

Breast Cancer Detection using Image Processing and Machine Learning: A Comprehensive Review and Improved Segmentation Approach

D.D.H. Erandika* and U.A. Piumi Ishanka

Department of Computing and Information Systems, Faculty of Applied Sciences,
Sabaragamuwa University of Sri Lanka, Belihuloya

erandikahansi95@gmail.com

Abstract

Breast cancer stands as one of the most prevalent health concerns for women. Early detection of breast cancer can significantly improve the chances of survival. In developed countries, more than 19.9% of women will die per year due to breast cancer. Regular breast cancer screening is an important way to detect cancer early. Image processing techniques are highly used for different types of cancer detection applications with medical screening. Segmentation of the breast tumor region is a critical step in image processing related to this manner. Lots of research work can be found on developing ways for detecting breast cancers. However still, there is a need for a standard and robust cancer region segmentation method. Right cancer region segmentation is significant for better feature extraction and better classification. This study presents a comprehensive literature review about technologies used with image enhancement and tumor segmentation. Further, the study proposes a robust approach to image enhancement and cancer region segmentation for breast cancer detection using image processing techniques and machine learning. In this work, mammograms are enhanced using Contrast Limited Adaptive Histogram Equalization and denoised using the Median blurring filter. This study enables an effective image segmentation approach with two stages: removing the background using thresholding, tumor region segmentation using a combination of thresholding, and Canny edge detection. Segmented tumor region's features are extracted using the Gabor filter. Here the accuracy of the approach is compared with three main machine learning classifiers: Decision Tree, Random Forest, Multinomial Logistic Regression and mammograms are classified into three classes (malignant, benign and normal). Finally, an ensemble approach is proposed using the hard voting mechanism to improve the accuracy. With the dataset of mini-MIAS this proposed approach achieved 78.89% accuracy. Results demonstrated that the proposed breast cancer detection approach improves the performance of segmentation breast tumor regions.

Keywords: Mammogram, Image enhancement, image segmentation, Image processing, Machine learning

1 Introduction

Cancer is a life-threatening disease that affects the majority of people worldwide. Thousands of people die every day as a result of cancer all over the world. Breast cancer is one of the leading causes of death in women around the world. When analyzing past research studies, it can be recognized that if a woman has a family history of this disease, she is also at risk of developing breast cancer..

By identifying breast cancer at an early stage, the mortality rate may be reduced. If breast cancer is diagnosed in an early stage, it can start treatment before spreading. The majority of breast cancer detection methods are based on

medical images. Mammography, thermal infrared imaging, microscopic slide images, ultrasound, MRI, and CT scans are some image acquisition techniques for breast cancer detection. Traditional cancer detection and diagnosis approaches rely on professional doctors using medical imaging to identify signs and symptoms that typically occur in the later stages of the disease.

In order to detect an early stage, it must correctly find out the tumor region. It may mostly depend on the segmentation techniques. Although research on breast cancer detection using digital image processing is not new, there is a standard and robust segmentation approach that still needs to be improved. Even so, advancements in diagnosing and treating it are still costly and

time-consuming. Here are some major challenges that were encountered in previous studies related to breast cancer detection.

- The object of interest can be quite small, resulting in the possibility of misidentification.
- Because mammograms indicate unique sizes, shapes, and types of microcalcifications, sample matching appears to be difficult.
- The Region Of Interest (ROI) may have low contrast. The distinction between questionable areas and the tissues that envelop them can be narrow.
- Suspicious regions are virtually undetected due to dense tissues and skin thickness, especially in young women.
- Dense tissues can be mistaken for calcifications, resulting in a significant number of false-positive cases.
- There is no standard and robust way to segment the tumor region from the breast image.

Image processing techniques consist of five basic steps as image acquisition, image pre-processing, image segmentation, image feature extraction, and image classification. In image processing, image acquisition is the first process that retrieves images from an external source for further processing. These images that are collected from various kinds of sources can be of different sizes. As well as there may be a lot of noises with those images and maybe in low contrast. So, as the second stage, image pre-processing is used to format images before training and testing models. This includes resizing, enhancing image quality, removing noises, removing unwanted parts like processes. Image segmentation is the method of dividing an image into separate regions based on various criteria. Image segmentation algorithms are used to split and group a specific set of pixels from an image. Image segmentation is important in image analysis since it covers detection, feature extraction, classification, and treatment. Segmentation assists physicians in quantifying the amount of tissue in the breast for treatment planning. Feature extraction is a process that is used to define the related shape details found in a pattern to make the task of classifying the pattern easier. The primary aim of

feature extraction is to extract the most important details from the original data and display them in a lower-dimensional space. Image classification algorithms hold the main role in image processing techniques. It is used to classify all pixels in a digital image into one of several categories. This approach pays high attention to a proper image enhancement method and image segmentation method to segment tumor regions. This paper is divided into five sections and the first section gives a brief overview of breast cancer detection. Then in the literature review section, some major past studies related to image enhancement techniques and image segmentation techniques for breast cancer detection will be discussed. In the methodology section, the proposed approach will be explained. Performance analysis of the approach is explained in the Results and Discussion section. Finally, the conclusion section provides a brief idea about the proposed approach, its result and suggestions for future works.

2 Literature Review

The approaches on breast cancer detection using mammogram images can include many activities like image enhancement, breast segmentation, mass detection, and microcalcification detection. When considering the past studies related to those activities, it can be identified that all of those have their own literature. Here a comprehensive literature review about image segmentation techniques used in the studies during the past decade with used datasets will be presented. Further, the mostly used image enhancement techniques to improve the quality of images will be analyzed.

2.1 Image enhancement techniques

When considering past reviews about image enhancement techniques for breast cancer detection, various kinds of techniques can be found. Histogram-based techniques, frequency-based techniques, fuzzy-based techniques, and filter-based techniques are some of them.

In a prior study in 2012 [1], there was a comparison between four image enhancement techniques on mammogram images. authors have analyzed unsharp masking, Contrast-Limited Adaptive Histogram Equalization (CLAHE), morphological operators, and wavelet-based enhancement. Mammogram images are collected from patients' medical records in Isfahan and the Contrast Improvement

Index (CII) used to evaluate the performance of image enhancement algorithms. It is calculated by dividing the contrast value of the processed images by original images. According to this study the wavelet transformation has been displayed.

In another study, some filters which use image preprocessing [2]. In this study, average filter, adaptive median filter, average or mean filter, and wiener filter on mammograms were analyzed. These filters were analyzed for mammogram images with salt pepper, Gaussian, and Speckle noises. In this study the performances were compared using mean square error, Peak signal to noise ratio, structural content and normalized absolute error parameters. According to this study, it is concluded that the adaptive median filter is a more appropriate method when compared with other filters.

K. Akila et al. in a prior study, presented a comparative study for some indirect contrast enhancement techniques in mammogram images [3]. In this study, Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), Brightness Preserving Bi-Histogram Equalization (BBHE), Minimum Mean Brightness Error Bi-Histogram Equalization (MMBEBHE), Recursive Mean Separate Histogram Equalization (RMSHE) were analyzed by evaluating Effective Measure of Enhancement (EME) and Peak Signal to Noise Ratio (PSNR). According to this study, for noise removal MMBEBHE provided better results and CLAHE and RMSHE provided better enhancement results when compared with others.

