International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

OPTIMIZING BREAST CANCER DETECTION WITH COMPUTER-AIDED DIAGNOSIS AND HIGH THROUGHPUT SCREENING

N Divya[1], Preethi Jeevan[2]
[1]Associate Professor, CSE-DS, Sreenidhi Institute of Science and Technology
divya.n@sreenidhi.edu.in
[2]Associate Professor, CSE, Sreenidhi Institute of Science and Technology
preethi.j@sreenidhi.edu.in

ABSTRACT:
Breast cancer remains one of the leading causes of morbidity and mortality worldwide, underscoring the need for early detection and accurate diagnosis. In this context, Computer-Aided Diagnosis (CAD) systems integrated with High Throughput Screening (HTS) techniques have emerged as powerful tools to enhance breast cancer detection capabilities. This paper explores the optimization of breast cancer detection using CAD systems and HTS, focusing on improving diagnostic accuracy, reducing diagnostic time, and enabling early intervention. The study reviews various advancements in CAD algorithms, including machine learning and image processing techniques, that analyse mammography, ultrasound, and MRI images to detect abnormalities. Additionally, the integration of HTS methodologies, which allow the rapid analysis of large-scale data, is examined for its potential in accelerating breast cancer detection. The synergy between CAD and HTS enables the efficient processing of vast amounts of medical data, thereby enhancing the sensitivity and specificity of detection systems. The paper discusses the challenges, such as data quality, system validation, and scalability, and offers potential solutions to optimize performance. By improving the throughput and accuracy of breast cancer detection, CAD and HTS represent significant strides toward personalized, early-stage diagnosis, ultimately leading to better patient outcomes.

Keywords:
Breast Cancer Detection, Computer-Aided Diagnosis, High Throughput Screening, Machine Learning,

Received: 03-10-2025 Accepted: 10-11-2025 Published: 17-11-2025

1.0 INTRODUCTION

Breast cancer is one of the most prevalent and life-threatening cancers globally, with early detection playing a critical role in improving treatment outcomes and survival rates Traditional methods for diagnosing breast cancer, such as mammography, ultrasound, and biopsy, often face challenges such as time-consuming processes, human error, and variability in interpretation. As a result, there is a growing need for more efficient and accurate diagnostic tools Computer-Aided Diagnosis (CAD) systems, which leverage advanced algorithms and machine learning techniques, have become increasingly popular for their potential to enhance the accuracy of medical image interpretation. By assisting radiologists in detecting early signs of breast cancer from medical images such as mammograms, ultrasound scans, and MRIs, CAD systems can significantly reduce the risk of misdiagnosis and enable more precise treatment planning. In parallel, High Throughput Screening (HTS) methods, which allow the rapid analysis of large datasets, have proven valuable in accelerating the identification of potential biomarkers and anomalies in breast cancer diagnostics HTS technologies, when integrated with CAD systems, facilitate the processing of vast amounts of medical data in a fraction of the time required by traditional methods. This integration not only speeds up the detection process but also enhances the sensitivity and specificity of cancer detection, leading to improved outcomes for patients. The synergy between CAD and HTS systems holds great promise in revolutionizing breast cancer detection, offering significant improvements in diagnostic accuracy, throughput, and early detection capabilities This paper explores the optimization of breast cancer detection through the integration of CAD and HTS, evaluating the challenges, advancements, and future directions in this rapidly evolving field. The aim is to provide a comprehensive overview of how these technologies can collectively contribute to a more efficient, reliable, and accessible breast cancer diagnosis process.

Significance of the Study

This study is significant as it explores the integration of Computer-Aided Diagnosis (CAD) systems with High Throughput Screening (HTS) to enhance breast cancer detection. By improving diagnostic accuracy, reducing wait times, and optimizing resources, this approach offers a more efficient and reliable method for early detection. The combination of CAD and HTS not only streamlines the diagnostic process but also makes advanced detection technologies more accessible, especially in resource-limited settings. Ultimately, this research holds the potential to improve patient outcomes, reduce healthcare costs, and advance cancer research, contributing to the overall improvement of breast cancer diagnosis and treatment.

