IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14825

BREAST CANCER DETECTION FROM MAMMOGRAM USING MACHINE LEARNING ALGORITHMS

Darshan R[1], Ananya[2], Darshan M[3], Inchara R[4], Latha S[5]

Student, ECE, SJB Institute of Technology, Bengaluru, India[1]

Student, ECE, SJB Institute of Technology, Bengaluru, India[2]

Student, ECE, SJB Institute of Technology, Bengaluru, India[3]

Student, ECE, SJB Institute of Technology, Bengaluru, India[4]

Assistant Professor, ECE, SJB Institute of Technology, Bengaluru, India[5]

Abstract: Breast cancer is a profound global health challenge for women, where early detection is critical for saving lives, and while mammography is the standard screening tool, the manual interpretation of these images is a complex and often subjective task that can sometimes lead to errors. To address this, machine learning algorithms are now being developed as powerful aids, capable of analysing mammograms with advanced image processing to automatically identify subtle signs of malignancy, and by training these models on extensive datasets, we can create systems that achieve remarkable accuracy, offering a reliable, complementary tool that enhances traditional diagnostics and holds the transformative potential to improve early detection rates and patient outcomes worldwide.

Keywords: Machine Learning, Image Processing, Segmentation, Early Detection, Artificial Intelligence.

I. INTRODUCTION

Breast cancer remains one of the leading causes of morbidity and mortality among women globally, with early detection being critical for improving survival rates. Mammography is widely regarded as one of the most effective imaging techniques for the early detection of breast cancer, enabling the identification of Tumors and abnormalities that may not yet be clinically visible. However, interpreting mammograms requires expert knowledge, and even experienced radiologists can sometimes miss subtle signs of cancer, leading to false negatives or false positives. This challenge highlights the need for advanced computational methods to assist in the detection process.
This project focuses on using machine learning algorithms for the detection and classification of breast cancer in mammograms. The goal is to improve diagnostic accuracy, reduce human error, and enhance early detection, thus contributing to better patient outcomes.

1.1 Motivation of Work

A project on breast cancer detection using Machine learning's can serve as a crucial contribution to healthcare by harnessing the power of machine learning and image processing to aid early diagnosis. Breast cancer is one of the most prevalent cancers among women, and early detection significantly increases the chances of successful treatment. By working on this project, you not only deepen your technical skills but also contribute towards developing innovative solutions that can save lives by enabling faster and more accurate detection methods. Several problems occur in segmentation and classification of MRI images. The segmentation is useful in partitioning the image in meaningful regions. The misclassification or wrong segmentation leads to several problems in diagnosis of Tumour. The accurate classification of the breast Tumour is necessary to save the patient's life. The exact treatment can be given on the basis of segmentation and classification of Tumour. The existing systems provided significant improvement in Tumour detection but the accuracy level is low and high level of noise was presented. These issues motivate us to provide a segmentation and classification algorithm which will improve the accuracy. Here, the breast Tumour is detected by the use of medical imaging techniques. The main motivation for an engineer would be to propose a work to develop the system of breast Tumour segmentation and detection which would provide better performance parameters. Identifying different Tumour classes or subclasses with similar morphological appearances present is an interesting problem and has an important implication in Tumour diagnosis and treatment. Present technique includes 'biopsy' procedure which is operative in manner. Classification based on the imaging techniques is not acceptable by the radiologist and oncologist due to required accuracy. This motivates the engineers to develop an easy, quick and reliable algorithms to be implemented in the device which acts as a substitute for biopsy and produce accurate results.