A new approach for noise removing and image enhancement is described in a study [4]. In this approach Fast Discrete Curvelet Transform via UneqiSpaced Fast Fourier Transform (FDCT via USFFT) is used for noise reduction and Modified Local Range Modification (MLRM) is used for contrast enhancement. In this study, many mammogram images have been tested, and as the results, increased contrast decreased mean square error, and increased peak signal to noise ratio when compared to existing methods. When testing this method with mammogram images this approach has given better results compared with existing methods.

In 2018, there was an approach for breast cancer detection [5]. Here the authors have used contrast improvement, noise lessening, texture scrutiny and

portioning algorithm. Here in image preprocessing stage images are converted to gray scale and noise removal was done using sharpening of the images and median filter. This study has shown that median and log transformation give better results when compared with others. They also suggested increasing the contrast and visibility of images using FCLAHE technique. Time saving, precision and reiterations are the main characteristics in this approach.

Jyoti Dabass et al. in 2019 [6], has proposed an approach for improving the quality of digital mammogram images. Here, contrast limited adaptive histogram equalization (CLAHE) and entropy-based intuitionistic fuzzy techniques were used with this study for contrast enhancement. The authors have used metrics like Discrete Entropy (DE) and Peak signal to noise ratio (PSNR) to measure the performance. Also they have compared the performance of this approach with another advanced fuzzy set that is a type II fuzzy set. They have shown that their proposed approach improves the visuality of mammogram images and it is an efficient method for correct elucidation of digital mammograms for breast cancer detection. It was an applicable approach for pre-processing all categories of breast images including dense glandular mammograms, fatty glands, and fatty mammograms.

T.Chaira et al. in 2020, proposed a mammogram enhancement approach using Atanassov's intuitionistic fuzzy set (IFS) [7]. IFS theory is most suitable for image enhancement with mammogram images due to low contrast of them and unclear image definitions. Here firstly, images were transformed to an intuitionistic fuzzy image using a novel intuitionistic fuzzy generator. Then the hesitation degree was determined, and two membership levels were computed using the hesitation degree to construct an interval type 2 fuzzy set. Then these two membership functions were combined using Zadeh's fuzzy t-conorm to form a new membership function. The restricted equivalency function was used to find the threshold of an interval type 2 fuzzy image. Modified fuzzy hyperbolization was performed using the threshold. This approach was performed well for mammogram enhancement with the real data.

More recent research has proposed an algorithm for the enhancement of breast tumors in mammograms and ultrasound images [8]. This study

consists of two steps. The first one is to identify various noises in images and the second one is for enhancing with algorithms like mean, median, Homomorphic, Weiner, Modified Gabor, Anisoid diffusion filter, Contrast Limited Adaptive Histogram equalization (CLAHE), Bilateral filter, and Guided filter. When compared with other filters the modified Gabor filter (Anisotropic diffusion with Gabor filter) is better in performance. The following table summarizes the previous studies related to image enhancement for breast cancer detection.

2.2 Image segmentation techniques

In the past studies about image segmentation for breast cancer detection, image segmentation based classical approaches, image segmentation based machine learning approaches, and image segmentation based deep learning methods can be identified. When considering classical approaches, there can be edge-based segmentation methods, threshold-based segmentation methods, region-based segmentation. Canny edge detection, Sobel, contour, energy minimisation are mostly found in edge detection techniques. Otsu thresholding, adaptive thresholding, morphological thresholding, global and local thresholding are generally used techniques as thresholding-based techniques. Watershed, partial region growing, rough set theory, and marker controller are mainly used region-based segmentation techniques.

When considering image segmentation based machine learning approaches, supervised machine learning methods and unsupervised machine learning methods can be identified. Support vector machine (SVM) is a mostly used supervised machine learning method and fuzzy C-clustering, k-means clustering, and hierarchical k-clustering are used as unsupervised machine learning.

U-Net, SegNet, Convolutional neural networks are the mostly used breast cancer segmentation-based deep learning approaches. Figure 1 shows the mammogram tumor segmentation techniques.

2.3 Image segmentation based on classical methods

In 2012, there was a region growing approach for mass segmentation in mammogram images [9]. The main activities in this approach were segmenting the breast from mammogram images without background, enhancing the contrast of images and

separating mass region from rest of the breast and. This segment breast from mammogram was done using manual thresholding technique. The closest value for thresholding was set manually by the user. For image enhancement and noise reduction used contrast limited adaptive histogram equalization (CLAHE). Then, mass segmentation was done using region growing based on local statistical texture analysis. Here, smoothness descriptor data was used to automatically detect the seed point.

Tolga Berber et al in 2013 [10], presented a new approach to enhance the breast mass boundary in ROI of mammogram that is called breast mass contour segmentation. That was based on the classical seed region growth. This study was tested using mammogram images from the Dokuz Eylul Mammography Set (DEMS). There were four steps with this approach as ROI detection, preprocessing, mass size estimation and segmentation. For ROI detection the authors have assumed that bounding rectangles of a mass region are drawn by the user and for preprocessing they used median, averaging and Laplacian filters. Mass size was estimated using OTSU segmentation. The performance was evaluated using matrices such as specificity, sensitivity, balanced accuracy, Yassnoff and Hausdorff error distances. Also this approach was evaluated with three other region based segmentation methods. They were level-set segmentation, seeded region growing segmentation and watershed segmentation. With this approach the authors could achieve 95.06

Neto et al. [11] presented an automatic mass segmentation approach in mammogram images that was based on thresholding and evolutionary algorithms. There were five stages in this approach as image acquisition, pre-processing, segmentation, reduction of false positives and mass and non-mass classification. The images were collected for this study from the Digital Database for Screening Mammography-DDSM database. In order to remove noises, edges, pectoral muscle and other artifacts, the authors used a method proposed by [12]. Then enhanced the linear structures of images using a method based on the histogram and a mean filter. In the segmentation phase Mean Standard Deviation of the Image(msdi) was calculated. Then applied the Otsu algorithm to find the first threshold of the image and divide it into two clusters. The Particle Swarm Optimization (PSO) was used to optimize, evolutionarily, the

Table 1: Previous studies related to image enhancement for breast cancer detection

Author	Image Source	Techniques/ Algorithms	Results
Hajar Moradmand et al. [1] (2012)	Patients' medical records, Isfahan	Unsharp masking, Contrast-Limited Adaptive Histogram Equalization (CLAHE), Morphological operators, Wavelet-based enhancement	Wavelet transformation perform better enhancement on mammograms
R. Ramani et al. [2] (2014)	MIAS database	Average filter, Adaptive median filter, Average or mean filter, Wiener filter.	The comparison showed the adaptive median filter is a more appropriate method than others.
K.Akila et al. [3] (2015)	MIAS database	HE, CLAHE, BBHE, MMBEBHE, RMSHE	CLAHE and RMSHE provide better enhancement compared to the other methods
Senthilkumar, B. et al. [4] (2017)	MIAS database and DDSM	FDCT via USFFT, MLRM	Increased contrast, decreased mean square error, and increased peak signal to noise ratio when compared to existing methods
Yousif M.Y Abdallah et al. [5] (2018)	Mammogram images	Median, Log transformation, Wiener function, Watershed transform	Proposed time saving, precision and reiteration approach
Jyoti Dabass et al. [6] (2019)	MIAS	CLAHE, Entropy-based intuitionistic fuzzy method	Removed the uncertainty in boundary or region of interest without affecting the information content and shape of the image
T.Chaira et al. [7] (2020)	Mammogram images	Intuitionistic fuzzy set, Global thresholding	Performed well with low illuminated mammogram images
S.V.Shwetha et al. [8] (2020)	Mammogram and ultrasound images	Mean, median, Homomorphic, Weiner, Modified Gabor, Anisotropic diffusion filter, Contrast limited Adaptive Histogram equalization (CLAHE), Bilateral filter and Guided filter	Modified Gabor filter achieved better performance on comparison with the other