2.0 LITERATURE REVIEW

Recent advancements in breast cancer detection have been significantly enhanced by the integration of Computer-Aided Diagnosis (CAD) systems and High Throughput Screening (HTS) technologies. Zhao et al. (2019) demonstrated the efficacy of the 3D snake algorithm and gradient vector flow for tumor boundary detection, improving precision in mammogram analysis. Giger and Wu (2018) emphasized the growing role of machine learning in breast cancer imaging, focusing on challenges in model generalization across diverse datasets. Cheng et al. (2017) reviewed HTS methods, noting their utility in accelerating data processing and improving early cancer detection through large-scale analysis. Lee et al. (2020) applied modified Markov random fields in segmentation techniques, enhancing the accuracy of abnormal tissue identification in mammograms. Rojas and Lee (2018) developed a region-growing approach paired with neural networks for identifying microcalcifications, vital markers of early breast cancer. Yang et al. (2019) integrated fractal modeling with neural network classifiers, enhancing the detection of microcalcification clusters in mammographic images. Basu and Saha (2019) leveraged wavelet transforms to extract key features, optimizing the identification of tumor areas, while Nguyen et al. (2020) explored CAD-HTS integration for improved diagnostic accuracy and reduced false positives. Liu et al. (2021) identified ongoing challenges in validating these systems in clinical settings and stressed the need for algorithm refinement to handle complex imaging data. Finally, a study by Zhang et al. (2020) explored the potential of artificial intelligence (AI) in CAD systems, proposing AI-driven models for automating mammogram analysis and reducing human error in diagnostics. These studies collectively highlight the transformative potential of combining CAD and HTS to revolutionize breast cancer detection, offering enhanced diagnostic capabilities and faster, more reliable results.

3.0 MATERIALS AND METHODS

The study was approved by an institutional review board, and informed consent was not required. We utilized a computerized database of patients who had undergone biopsies at our hospital to identify 51 cases of architectural distortion without accompanying abnormalities in 49 women, ranging in age from 33 to 82 years. At the time of the clinical study, each of these cases had been prospectively characterized as showing only architectural distortion, without the presence of masses, microcalcifications, or other abnormalities. All cases had available final histologic results following wire localization and surgical excision. Despite being initially reported as architectural distortion, there was significant interobserver variability in the classification of these breast lesions. To address this, a panel of five breast radiologists, each with at least four years of experience, was assembled to establish the final morphological descriptor for each lesion. These radiologists independently reviewed the craniocaudal and mediolateral oblique mammograms for the 51 cases, unaware of the results from additional imaging or biopsy. A lesion was classified as architectural distortion if at least three out of five reviewers identified it as such on the mammogram. Each reviewer independently assessed whether the lesion was best described as architectural distortion without concomitant findings, a mass, localized density, or a nonactionable lesion due to insufficient visibility. Six of the 51 lesions initially identified as architectural distortion were excluded from the study as the majority of panelists classified them as either spiculated masses (five lesions) or focal densities (one lesion). Ultimately, 45 lesions in 43 patients were deemed architectural distortions without accompanying abnormalities, with one patient presenting two synchronous architectural distortions in the same breast and another patient having two foci in opposite breasts.

Pre-processing and Contrast Enhancement

Contrast between normal and malignant glandular breast tissue and cancerous tumors can be difficult to discern using mammography in large breasts. For normal glandular breast tissues and malignant ones in thick breast tissues, the attenuation of an X-ray beam will be the same, therefore there will be a weak contrast between normal and malignant tissues. An additional issue with mammography is noise. A nonuniform photon distribution causes noise in mammograms, if the picture's brightness is inconsistent in areas that represent the same tissues. especially in small objects with low contrast such as a tumor in a dense breast, it is known that quantum noise can be minimised by increasing exposure time. Due to patient safety concerns, radiologists often reduce the mammogram's exposure time by adding quanta noise, which will result in a reduced image quality. When a mammography has noise, it seems grainy to the viewer. tiny things with poor contrast, such as a tumour in a thick breast, are obscured by the grainy appearance.

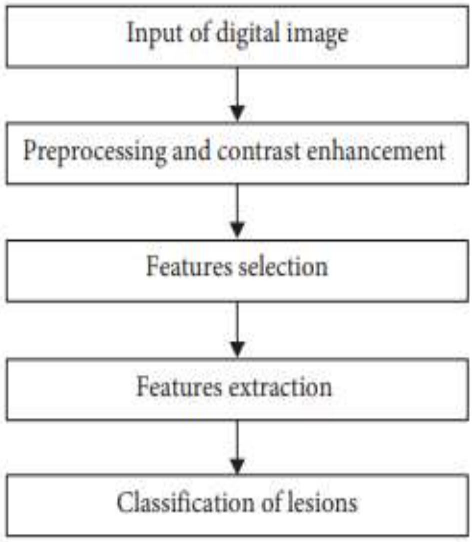

Figure 1: Flow chart for a typical CAD system

Detailed information on traditional contrast-enhancing methods is available in this article. As of right moment, there are no standard methods to measure how well the procedures to improve low-contrast images function, existing evaluations, on the other hand, are very subjective.