1.2 Objective of work

- Early and Accurate Detection: Develop machine learning models capable of identifying breast cancer at its earliest stages with high sensitivity and specificity, reducing false negatives and false positives.
- Enhancement of Diagnostic Accuracy: Use machine learning to assist radiologists in interpreting mammograms, minimizing subjectivity and improving overall diagnostic precision.
- Feature Extraction and Analysis: Automate the process of extracting and analyzing elevant features (e.g., texture, shape, and density) from mammogram images for better classification of malignant and benign tissues.
- Integration of Diverse Data: Combine imaging data with patient-specific information (e.g., age, genetic predisposition, and medical history) to create personalized risk prediction models.
- Scalable and Cost-Effective Solutions: Design machine learning-based systems that are scalable and can be implemented in low-resource settings, making advanced diagnostic tools accessible worldwide.
- Real-Time Application: Enable real-time processing of mammograms to provide immediate diagnostic support and reduce the workload on healthcare professionals.
- Model Validation and Robustness: Validate the machine learning models on diverse and large datasets to ensure their generalizability and robustness across different populations.
- Reducing Human Error: Use automation to minimize the variability in human interpretation and enhance the reliability of breast cancer detection.
- Support for Decision-Making: Provide actionable insights and decision-support tools for oncologists and radiologists, aiding in effective treatment planning and monitoring.
- Contribution to Research: Advance the field of medical imaging and machine learning by contributing novel methodologies and algorithms tailored for breast cancer detection.

II. LITERATURE REVIEW

1. "Performance Comparison of Different Machine Learning Techniques For Early Prediction of Breast Cancer using Wisconsin Breast Cancer Dataset"

Authors: Atajan Rovshenov and Serhat Peker (IEEE-2022)

A significant health issue, cancer is becoming more prevalent globally and is a leading cause of mortality. Recent studies have shown that breast cancer is one of the most prevalent cancer type, particularly among women. Early detection can increase the chances of survival for those with breast cancer and lower treatment cost. However, there are drawbacks to the early diagnosis methods utilized in today's healthcare systems. These include the need for substantial human resources, long-term effects, and difficult access to these services for everybody. For early breast cancer diagnosis, technologies that are simple to use, yield reliable findings compared to scientific methodologies, and are available to everyone are required. Artificial Intelligence techniques enable the early diagnosis of breast cancer. This study aims to classify benign and malignant breast cancer image features. Artificial Neural Network, Support Vector Machine and Random Forest algorithms were used to classify features obtained from images. Experiments were performed on the Wisconsin Breast Cancer dataset. Experimental evaluation shows that 99% of the most successful results were achieved with the Artificial Neural Network algorithm. According to experimental findings, the classification technique can identify breast cancer in its early stages. The findings of the study are expected to shed on light new researches for investigation into breast cancer early detection.

2. "Breast Cancer Diagnosis Using Adaptive Voting Ensemble Machine Learning Algorithm"

Authors: Naresh Khuriwal and Nidhi Mishra (IEEE-2018)

According to Breast Cancer Institute (BCI), Breast Cancer is one of the most dangerous type of diseases that is very effective for women in the world. As per clinical expert detecting this cancer in its first stage helps in saving lives. As per cancer.net offers individualized guides for more than 120 types of cancer and related hereditary syndromes. For detecting breast cancer mostly machine learning techniques are used. In this paper we proposed adaptive ensemble voting method for diagnosed breast cancer using Wisconsin Breast Cancer database. The aim of this work is to compare and explain how ANN and logistic algorithm provide better solution when its work with ensemble machine learning algorithms for diagnosing breast cancer even the variables are reduced. In this paper we used the Wisconsin Diagnosis Breast Cancer dataset. When compared to related work from the literature. It is shown that the ANN approach with logistic algorithm is achieved 98.50% accuracy from another machine learning algorithm.

DOI: 10.17148/IJARCCE.2025.14825

3. “Breast Detection Using Nanoparticle Sensor with Machine Learning Algorithms”
Authors: J.N.V.R. Swarup Kumar, Charishma Karri, Naga Srihitha Vatsavayi, Harshitha Lekkala and Saaketh Choudarapu (IEEE-2024)
This research is mainly focused on the early detection of breast cancer in women by their urine samples with nanoparticle sensors. It detects certain enzymes and proteins that can be the main cause of cancer by machine learning algorithms, correlation analysis, and logical regression methods and created a web-based breast cancer prediction website using started nanoparticle urine analysis data, which contains the DNA barcode sequences that match with DNA of patient urine which is user friendly to use for every person. This study mainly focuses on non-invasive cancer. We analyze non-invasive cancer by using urine sample test data collected from the patients, utilizing advanced technology to access DNA signatures associated with breast cancer biomarkers. Our approach involves the barcode of nanoparticle sensors, which matches the urine samples. After the urine samples match, they are applied to the sensors, translated into digital data, and transmitted to a centralized system. Now, the centralized system collaborates with datasets derived from previous breast cancer cases. The algorithm now analyses the urine data and identifies the patterns. Then, it correlates with different stages of breast cancer.