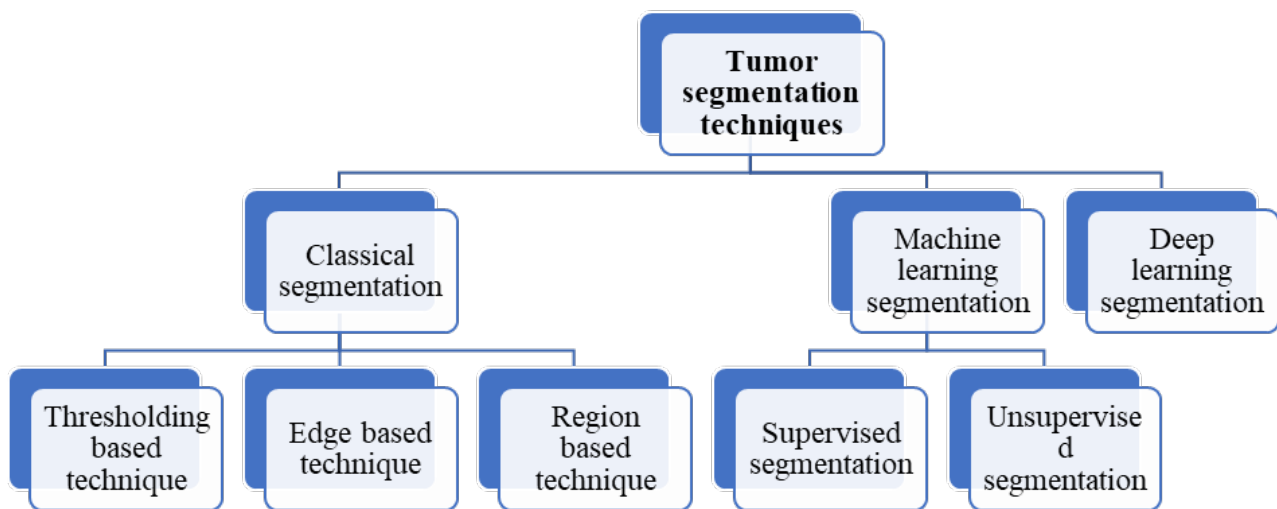


Figure 1: Mammogram tumor segmentation techniques

search for the global maximum of the thresholds in order to achieve a better segmentation. In the Reduction of False Positives stage, region growing was applied to isolate the Region of Interest (ROI) of the object, and after that applied an area filter and the Graph Clustering. This proposed approach was most suitable for being safely integrated to a computer-aided detection tool.

In a study, morphological threshold-based segmentation techniques were used for mass detection in mammograms [13]. In this work, firstly images were converted to gray-scale and created binary images using thresholding techniques. Then used modified region growing technique to remove the pectoral muscle. Then, images were enhanced using median filter and contrast limited adaptive histogram equalization technique (CLAHE). After that for ROI extraction, the Sobel operator was applied for gradient calculation and watershed transformation was used for finding the regional minima. Then the morphological operations such as opening and closing were applied to minimize the regions. Then the level-wise thresholding was applied to select the appropriate threshold point. This study was analyzed using 55 mammogram images from the Mini-MIAS database. In this approach the mass detection accuracy in mammograms was 94.54% and false positive rate was 5.54%. When compared with the traditional OTSU thresholding method, the proposed method achieved more effective results.

In 2016 [14], proposed a new method for segmentation and classification masses from mammogram images. Preprocessing, segmentation, feature extraction and classification were main stages in this approach. In preprocessing, border clearing and contrast enhancement were done. Then to detect regions of interest (mass) applied the Otsu's threshold technique and morphological operations like erosion and opening. Then area, perimeter, major axis, minor axis, dispersion, compactness, circularity, roundness, elongation, shape index and entropy, thinness ration, eccentricity, equivdiameter like texture and shape features were extracted from segmented regions. After extracting these features, they were normalized and ranked using absolute value two-sample t-test for getting significant features. This approach was evaluated using mammogram images from the Women Health Care Program (WHC) and the Digital Database for Screening Mammography (DDSM). The result of this study was analyzed using ANN, K-Nearest Neighbor and SVM classifiers. According to the result of this approach, it could achieve accuracy of 100% and 90% for the WHC and DDSM datasets. When comparing the classifiers ANN performed better than the KNN and SVM.

In another study A. Rampun et al. [15] proposed a new method for breast boundary and pectoral muscle segmentation in mammograms using edge detection techniques. Accurate mammogram image, 2D modeling, breast alignment, remove

artifacts, noise reduction, initial breast boundary, active contour, post processing were main processes in this approach. Noises were removed using a combination of median and anisotropic diffusion filters. A 2D breast model was used to find the threshold value to create a binary mask. In the breast alignment stage, if the breast is located on the right hand side, they flip. Then always the pectoral muscle was located at the top left of the image. Canny edge detection was used to get all possible edges and edge features were used to detect the most probable pectoral contour for the pectoral boundary estimation. The results of this approach were evaluated using the Mammographic Image Analysis Society (MIAS), INBreast and Breast Cancer Digital Repository (BCDR) databases. In this work, they could achieve γ coefficients of 98.8% and 97.8% (MIAS), 98.9% and 89.6% (INBreast) and 99.2% and 91.9% (BCDR) respectively.

In [16] proposed an approach for automatic detection of breast masses using an optimized region-growing technique. This study was focused on developing a robust method for generating an optimal seed point and threshold for the optimized region growing using Dragon Fly optimization technique. In this work, mammogram images were preprocessed using a Gaussian filter. Then Regions of Interest (ROI) were extracted using optimized region growing techniques. In feature extraction of segmented ROI, Gray level co-occurrence Matrix (GLCM) and Gray level Run Length Matrix (GLRLM) techniques were used to extract texture features. Then feed them to a Feed Forward Neural Network (FFNN) classifier with a back propagation learning algorithm. This approach was tested using mammogram images from Digital Database for Screening Mammography (DDSM) database. This approach achieved 98.1% sensitivity and 97.8% specificity with 300 images.

In 2019 there was an approach for detect tumor regions more accurately [17]. In this work different classical segmentation methods were used for segment tumor regions. Mammogram and MRI images were used for analyze this approach. In this work, firstly images were converted to grayscale. Then applied Gaussian filter for noise removing, smoothing and enhancing the visual quality of images. After that histogram analysis has done for get complete tonal distribution of image. Then images were segmented using

thresholding and edge detection methods to extract tumor region. Then these thresholding and edge detection method were compared with MRI and mammogram images by measuring the entropy. According to this study, it is obvious that edge detection technique provides better results for MRI and mammogram images. Then this study was extended to analyze the performance of Sobel, Canny, and Prewitt operators which are edge detection operators. When considering this analyze, Canny operator provide better entropy than n Sobel and Prewitt with MRI and mammogram images.

When considering the recent approaches, an approach can be found that uses threshold-based segmentation for normalizing the image and the morphological operation for segmentation of different parts of mammograms [18]. 19 cancerous images and 22 noncancerous images are done for this study. Also, an FCM classifier is used for image classification. This research presented higher accuracy than 80 percent and got better accuracy than k-means based classification approach when compared with previous studies.