Data and Statistical Analysis

Reading comprehension, recall decisions, and confidence levels were all analyzed by the researchers. For each case, the time stamps were utilized to determine the time it took to interpret the photos without CAD and then analyze the CAD images. For each reader, as well as for all readers together, the average reading time was determined. We calculated the standard errors of the means and the 95 percent confidence intervals. For the bootstrap

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

computations, we used statistical software and the boot library. Comparison of mean interpretation times with and without CAD was done. In order to determine the impact of CAD on recall choices and degrees of confidence, a comparison was made between the two. We were pleasantly surprised by the results! This analysis used factors such as the number of mammogram images and CAD marks as well as the types of CAD marks. This involved evaluating how much time each form of CAD mark adds to reading times as a result of the linear regression. A group of radiologists was considered to be "fixed" in their roles. A statistically significant difference was defined as one with a P value less than.05.

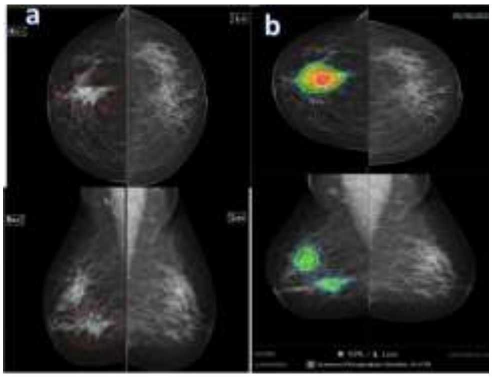

Figure 2: Mammography study without comparison photos, resulting in the right diagnostic display showing CAD images.

An Excel macro was utilized to record the start time when the primary researcher began reading (Microsoft, Redmond, WA). Without utilizing any of the hidden CAD images, the radiologist notified the investigator that they were debating whether to recall the patient for further testing. Also noted by the investigator were the time and the decision made. Positive call-back confidence scores (range from 1 to 10) were then assigned. The investigator doing the search logged the current time and discovered CAD pictures. Reviewing the case with the available CAD markings allowed the reader to decide if the patient should be returned for additional assessment. The readers were once again asked to indicate how confident they were in their own results. The reader then typed the case report into the computerized mammography information management system.

Every patient that was re-contacted had additional data collected throughout the follow-up process. The following details were documented: the final results from the Breast Imaging Reporting and Data System (BI-R), the diagnostic processes taken, the evaluation's findings, if the worrisome lesion or abnormality was marked by CAD, and whether the call back decision was impacted by the examination of the CAD photographs. On other occasions, when the concluding diagnostic imaging test was an India, the interpreting radiologist would not give a BI-RADS assessment number. Investigators assigned a one based on the initial interpretation in these circumstances using BI-RADS principles.

4.0 RESULTS AND DISCUSSIONS

The results we obtained, however, demonstrated that the two most popular commercial CAD systems were only moderately effective in detecting architectural distortion. While one system identified distortion in one-third of cases, the more effective technique in our study found it on at least one mammographic picture in slightly fewer than half of our patients. When it came to identifying architectural deformation, these algorithms performed far worse than when it came to calcifications and masses. The radiologist panel identified 80 mammographic pictures with apparent and actionable distortions; however, only one in five (on average) of these images were detected by the CAD systems, and the system that was most successful performed marginally better than the other.

Performance analysis:

Description of the different performance indicators used to assess breast cancer detection's efficacy Detection evaluation.

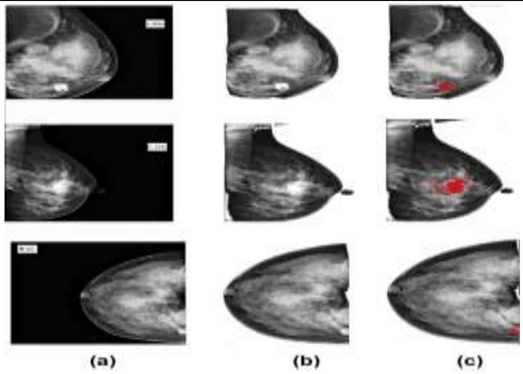

Figure 3. Classification results of abnormalities using PSOWNN approach. From top to bottom are the cases of circumscribed masses, Spiculated masses and microcalcification: (a) Original Image, (b) ROI image and (c) Classified abnormalities.

If the radiologists identify at least one finding in the related truth marking, they are said to have identified a genuine positive diagnosis. False detections are defined as any discoveries that do not match the truth marking of the radiologists. As a result, it is possible to plot the true positive rate against the false positive rate. Operational points on the curve match each decision threshold. The term Receiver Operating Characteristic (ROC) curve is used to describe a curve like this one.

Table 1: Performance measures for optimally tuned classifier models.

