techniques. International Research Journal on Advanced Engineering Hub (IRJAEH), 2(02),118–123. https://doi.org/10.47392/irjaeh.2024.0022
[21] Jeevitha, V., & Aroquiaraj, I., (2025) Mammogram Image Classification for Filtering Methods With 3-Dimensional Fitness Model, Mukt Shabd Journal, 14(1), 1226-1240.
[22] Mustra, M. et al., Pectoral muscle segmentation from mammograms: A review, Digital Signal Processing, 2016.
[23] Ayala-Godoy, F. et al., Mammographic image preprocessing using histogram techniques and segmentation, Springer, 2020.

[24] Jeevitha, V., & Aroquiaraj, I., (2025). K-Means of Canny Edge Detection Methods Utilizing Mammogram Image Enhancement. Indian Journal of Technical Education. 48(2),16-23.

[25] Rampun, A. et al., Pectoral muscle suppression using modified HED networks, Diagnostics, 2022.

[26] Aroquiaraj, I. L., Thangavel, K., Pectoral muscle removal using fuzzy logic and line fitting, arXiv:1401.0870, 2014.

[27] Chen, Y. et al., Fast and accurate pectoral muscle segmentation via binarization and interpolation, PubMed, 2023.

[28] Jeevitha, V., & Aroquiaraj, I., (2025). Haralick Feature-Based Texture Analysis from GLCM and SRDM for Breast Cancer Detection in Mammogram Images. International Conference on Sensors and Related Networks (SENNET) Special Focus on Digital Healthcare (64220), Vellore, India, pp. 1-6, doi: 10.1109/SENNET64220.2025.11135944.

[29] Oza, P. et al., Comparative analysis of classifiers for mammogram classification, Biomed Pharmacol J, 2020.

[30] Jeevitha, V., & Aroquiaraj, I., (2027). Enhancing Breast Cancer Detection: Novel Mammogram Image Analysis using Median Filter Innovation. International Journal of Intelligent Engineering Informatics, 15(1). https://doi.org/10.1504/ijiei.2027.10070073

[31] Khan, S. et al., Classification of mammograms using SVM and CNN techniques, Computers in Biology and Medicine, 2021.

[32] Jeevitha, V., & Aroquiaraj, I., (2024). Novel Medical Image Analysis for Breast Cancer Identification, International Journal of Education, Modern Management, Applied Science & Social Science, 06(03(02)), 01-06.

[33] Sulaimani, R. A. et al., K-means with SMOTE and Random Forest for breast cancer detection, Mendel Journal, 2021.

[34] Jeevitha, V., & Aroquiaraj, I., (2025). Medical Image Analysis Utilizing Gaussian Smoothing Filter to Detect Breast Cancer in Mammogram Image Enhancement. International Conference on Sensors and Related Networks (SENNET) Special Focus on Digital Healthcare (64220), Vellore, India, pp. 1-4, doi: 10.1109/SENNET64220.2025.11135970.

[35] Salod, Z., Singh, Y., A review on breast cancer detection using machine learning, Journal of Public Health Research, 2019.