In [19], proposed an approach to breast tumor detection using Improved Awareness Probability-based Crow Search Algorithm (IAP-CSA)-based Region growing. There were four phases in this study. They are image pre-processing, tumor segmentation, extracting features, and detection. In image pre-processing, a median filter was used to remove noise and enhance the images. For tumor segmentation, an optimized region growing algorithm was used that was optimized using IAP-CSA. GLCM and GLRM techniques were used for feature extraction from the segmented tumor regions. Then, modified Fuzzy Logic classifier used the classification procedure. The classification accuracy was improved by optimizing the membership function of the classifier with the proposed IAP-CSA algorithm. The novelty of this study was optimized region growing method and optimized Fuzzy classifier. In this study, the performance of IAPCSA-fuzzy was compared with the past studies like PSO-fuzzy, GWO-fuzzy, WOA-fuzzy, and CSAfuzzy. According to the result IAP-CSA-based fuzzy achieved the accuracy 41.9% improved than the fuzzy classifier, 2.80% improved than PSO, WOA, and CSA, and 2.32% improved than GWO-based fuzzy classifiers. It also achieved better accuracy than NN, SVM and other existing

fuzzy classifiers .

Neeraj Shrivastava et al. in their study [20] proposed an approach for breast tumor detection using Efficient Seeded Region Growing segmentation technique. This algorithm was divided into four parts. First part is image preprocessing. ROI extraction, contrast enhancement, noise removing were the main purpose of this part. Here to extract ROI background and the pectoral muscle removal were done. Thresholding was applied to background removing. Then region growing technique was applied to pectoral muscle removal . Here the seeder point was automatically calculated using the information about orientation about mammograms . In this process, if an image is left oriented, , the left upper pixel is selected as a seed point, otherwise the right upper pixel is selected as a seed point. Then using this seed point, seed region growing image segmentation technique was applied to remove the pectoral muscle mammogram images. After the contrast adjustment, a median filter was applied to remove salt and pepper noise. The second part of this algorithm was selecting an automatic threshold value for binarization. It was obtained from calculating the intensity differences of 4-neighbor pixels. The mean intensity difference of this was selected as threshold. In this work, performance were analyzed using a set of matrices such as TPF, True Negative Fraction (TNF), False Positive Fraction (FPF), False Negative Fraction (FNF), Dice Similarity Co-efficient (DSC), Relative Overlap (RO), Sum of true volume fraction (SVTF), Misclassification Rate (MCR), Accuracy, Precision (P), and F-Measure. This proposed approach achieved 94.6% and 95.3% True Positive Fraction for MIAS and DDSM datasets respectively and achieved better performance for accuracy, precision, and F-measure. However, this work did not differentiate the tumor type as benign or malignant.

Previous studies related to image segmentation for breast cancer detection that will be discussed in this study are presented in the following table.

2.4 Image segmentation based on machine learning methods

Cordeiro F.R. et al. in 2012 [21] presented an automatic approach for detecting tumor regions using the Extreme Learning Machine (ELM) neural network. There were two metrics used to evaluate the accuracy of the classifier. They were

the global accuracy index, the confusion matrix, and the Kappa index. When compared with traditional networks the ELM had a fast-learning phase, only having a need for a single iteration, besides requiring less configuration parameters are advantages.

There was an approach based on K-means clustering for eliminating the triangular area of pectoral muscle in 2014 [22]. There were several image processing steps with the approach . Firstly, images were identified as left or right oriented, and labels, artifacts were removed in the background. Then image morphology and image filtering techniques were used to remove noise and artifacts . For enhancement the contrast of images used histogram equalization and contrast enhancement. Then for image binarization thresholding technique was used. After that, the biggest blob that contains the pectoral muscle, fatty tissues, ligaments, sinus, lobules and ducts, is extracted from the original image. Then pectoral muscles were removed using image morphology and K-means clustering algorithm. This approach achieved 94.4% true positive, 5% false positive and 1% false negative results.

Neeraj Dhungel et al. [23] proposed a novel approach to mass segmentation based on a structured support vector machine (SSVM). This study used a potential function based on deep belief networks (DBN) that can learn complex features directly from mammograms. In this work, to improve the contrast of images, the authors have adopted the preprocessing that was described in [24]. They used the Dice index (DI) for measuring the segmentation accuracy quantitatively. Here the ROI was taken by a manual annotation of the mass center and scale. This proposed approach achieved better performance than other state-of-the-art methods with 0.87 Dice Index and 0.8s computation time.

Arnau Oliver et al in their study [25], proposed an automatic density segmentation method based on a supervised pixel-based classification. In this work, for automatically detecting the pectoral muscle used a method by [26]. That method was applicable with most of mammograms and, pectoral muscle in rest of images were manually removed. Then applied enhancement method, that was proposed by [27]. For validation this approach the authors have used mammogram images from the Spanish screening programme specifications

Table 2: Image segmentation for breast cancer detection based on classical methods

Author	Image Source	Techniques/ Algorithms	Results
Nor Ashidi Mat Isa et al. [9] (2012)	Mini MIAS database	Thresholding, CLAHE, Smoothness descriptor, Region growing	Achieved excellent segmentation results with 94.59% of sensitivity and 3.90 false positive per image
Tolga Berber et al. [10] (2013)	DESM dataset	Median filter, Averaging filter, Laplacian edge enhancement filter, OTSU segmentation, Seeded region growing (SRG) algorithm	Minimized false positive pixels while obtaining high accuracy rates and least distance values to reference region
Neto et al. [11] (2015)	DDSM database	Histogram technique, Mean filter, Mean Standard Deviation, Otsu algorithm, PSO, Area filter, Graph Clustering	Proposed suitable approach for being safely integrated to a computer-aided detection tool, Got 95.2% accuracy
Aziz Makandar et al. [13] (2015)	Mini MIAS database	Median filter, CLAHE, Thresholding technique, Region growing method, Sobel operator, Watershed transformation	Achieved more effective results than traditional Otsu thresholding method
Basma A. et al. [14] (2016)	WHC database, DDSM database	Otsu's threshold technique, Morphological operations, Global thresholding technique	Achieved accuracy of 100% and 90% for the WHC and DDSM datasets, ANN showed better results than KNN and SVM classifiers for this approach
A. Rampun et al. [15] (2017)	MIAS database, INBreast database, BCDR database	Median filter, Anisotropic diffusion filters, 2D breast model, Canny edge detection, Otsu's thresholding, KNN, SVM, ANN	Achieved high accuracy over a range of evaluation metrics, Performed better results than existing methods in the literature
S.Punitha et al. [16] (2018)	DDSM database	Dragon Fly optimization technique, Gaussian filter, Optimized region growing method, GLCM, GLRLM, Feed Forward Neural Network	Achieved 98.1% sensitivity and 97.8% specificity with 300 images
P.Sahni et al. [17] (2019)	Mammogram, MRI	Gaussian filter, Thresholding methods, Edge detection methods, Histogram analysis, Sobel operator, Canny operator, Prewitt operator	Edge detection algorithms achieved better result, Canny operator showed better results than Sobel and Prewitt
P.B.Chanda et al. [18] (2020)	Mammograms	Threshold-based segmentation, Morphological operation, FCM classifier	Achieved higher accuracy than 80%
Rajeshwari S. Patil et al. [19] (2020)	DDSM database	Optimized region growing algorithm, Median filter, GLCM, GLRM, Modified Fuzzy Logic classifier	IAP-CSA-based fuzzy achieved better accuracy than NN, SVM and other existing fuzzy classifiers
Neeraj Shrivastava et al. [20] (2020)	MIAS database, DDSM database	Thresholding technique, Region growing technique, Median filter, Seeded region growing, Morphological operating	Achieved better performance, Had the capability of automatic selection of seed point and threshold

database. The main limitation of this work was computation cost on SVM and computation time. They claimed that their approach was outperformed threshold-based approaches or clustering approaches.