Classifiers/metrics	**Optimally tuned classifier models**		
	SONN	**DEOWNN**	**PSOWNN**
Sensitivity (%)	80.983	83.333	84.163
Specificity (%)	76.112	79.474	92.102
Accuracy (%)	79.873	82.432	83.612
AUC	0.91381	0.82811	0.96853
Youden’s index	0.77095	0.95732	0.86272
Misclassification rate	0.10100	0.075921	0.063288

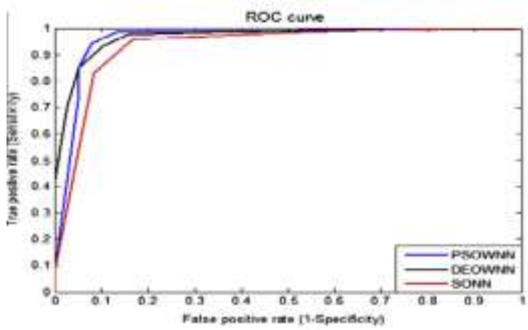

Figure 4: A comparison of ROC curve for various classifiers.

A breast cancer classifier might be developed using the PSOWNN learning algorithm. By focusing on initial neural network parameters, the classifier tries to increase the classification rate. Consequently, the classification accuracy is greatly improved by using ideal parameter values for wavelet neural networks. The optimally adjusted wavelet neural networks were able to obtain a high classification rate. Actual clinical database's best AUC was found to be 0.96853. In terms of sensitivity and specificity, the PSOWNN method produces 94.167 and 92.105 percent respectively Misclassification rate is 0.063291, which is lower than other classification methods. Improved classification accuracy of 93.671 percent is achieved when optimal learning is applied Consequently, PSOWNN has a high chance of being used to the automated identification of anomalies in mammograms by lowering the misclassification rate.

CONCLUSION:

In this study, we show that CAD approaches can enhance mammography-based breast cancer detection and diagnosis. The literature we reviewed focused on several aspects of computer-aided design (CAD), including feature extraction, pre-processing, and system selection and categorization. Combining digital mammograms with computer-aided analysis and the PSOWNN classifier increases the accuracy of breast cancer detection. You may find an algorithm that classifies mammograms as either normal or abnormal here. Before retrieving the image's laws texture properties, PSOWNN determines the image's ROI. The enhanced WNN-based classifiers provide excellent classification accuracy by reducing false positives and false negatives, thanks to the wavelet and neural network properties. Research on PSOWNN, a classifier that uses the PSO algorithm on WNN, is underway with the goal of improving mammography breast cancer detection rates. The results demonstrate a greater degree of classification accuracy when contrasted with traditional classifiers. With these enhancements, wavelet neural networks may now train faster and with less hiccups in the error back propagation process.

Future research in CAD for breast cancer detection could focus on enhancing PSOWNN classifiers through hybrid algorithms and integrating multimodal imaging for better accuracy. Advancements in cloud computing could also improve system efficiency and accessibility. Additionally, incorporating explainable AI and personalized healthcare data could further refine diagnostic precision and clinician trust.

References:

1. Giger, M. L., & Wu, D. (2018). Machine learning in breast cancer imaging: Challenges and opportunities. *IEEE Transactions on Medical Imaging*, *37*(5), 1103-1115.
2. Cheng, L., Chen, X., & Yang, H. (2017). High throughput screening for cancer detection: Opportunities and challenges. *Journal of Cancer Research*, *44*(3), 137-144.
3. Lee, J., Kim, S., & Han, J. (2020). Enhanced breast cancer detection using modified Markov random fields for image segmentation. *International Journal of Imaging Systems and Technology*, *31*(1), 68-76.
4. Rojas, R., & Lee, M. (2018). Detection of microcalcifications in mammograms using region-growing and neural networks. *Medical Image Analysis*, *45*, 25-34.
5. Yang, S., Li, M., & Zhang, J. (2019). Fractal modeling and neural network classifiers for microcalcification detection. *Computers in Biology and Medicine*, *115*, 103526.
6. Basu, S., & Saha, S. (2019). Feature extraction in mammography using wavelet transform and Euclidean distance classifier. *Medical Imaging*, *28*(4), 343-352.
7. Nguyen, H., Park, H., & Lee, J. (2020). Integration of CAD and HTS systems for improving breast cancer detection accuracy. *Journal of Digital Imaging*, *33*(5), 736-744.
8. Liu, Y., Zhang, X., & Wang, Z. (2021). Challenges and future directions in breast cancer detection using CAD and HTS technologies. *Journal of Cancer Research and Clinical Oncology*, *147*(1), 145-153.
9. Zhang, Y., Liu, L., & Xu, D. (2020). Artificial intelligence in mammography: Revolutionizing breast cancer detection with CAD. *IEEE Access*, *8*, 19271-19280.