4. “Using Supervised Learning for Breast Cancer Detection using AI&ML”
Authors: Pratyaksh Singh, Jaideep Nagill and Dr. Kavita Saini (IEEE -2023)
Breast cancer is the leading cause of cancerrelated deaths among women worldwide. A significant amount of research has been conducted to improve early detection of breast cancer, which is crucial for effective treatment and increased chances of survival. While mammograms have been the most reliable detection method, there is a need to explore alternatives that are cost-effective, safe, and accurate across different datasets. A hybrid paradigm of machine learning methods is presented in this work, proposed for effective breast cancer detection. The model combines several machine learning algorithms, including ANN, SVM, KNN and Decision Tree (DT), and can be applied to various data types, including images and blood tests. The proposed model aims to provide accurate results that are close to perfection.

5. “Breast Cancer Prediction System Utilizing Machine Learning Algorithms”
Authors: Chirayou Bista, Asreetha M, Salahuddin Slimanzay, Md Solaiman Sheikh and Dr. P Srinivasa Rao (IEEE-2024)
Breast cancer remains a significant global health concern, necessitating advanced predictive tools for early detection and effective prognosis. In response to this challenge, this research introduces a Breast Cancer Prediction System utilizing state-of-the-art machine learning algorithms. The system leverages the power of Random Forest, Support Vector Machine (SVM), and Gradient Boosting Ensemble to enhance accuracy and reliability. The Random Forest algorithm efficiently captures complex relationships within the breast cancer dataset, creating an ensemble of decision trees for robust predictions. Meanwhile, the Support Vector Machine optimally classifies data points by identifying hyperplanes, thereby enhancing the model's ability to discriminate between benign and malignant cases. Furthermore, the Gradient Boosting Ensemble technique synthesizes weak predictive models into a strong learner, boosting the overall predictive performance of the system. The Breast Cancer Prediction System offers a thorough and precise evaluation of the likelihood of breast cancer by combining the advantages of all three algorithms. By utilizing the discriminative power of the SVM, the boosting mechanism of the Gradient Boosting Ensemble, and the Random Forest's capacity to handle complex datasets, the system attains increased accuracy in early detection and prognosis. The suggested system has great promise for useful application in clinical settings, resulting in prompt interventions and better patient outcomes

6. “Exploring Machine Learning Techniques for Enhanced Breast Cancer Detection”
Authors: Nagesh Sharma and Sandeep Singh Kang (IEEE-2024)
Breast cancer is the second-greatest cause of death for women worldwide, affecting the majority of them. On the other hand, if cancer is identified early and adequately treated, it may be cured. Patients' chances of survival and prognosis can be greatly improved by early identification of breast cancer and prompt treatment intervention. Additionally, accurate benign tumor classification might assist patients in avoiding unnecessary therapy. This study provides a thorough overview of several studies that examined how ML algorithms may be used to find breast cancer. The main goal is to evaluate these algorithms' performance in terms of precision, accuracy, recall, and overall effectiveness. This evaluation tries to identify the most promising approaches and indicate areas for further development by looking at a wide variety of algorithms

7. “An Analysis of Ensemble Machine Learning Algorithms for Breast Cancer Detection: Performance and Generalization”
Authors: Rakesh Kumar, Meeta Chaudhry, H. K. Patel, Navin Prakash, Abhinav Dogra and Sunil Kumar (IEEE-2024)
Breast cancer is an explorative area, now a days, it is common dieses in females. So, the diagnosis of breast cancer for a patient at early stage can help to prevent their lives. The prediction of breast cancer by using machine learning can be

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14825

done after applying the various machine learning algorithms. Here, we applied ensemble machine learning algorithms and got better results in terms of performance and generalization. In this paper, the comparative analysis of Light Gradient Boosting (LightGBM) and Gradient Boosting have been done and experiments done by using labelled dataset of breast cancer. The LightGBM algorithm has been found to be less accurate in comparison to Extreme Gradient Boosting (XGBoost) algorithm.