In 2018 [28], proposed a method for tumor segmentation based on a fuzzy set and thresholding technique. In preprocessing images were converted to binary using thresholding techniques. Then images were converted to fuzzy images and used morphological operators. In this work, noises were removed using a Gaussian filter. Then found all edges using thresholding technique. When compared with classic morphology the algorithm that adopted the Fuzzy Morphology has shown more effectiveness in diagnosing the tumor boundary.

Poulomi Das et al. in their study [29], proposed an automated masses segmentation approach using kernel based fuzzy c-means (FCM) clustering technique. This kernel based fuzzy c-means clustering technique captured significant regional features of mammograms which are responsible for identifying the presence of masses even in vague and diffused mammograms. In this approach the significant regional features, regional entropy and brightness of mammograms were intelligently used to segment the masses. The performance of this study was evaluated by comparing the existing methods of clustering-based segmentation techniques such as fuzzy c-means clustering and k-means clustering approaches. This approach allowed for a more accurate capture of the suspicious area. Furthermore, dense glandular tissues in mammograms might sometimes occupy similar brightness of masses for significant X-ray attenuation. However, because the textural distribution of those glandular tissues is random, the proposed technology could completely suppress them.

The table below depicts the previous studies related to image segmentation based on machine learning for breast cancer detection that will be analyzed in this study.

2.5 Image segmentation based on deep learning methods

In 2015 there was a new approach based on deep learning for mass segmentation using Tree-reweighted belief propagation [30]. To increase the contrast of images, the authors have used

a preprocessing method that was described in [24]. This approach was implemented using a conditional random field model (CRF). Accuracy of this study was tested with Dice index (DI). The proposed approach could achieve an 89.0

M.S. Hossain in [31] proposed a microcalcification segmentation method in mammograms using a modified U-Net segmentation network. Image preprocessing, segmentation of breast regions, extraction of suspicious patches, selection of positive patches and training the segmentation network were the five stages of this study. In this study, images were enhanced using Laplacian filters and for removing artifacts, minor gradient noises were eliminated by the image erosion method. Then pectoral tissues were removed using K-means pixel-wise clustering. High pass filter was applied to enhance the contrast of the suspicious region. Here a fuzzy C-means clustering method was used to detect the optimal threshold for the frequency histogram method. After automatically identifying the suspicious regions, the positive and negative marks for the bounding boxes were observed manually. To train the segmentation network, positive boxes were fed to the network which was modified and developed based on the conventional U-net segmentation network.

In another study proposed a method for mass segmentation based on deeply supervised U-Net (DS-U-Net) [32]. Here contrast limited adaptive histogram equalization (CLAHE) used contrast enhancement. DS U-Net used for segmentation of lesions. DS U-Net was also used to improve the recovery of object boundaries and to reduce false positives. It identified the location of the mass in mammograms. The dense conditional random fields (CRF) model was used to fine-tune the resulting segmentation map of DS-U-Net. The authors achieved 82.9% and 79% dice scores for CBIS-DDSM and INBREAST databases respectively. When compared with other models DS U-Net worked better from mammogram segmentation.

V.K.Singh et al. in their study [33], proposed an approach for breast tumor segmentation and shape classification in mammograms using the cGAN and convolutional neural networks. Before feeding images to the network, the Single Shot Detector (SSD) [34] was applied to the mammogram images to locate tumor positions and fit bounding boxes around them. The losing

Table 3: Image segmentation for breast cancer detection based on machine learning methods

Author	Image Source	Techniques/ Algorithms	Results
Cordeiro F.R. et al. [21] (2012)	Mini-MIAS database	ELM network,	ELM provided an over 81% classification rate
Nashid Alam et al. [22] (2014)	Mini-MIAS database	Image morphology, Histogram equalization, Contrast enhancement, Median filter, Thresholding technique, K-means clustering algorithm	Provided about 94% acceptable segmentation results, Detection step is most time consuming
Neeraj Dhun-gel et al. [23] (2015)	DDSM-BCRP database, INbreast database	SSVM, Potential Functions,,	Overcomed other state-of-the-art methods with 0.87 Dice Index and 0.8s computation time
Arnau Oliver et al. [25] (2015)	SSPS database	SVM, Median filter,	Outperformed threshold-based approaches or clustering approaches, Performed better result than neural network and the boosting classifier
A.M. Salih et al. [28] (2018)	Mini-MIAS database	Threshold technique, Fuzzy morphology algorithm, Gaussian filter	Fuzzy Morphology has showed more effectiveness in diagnosing the tumor boundary than classic Morphology
Poulomi Das et al. [29] (2019)	MIAS database, DDSM database, BIRADS database	Kernel based fuzzy c-means (FCM) clustering technique	Showed extensive feature based analysis and classification outcomes than other clustering-based segmentation techniques

frame was applied for cropping tumor areas and healthy regions. The cGAN was segmented a breast tumor within an ROI in a mammogram. In preprocessing the ROI, Gaussian filter was used for noise removal and histogram equalization used to contrast enhancement. Then applied normalization to rescale images. After feeding these processed images to the cGAN, they could get a binary mask of the tumor by using morphological operations. Then classifying the binary masks was done using a CNN based shape descriptor.

3 Methodology

3.1 Image acquisition

The early-stage tumors rarely cause harm to patients, and mammography is the only established approach for consistently detecting these tumors. It can also identify all forms of breast cancer, including invasive ductal and invasive lobular cancers. For this study, mammogram Images are collected from the mini-Mammographic Image Analysis Society (mini-MIAS) database. It is generated by an organization of a UK research interest group. All of these images are in the center of the matrix and PGM file format (Portable Gray Map). Those mammogram images are categorized into three classes as benign, malignant, and normal.

3.2 Image preprocessing

In this stage, images are resized and converted to grayscale. If all of the images are the same size, computations will use the same memory space for all of them. If the sizes of the images are different, memory will need to be rearranged. So, when this happens it will take more time. If there is insufficient memory, the procedure may potentially fail. So, image resizing is a critical step in preprocessing.

Mammogram images are low in contrast. So before using further processes we have to improve the quality of images. Here we applied After resizing images are Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance the quality of images. It is a variation of Adaptive Histogram Equalization (AHE) that addresses contrast over-amplification.

Then noises are removed using the median blur filter. Median blurring is more successful at eliminating salt-and-pepper type noise from

an image because each central pixel is always replaced with a pixel intensity that exists in the image. Furthermore, because the median is robust to outliers, salt-and-pepper noise will have less influence on the median than other statistical measures, such as the average. Figure 2 shows the processes in the image preprocessing stage.

3.3 Image segmentation

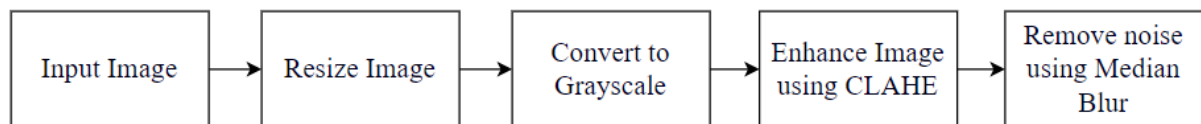
In this study, image segmentation is done in two steps. There are some labels and letters with mammogram images. Therefore, as the first step of image segmentation breast images are segmented from removing the unwanted labels in the background using thresholding techniques. The second step in image segmentation is to segment tumor regions from the breast by removing the pectoral muscles in the upper area of the mammogram images.

The below processes are applied for the first image segmentation step which is unwanted label removal.

1. Choose a suitable value for thresholding using the `threshold_otsu()` method.
2. Get a thresholding image. For this, we use `cv2.threshold()` function in OpenCV. The formula of this function is, `cv2.threshold(src, thresh, maxval, type[, dst])`. Here,
 - `src` - this is the grayscale source image.
 - `thresh` - threshold value which is used to classify the pixel values.
 - `maxval` - maximum value which is used with the `THRESH_BINARY` and `THRESH_BINARY_INV` thresholding types.
 - `type` - thresholding type. Maybe one of
 - `cv2.THRESH_BINARY`
 - `cv2.THRESH_BINARY_INV`
 - `cv2.THRESH_TRUNC`
 - `cv2.THRESH_TOZERO`
 - `cv2.THRESH_TOZERO_INV`
 - `dst` - output array of the same size and type as `src`
3. Get all external contours of the thresholding image and keep the maximum one.
4. Make a mask in the same size and same data type with the threshold image and draw the maximum contours on it.

Table 4: Image segmentation for breast cancer detection based on deep learning methods

Author	Image Source	Techniques/ Algorithms	Results
Neeraj Dhun- gel et al. [30] (2015)	Inbreast database, DDSM-BCRP database	Tree Re-weighted Belief Propagation, Potential Functions	Gained in terms of accuracy and efficiency in comparison with the current state of the art
M.S. Hossain [31] (2019)	DDSM database,	Laplacian filter, Erode method, K-means pixel- wise clustering, High pass filter, Fuzzy C-means clustering, U-net	Achieved 98.2% average accu- racy with better performance than state-of-the-art methods
Ravitha et al. [32] (2020)	CBIS-DDSM database, INBREAST database	CLAHE, Histogram equal- ization, DS U-Net, Dense CRF model	Performed better than AU- Net and U-Net
V.K. Singh et al. [33] (2020)	DDSM database, IN- breast dataset, Hos- pital Sant Joan de Reus dataset	Single Shot Detector, Losing frame, cGAN, Gaussian filter, Histogram equalization, Morphological operation	Achieved a Dice coefficient of 94.0% and an intersection over union (IoU) of 87.0%., Outperformed is the current state-of-the-art with overall accuracy of 80%.

**Figure 2:** Pre-processing stage in the proposed approach

5. Extract breast from the image using bitwise_and function in OpenCV. Formula of this function is, result = cv2.bitwise_and(src1, src2, mask=mask). Here

- src1: the first object for merging
- src2: the second object for merging
- mask: optional operation mask. The first object and the second object are merged without getting the black region of the mask.

Then the second segmentation step is segmenting tumor regions by removing the pectoral muscles in the upper area of the mammogram images. In this work, a method was used to do this by combining a threshold-based segmentation technique with an edge-based image segmentation algorithm. By applying the THRESH_BINARY thresholding the areas of tumors and the pectoral

muscles area can be segmented because pixel values of those tissues are higher than other tissues in the mammogram images. the threshold_yen() function was applied for finding the threshold value for this. It is an automatic thresholding method [28]. Figure 4 (b) shows the result of using this thresholding technique.

After that, the canny edge detection technique which is given by OpenCV was utilized to detect the edges of the images. Canny edge detection is most sensitive to noises. So before using this, applied a 5x5 Gaussian filter to remove the noise. Figure 4 (c) shows the detected edges of a mammogram image. Then got external contours using these edges for creating a mask to detect the upper area of the image. Using this mask can segment the tumors of the mammogram as shown in figure 4 (d).

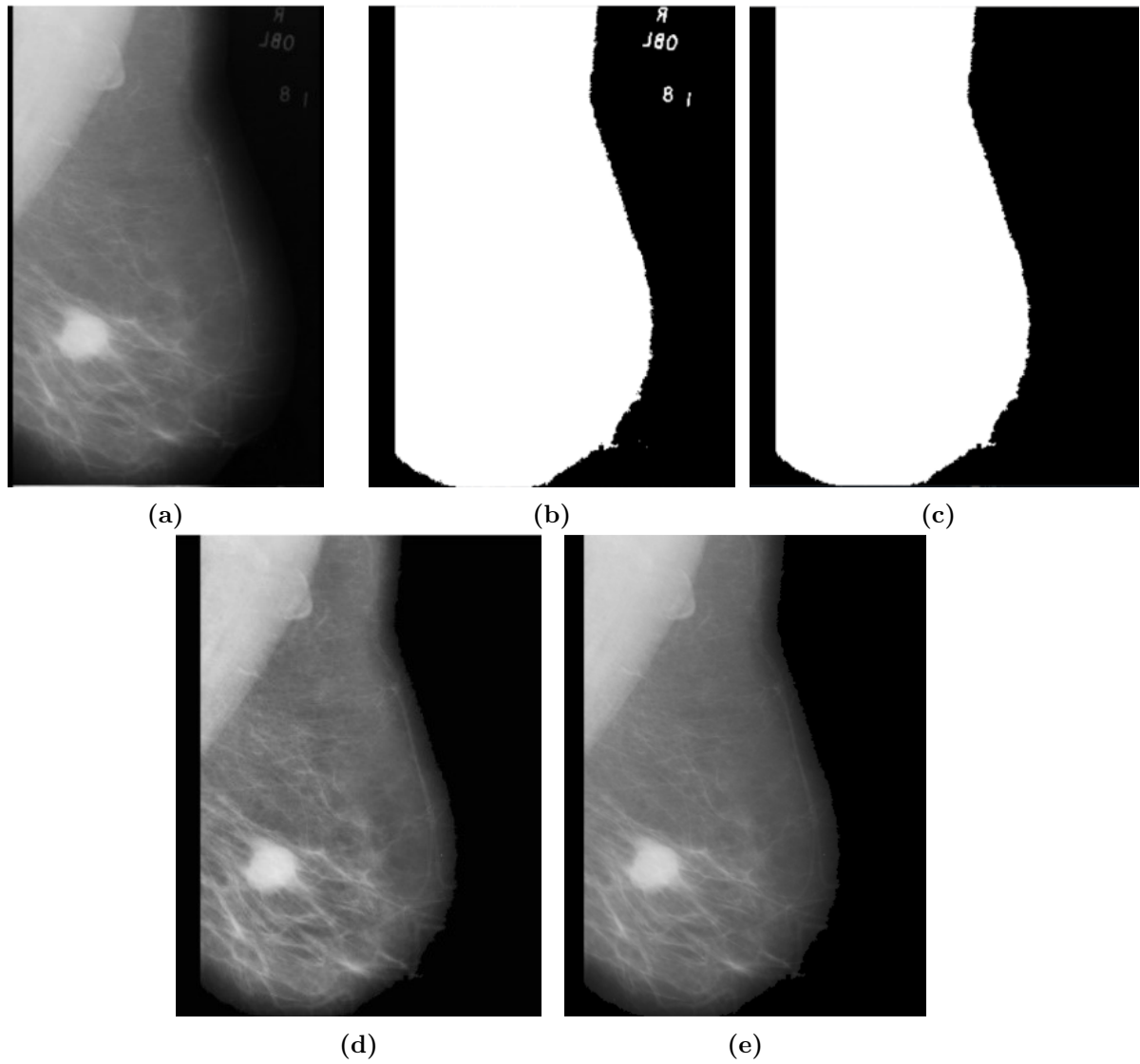


Figure 3: mdb028 from the Mini-MIAS database (a) Original image (b) Binary image (c) Mask to crop labels (d) Label-free image (e) Image enhancement using CLAHE