III. DESIGN AND IMPLEMENTION

Input Images: Images are collected from a source for analysis.
Adaptive Median Filter: Both input images and database images go through an adaptive median filter, likely to reduce noise while preserving edges.
Segmentation: The filtered images undergo segmentation to isolate regions of interest.
Feature Extraction: Features are extracted from the segmented regions to create a feature set that represents the data.
Classifier: Extracted features are fed into a classifier that differentiates between two outcomes: Benign or Malignant.

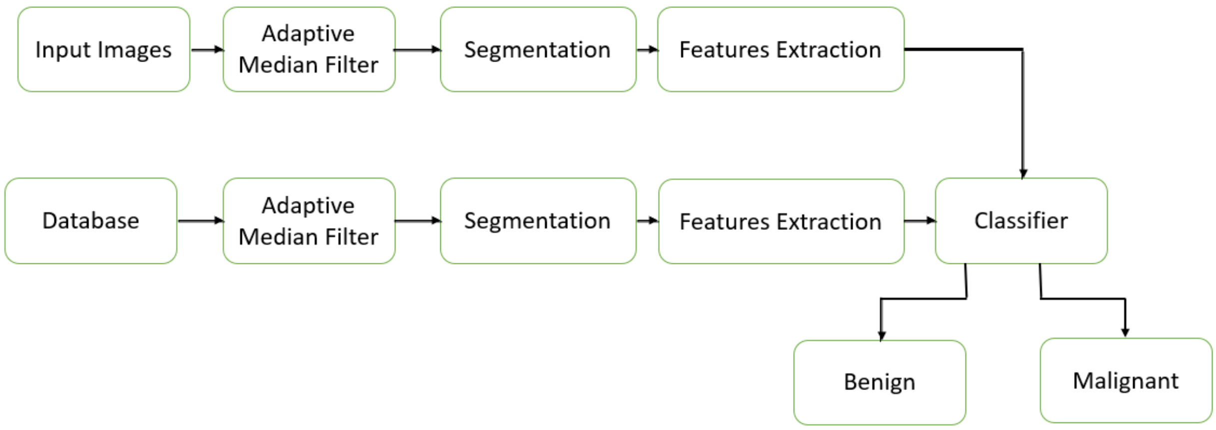

Fig 3.1.1 Block Diagram

3.1 IMAGE PROCESSING

Image processing is a technique that converts the ordinary image into digital form in order to enhance the quality of an image which gives useful information. The Medical image processing can be defined as picturing of body parts, tissues or organs for clinical analysis and treatment. It is one of the techniques used to create an image of the human body. Medical image processing is a highly challenging research area. The internal parts of the human body are diagnosed through medical imaging technique. Medical imaging has high importance because of correct diagnosis and treatment of diseases in health care system. The image of internal body parts where produced by the equipment's like CT scanner, MRI, etc. The Medical image processing consists of medical signal assembly, image developing, processing the image and image display to medical analysis based on feature extraction. The outlining, noise cleaning, search, filtering, de-blurring and texture analysis are some of the basic techniques of image processing. Image processing covers four main areas namely image analyzation, visualization, information management, and image formation. The noise in the images may produce inaccurate data which can be rectified in the image processing. The clear view of the image is very much useful in diagnosis of the disease. The image analyzation produces detail information about the image which help in noise reduction. The various steps in image processing is shown in Fig.3.1.1

3.2 PRE-PROCESSING

The pre-processing phase has a great importance in the applications of image processing and specially segmentation. Generally in the pre-processing phase, the main goal is to remove the noise from the images. Undoubtedly MRI images have noises which have to be removed. But the noise deletion shouldn't destroy the edges of the image and decrease the clarity and quality of it. There are several methods for removing noise, including: Gaussian filter, contourlet transform approach and wavelet thresholding approach, median filter, anisotropic diffusion filter. Anisotropic diffusion filter is a method for removing noise which is proposed by Persona and Malik. This method is for smoothing the image by preserving needed edges and structures. Fundamental idea is to adjust the smoothing level in a region based on the edge structure in the neighbourhood. Homogenous regions are highly smoothed and strong edge regions are barely smoothed (to preserve the structure).