3.4 Feature extraction

In this work, the texture features in those images are extracted using the Gabor filter. From this filter, a set of features that can be used for the next image classification phase can be created. A bank consists of many Gabor filters that have been adjusted with various parameter settings, resulting in quite varied filter responses. In general, parameter settings are chosen in a data-independent manner.

3.5 Classification

The last phase in this proposed approach is classifying the mammogram images into three classes. In this study, an ensemble approach was employed, which is a technique used in machine

learning to combine multiple models in order to produce a more accurate and robust solution. The principle behind this technique is that by leveraging the strengths and compensating for the weaknesses of individual models, one can achieve better performance. The ensemble consists of three base machine learning classifiers: Random Forest, Decision Tree, and Multinomial Logistic Regression. Each of these classifiers offers a unique approach to the classification task:

- **Random Forest:** This is an ensemble of decision trees itself, typically trained with the "bagging" method. It is known for reducing overfitting, handling large data sets with higher dimensionality, and its ability to handle missing values.

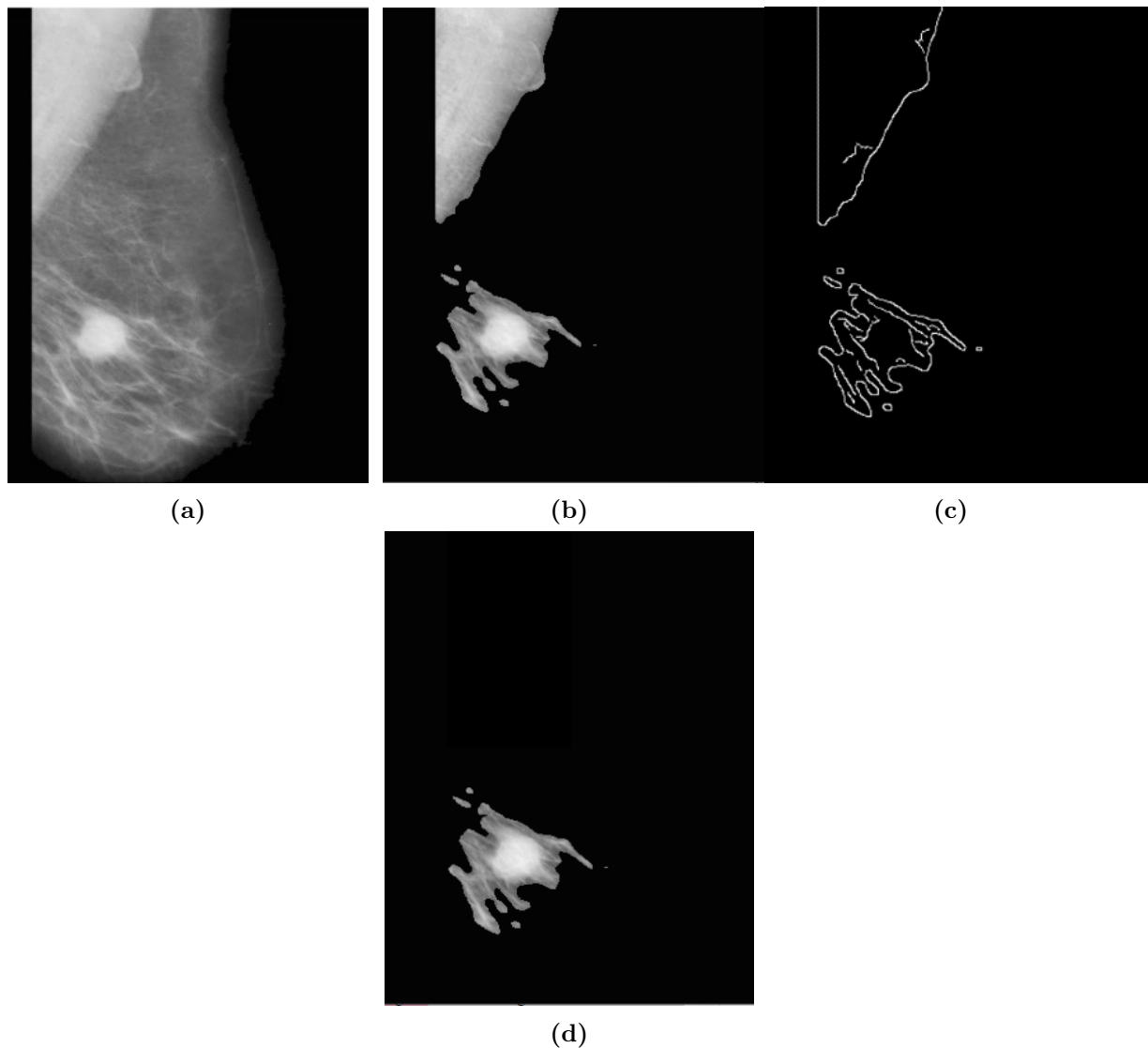


Figure 4: (a) Background removed image (b) Threshold image (c) Edge detected image using canny filter (d) Upper pectoral muscles cropped image

- **Decision Tree:** A flowchart-like structure in which internal nodes represent features, branches represent decisions, and each leaf node represents an outcome. They're simple, interpretable, and can handle both categorical & numerical data.
- **Multinomial Logistic Regression:** It is an extension of logistic regression, suitable for scenarios with more than two classes in the target variable. It models the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables.

Once each classifier produces its prediction, a 'hard voting' mechanism was applied. In hard

voting, each individual classifier in the ensemble 'votes' for a class, and the class that gets the most votes is chosen as the final prediction. The rationale behind this is to capitalize on the wisdom of the 'crowd'; individual errors from classifiers can be compensated by the correct predictions of others, leading to a more reliable and often more accurate final prediction.

This ensemble approach, combined with the hard voting mechanism, is aimed at enhancing the robustness and accuracy of the breast cancer detection system. By leveraging the different strengths of each individual classifier, we intend to achieve higher predictive accuracy than any single classifier could achieve on its own.

Table 5: Classification result for Experiment 1

Machine Learning Classifier	Accuracy
Decision Tree	68.91%
Random Forest	63.51%
Multinomial Logistic Regression	60.81%
Support Vector Machine	56.75%
Naive Bayes	55.4%
K-Nearest Neighbor	47.92%

4 Results and Discussion

The proposed approach is developed from completing several experiments. Images in the mini-MIAS database are used for all experiments. In this section, the result of major experiments will be discussed. The first major experiment is that these mammogram images are classified without using any image enhancement or noise removing technique.

Here features are extracted using image pixels. For classification, several machine learning classifiers were selected.