DOI: 10.17148/IJARCCE.2025.14825

3.3 ALGORITHM USED:

The Integrated Mammogram Analysis Process: From Pixels to Diagnosis This process is a sophisticated pipeline of image processing and machine learning techniques designed to automatically detect and classify potential breast cancer from mammogram images. While the final classification is rule-based (using thresholds), the entire pipeline relies on foundational algorithms to prepare and analyze the data, many of which are inspired by or directly use probabilistic and statistical methods.

1. Image Preprocessing: Enhancing Data Quality

The raw mammogram is first prepared for analysis. This involves improving its quality and removing artifacts that could confuse the detection algorithms.
Histogram Equalization: This algorithm redistributes the pixel intensity values across the entire available range (0-255). It flattens and stretches the image's histogram, enhancing the contrast between tissues. This makes subtle features, like the spiculated edges of a malignant tumour, much more visible to subsequent steps.
Median Filtering: A non-linear digital filtering technique, excellent for removing "salt-and-pepper" noise while preserving crucial edges. It works by scanning the image with a small window (e.g., 3x3 pixels) and replacing the central pixel's value with the median value of the neighbours. This smooths the image without blurring the important boundaries that define potential tumours.

2. Feature Highlighting: Identifying Regions of Interest

This phase isolates potential abnormalities from the surrounding breast tissue.
Canny Edge Detection (The Gaussian Method): This is a multi-stage algorithm and a prime example of using the Gaussian distribution in image processing, a core concept in machine learning.
Gaussian Smoothing: The image is first smoothed using a Gaussian filter to reduce noise. This filter uses a Gaussian kernel (a bell-shaped curve) to compute a weighted average of pixels, effectively suppressing high-frequency noise.
Gradient Calculation: The algorithm then finds the intensity gradient of the image, highlighting regions with high spatial derivatives (i.e., edges).
Non-Maximum Suppression & Hysteresis Thresholding: The resulting edges are thinned and finalized by suppressing pixels that are not part of the strongest edges. The result is a clean, binary map of the most prominent edges in the mammogram.
Morphological Operations: The edge map often has gaps. Morphological operations are applied to refine it:
Dilation: Using a disk-shaped structuring element, this operation expands the white edge regions, bridging small gaps between them to form continuous contours.
Hole Filling: This operation fills the enclosed areas within these contours, resulting in solid white "blobs" that represent potential masses or calcifications.

3. Region Analysis: Feature Extraction

Each detected blob is now analysed as a distinct object.
Connected Component Labelling: The label algorithm scans the binary image and assigns a unique label to every connected group of white pixels. This allows the system to process each potential tumour individually.
Region Properties & Feature Calculation: For each labelled region, geometric properties are computed using region props:
Area: The size of the region.
Perimeter: The length of its boundary.
Circularity: A key shape descriptor calculated as (4 * π * Area) / (Perimeter^2). This metric is crucial:
A value near 1.0 indicates a perfect circle (often associated with benign cysts).
A value closer to 0 indicates a complex, irregular shape (a classic hallmark of malignant tumours, which grow invasively).

4. Classification: From Features to Diagnosis

This is the "supervised" decision-making step. While not a trained model like a Gaussian Naive Bayes classifier, it uses a rule-based approach that mimics a simple classifier based on expert-defined thresholds.
Threshold-Based Classification: Each region is classified by comparing its extracted features to pre-defined thresholds:
Normal Tissue: Regions that are too small (area < threshold area) or too irregular are dismissed as normal tissue structures.
Benign Tumour: A region that is sufficiently large and has a high circularity score is classified as a benign mass (e.g., a fibroadenoma).
Malignant Tumour: A region that is both *large* (area > malignant_area_threshold) and *highly irregular* (circularity < malignant_circularity_threshold) is flagged as potentially malignant.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14825

5. Visualization: Presenting the Results

The final step is to present the results in an interpretable way for a medical professional.

Bounding Box Visualization: The original mammogram is displayed with color-coded bounding boxes drawn around the detected regions:

Blue: Normal

Yellow: Benign

Red: Malignant

This provides an immediate, clear visual assessment of the algorithm's findings, effectively acting as a computer-aided detection (CAD) system.

IV. STEPS INVOLVED IN DETECTION

The Mammogram Analysis Process: A Step-by-Step Explanation

This process transforms a raw mammogram image into a diagnosed output, highlighting potential cancerous regions. It can be broken down into four main phases: Image Preparation, Feature Highlighting, Region Analysis, and Diagnosis & Output.

Phase 1: Image Preparation (Preprocessing)

The goal here is to clean up the image and improve its quality to make the subsequent analysis more accurate.

Loading and Conversion to Grayscale: The process starts by loading the original mammogram image. Since color information is not needed for anatomical analysis, the image is converted to grayscale, reducing complexity.

Contrast Enhancement (Histogram Equalization): Mammograms can often have low contrast, making it difficult to distinguish subtle details. This step adjusts the image's intensity values to spread out the most frequent intensity levels, significantly improving contrast and making potential abnormalities more visible.

Noise Reduction (Median Filtering): Medical images can contain random noise (speckles or graininess). A median filter is applied to remove this noise while preserving the important edges of structures within the breast. It works by scanning the image and replacing each pixel's value with the median value of its neighbours, effectively smoothing the image without blurring edges.

Fig 4.1.1 Image Preprocessing

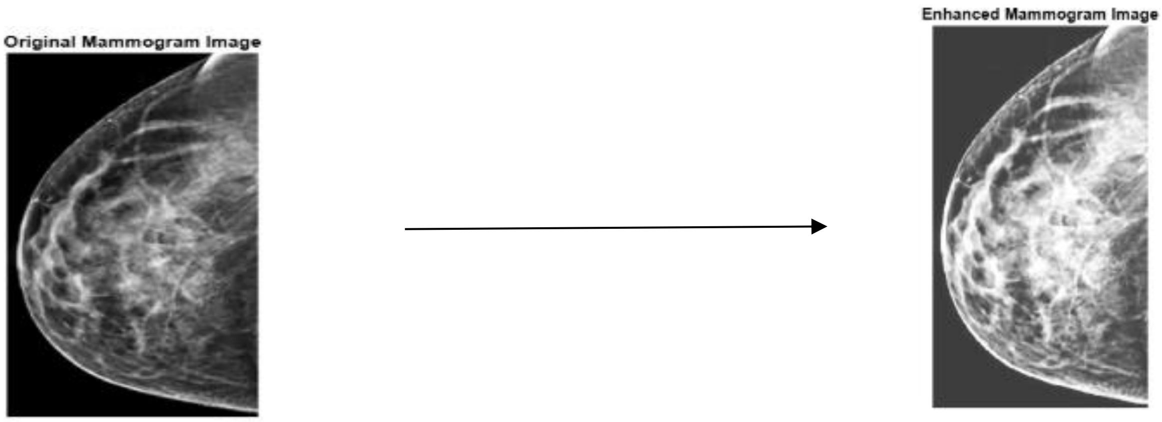

Fig 4.1.2 Image enhancement

Phase 2: Feature Highlighting (Segmentation)

This phase aims to isolate and identify distinct shapes and regions within the enhanced image.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14825

Edge Detection (Canny Algorithm): This is a crucial step for finding the boundaries of all structures in the image. The Canny algorithm is a sophisticated method that identifies points of sharp intensity change, outputting a binary image where white lines represent the edges of tissues, potential masses, or calcifications.
Morphological Operations (Dilation and Filling): The edge-detected image often has broken or discontinuous lines. To create complete, solid shapes, two operations are performed:
Dilation: This thickens the white edge lines, helping to bridge small gaps between them.
Filling: This operation looks at the closed contours created by the dilated edges and fills their entire interior, resulting in solid white "blobs" or regions of interest (ROIs).

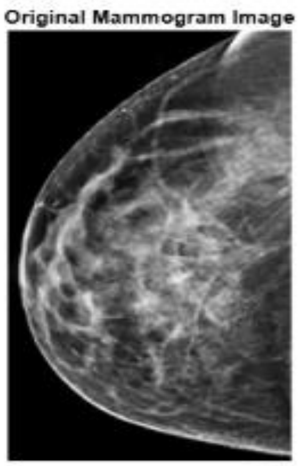

Fig 4.2.1Originaal Image

Fig 4.2.2: Edge detection

Phase 3: Region Analysis (Feature Extraction)
Now that potential regions are identified, the system measures their properties to understand their characteristics.
Labelling and Measurement: Each separate white blob is labelled as a unique object. The system then calculates key geometric properties for each one:
Area: The total number of pixels in the region.
Perimeter: The distance around the boundary of the region.
Circularity: A calculated metric $(4 * \pi * \text{Area}) / (\text{Perimeter}^2)$ that describes the shape's roundness. A perfect circle has a circularity of 1.0, while irregular, spiculated shapes have a value closer to 0.

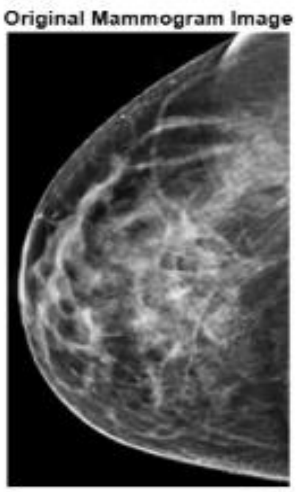

Fig 4.3.1Original Image

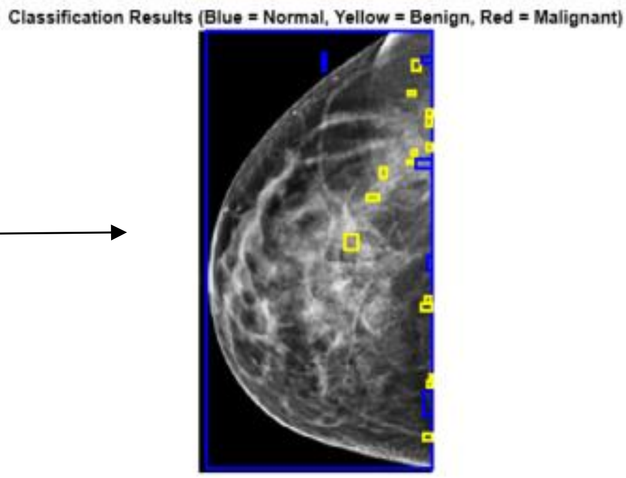

Fig 4.3.2 classification results

Phase 4: Diagnosis and Output (Classification)
The extracted features are used to classify each region as normal, benign, or malignant.
Classification Based on Thresholds: The system uses simple rule-based logic to make a diagnosis:
Any region too small or too irregularly round is ignored as normal tissue.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 14, Issue 8, August 2025
DOI: 10.17148/IJARCCE.2025.14825

A region that is sufficiently large and fairly round is flagged as a potential benign (non-cancerous) mass, like a fibroadenoma.
A region that is both very large and highly irregular (low circularity) is classified as potentially malignant (cancerous), as cancers often grow in an invasive, spiculated pattern.
Visualization and Result: The final result is presented by superimposing color-coded bounding boxes onto the original mammogram image:
Blue Box: Normal tissue (no cause for concern).
Yellow Box: A detected benign mass.
Red Box: A detected potential malignant tumour. This provides a clear, immediate visual aid for a radiologist to review the algorithm's findings. A message box also pops up to directly state the final diagnosis of whether the detected region is Benign or Malignant.

Fig 4.4.1 Detection results

V. RESULTS

Machine learning (ML) algorithms, particularly deep learning models like Convolutional Neural Networks (CNNs), have significantly improved breast cancer detection from mammograms. These models offer higher accuracy, sensitivity, and specificity compared to traditional methods, often reducing false positives and false negatives. ML can also detect subtle patterns that radiologists may miss, leading to earlier cancer detection. In addition, ML-based automated screening systems have been integrated into clinical practice, helping radiologists prioritize cases and speed up diagnoses. Combining mammography with other imaging modalities like ultrasound or MRI further enhances detection accuracy. However, challenges such as data imbalance, model generalization, and interpretability remain. Despite these, ML's potential in early detection and improving patient outcomes is increasingly recognized in healthcare settings.

VI. CONCLUSION AND FUTURE WORK

The application of machine learning algorithms in breast cancer detection using mammograms has significantly improved diagnostic accuracy, sensitivity, and efficiency. Traditional algorithms like SVM and Random Forest provide reliable classification, while advanced deep learning models, particularly CNNs and transfer learning approaches, have set new benchmarks in performance, achieving accuracies as high as 99%. Despite challenges such as data imbalance, model interpretability, and generalizability, ongoing research and the development of explainable AI are addressing these limitations. When integrated into clinical workflows, these models serve as powerful decision-support tools, complementing radiologists and enhancing the overall effectiveness of breast cancer screening programs. As the field continues to evolve, machine learning is poised to play an increasingly vital role in early cancer detection and improving patient outcomes. Future work in the field of machine learning (ML) for breast cancer detection using mammograms focuses on addressing current limitations and expanding capabilities. One key area of research is developing more robust models that generalize well across diverse populations, imaging systems, and clinical settings. This involves training on larger, more diverse datasets and integrating advanced techniques like domain adaptation to ensure consistent performance globally.

REFERENCES

[1]. "Performance Comparison of Different Machine Learning Techniques For Early Prediction of Breast Cancer using Wisconsin Breast Cancer Dataset" by Atajan Rovshenov and Serhat Peker (IEEE-2022)
[2]. "Breast Cancer Diagnosis Using Adaptive Voting Ensemble Machine Learning Algorithm" by Naresh Khuriwal and Nidhi Mishra (IEEE-2018)
[3]. "Breast Detection Using Nanoparticle Sensor with Machine Learning Algorithms" by J.N.V.R. Swarup Kumar, Charishma Karri, Naga Srihitha Vatsavayi, Harshitha Lekkala and Saaketh Choudarapu. (IEEE-2024)

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14825

[4]. "Using Supervised Learning for Breast Cancer Detection using AI&ML" by Pratyaksh Singh, Jaideep Nagill and Dr. Kavita Saini (IEEE -2023)

[5]. "Breast Cancer Prediction System Utilizing Machine Learning Algorithms" by Chirayou Bista, Asreetha M, Salahuddin Slim anzay, Md Solaiman Sheikh and Dr. P Srinivasa Rao (IEEE-2024)

[6]. "Exploring Machine Learning Techniques for Enhanced Breast Cancer Detection" by Nagesh Sharma and Sandeep Singh Kang (IEEE-2024)

[7]. "An Analysis of Ensemble Machine Learning Algorithms for Breast Cancer Detection: Performance and Generalization" by Rakesh Kumar, Meeta Chaudhry, H. K. Patel, Navin Prakash, Abhinav Dogra and Sunil Kumar (IEEE-2024)

[8]. https://www.who.int/cancer/en/

[9]. https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.12763

[10]. https://www.cancer.gov/types/breast/patient/breast-treatment-pdq

[11]. Genetic Factors in Breast Cancer Jhonson & Lee 2019

[12]. Trends in Breast Cancer Incidence Smith et al 2021

[13]. Psychosocial Impacts Robinson & White 2022

[14]. Advancements in Radiation Therapy Thompson et al 2023