According to the result of the table 5, Decision Tree, Random Forest and Multinomial Logistic Regression were selected for further development of the selected research approach.

In the second main experiment, several feature extraction techniques for this study were used which are mostly used in previous studies. When considering the result of table 6 the Gabor filter was selected as the feature extraction technique in our study.

Then various image enhancement techniques were utilized with this study and when comparing the images with tested technologies it is obvious CLAHE were enhanced images better. Figure 5 depicts how contrast is enhanced with applying CLAHE for an image.

Then in the next main experiment, CLAHE was applied to enhance the quality of images and the Median blur filter for denoising. Then applied a proposed image segmentation technique to remove the background and segment tumor region. The results for this experiment are displayed in table 7.

Then finally, these processed images are classified using the ensemble approach. The hard voting mechanism was applied for this experiment. The

accuracy result of applying the ensemble approach for the classification is shown in table 8.

When comparing the results of table 8, it is clear that the highest accuracy has got the hard voting mechanism with Decision Tree and Multinomial Logistic Regression machine learning classifiers.

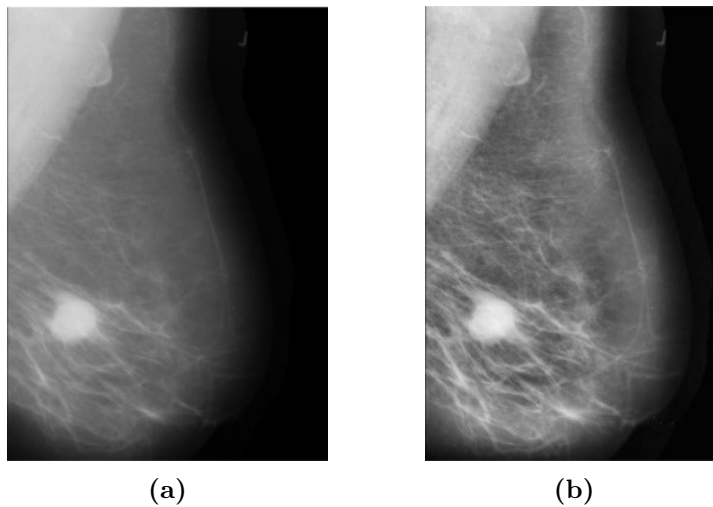
When considering the difference in the accuracy before and after applying the proposed approach, the accuracy is improved using this approach.

5 Conclusion

Early detection and early treatment are the main solutions for reducing the mortality rate of breast cancer. When considering the early detection of breast cancer, mammography which is an image acquisition technique plays a critical part. There are various kinds of tissues, labels, and noises in a mammogram. Therefore, when using mammograms for breast cancer detection, we have to remove noises, labels in the background, pectoral muscles like things and segment tumor regions from the mammogram. This study is conducted around the mini-MIAS database and image enhancement is done using CLAHE, noise removal is done using Median blur. Unwanted labels in the mammogram are removed using global thresholding techniques. For this, the closest value for thresholding is found using the threshold_otsu method which is provided by OpenCV. After that cancer tumor region is segmented using an approach that combines the threshold-based segmentation method and edge-based segmentation method. As a threshold-based technique, global thresholding is applied with finding the closest value to threshold by the threshold_yen method that is provided by OpenCV. As an edge-based method, the Canny edge detection technique is applied. Before applying the Canny method, images are denoised and smoothed using a Gaussian blurring filter. Here we used an ensemble approach which is the hard voting mechanism to improve the classification accuracy. The combination of Decision Tree and Multinomial Logistic Regression classifiers gave the highest accuracy with the hard voting mechanism for this approach. In future, improvement in the accuracy of this approach is expected by adjusting the parameters of the Gabor filter as appropriate with these images and extracting more features from the images.

Table 6: Classification result for Experiment 2

Feature Extraction Technique	Decision Tree	Random Forest	Multinomial Logistic Regression
Gabor filter	70.27%	65.89%	63.72%
Sobel filter	63.51%	62.16%	63.15%
Gaussian filter	64.86%	64.86%	62.86%
Gabor filter + Sobel filter	68.91%	58.10%	63.21%
Gabor filter + Gaussian filter	70.07%	60.81%	63.18%
Sobel filter + Gaussian filter	68.89%	59.45%	62.21%

**Figure 5:** (a) Original image before contrast enhancement (b) Contrast-enhanced image**Table 7:** Classification result with the proposed approach

Machine Learning Classifier	Accuracy
Decision Tree	75.91%
Random Forest	69.41%
Multinomial Logistic Regression	68.17%

Table 8: Classification result with the hard voting mechanism

Machine Learning Classifier	Accuracy
Decision Tree + Random Forest	78.75%
Random Forest + Multinomial Logistic Regression	74.94%
Decision Tree + Multinomial Logistic Regression	78.89%
Decision Tree + Random Forest + Multinomial Logistic Regression	77.61%

References

- H. Moradmand, S. Setayeshi, A. R. Karimian, M. Sirous, & M. E. Akbari (2012) Comparing the performance of image enhancement methods to detect microcalcification clusters in digital mammography, *Iran. J. Cancer Prev.*, vol. 5, no. 2, pp. 61–68, Spring 2012.
- R. Ramani, N. Suthanthira Vanitha & S. Valarmathy (2013) The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images, *I.J. Image, Graphics and Signal Processing*, vol. 5, no. 5, pp. 47-54.
- K. Akila, L. S. Jayashree, & A. Vasuki (2015) Mammographic image enhancement using indirect contrast enhancement techniques – A comparative study, *Procedia Comput. Sci.*, vol. 47, pp. 255–261, 2015.
- B. Senthilkumar, R. Gowrishankar, M. Vaishnavi, & S. Gokila (2017) Combination of noise removal and contrast enhancement methods for the preprocessing of mammogram images-towards the detection of breast cancer, *Biosci. j.* (Online), pp. 1653-1658.
- Yousif Mohamed Y. Abdallah, Sami N.A. Elgak, Hosam Zain, Mohammed Rafiq Rafiq, Elabbas Ebaid and Alaeldein Elnaema (2018) Breast cancer detection using image enhancement and segmentation algorithms, *Biomedical Research*, vol. 29, no. 20, pp. 3732-3736.
- J. Dabass, S. Arora, R. Vig, & M. Hanmandlu (2019) Mammogram image enhancement using entropy and CLAHE based intuitionistic fuzzy method, in 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN).
- T. Chaira (2020) Intuitionistic fuzzy approach for enhancement of low contrast mammogram images, *Int. J. Imaging Syst. Technol.*, vol. 30, no. 4, pp. 1162–1172, 2020.
- S.V. Shwetha, L. Dharmanna & S.B. Anami (2020) Design and methodology of algorithm for the enhancement of breast tumor images, In *IOP Conference Series: Materials Science and Engineering*, vol. 994, no. 1, p. 012020), IOP Publishing.
- N. A. M. Isa and T. S. Siong (2012) Automatic segmentation and detection of mass in digital mammograms, in *Proceedings of the 11th international conference on Telecommunications and Informatics, Proceedings of the 11th international conference on Signal Processing*, 2012, pp. 143–146.
- T. Berber, A. Alpkocak, P. Balci, & O. Dicle (2013) Breast mass contour segmentation algorithm in digital mammograms, *Comput. Methods Programs Biomed.*, vol. 110, no. 2, pp. 150–159, 2013.
- O. P. S. Neto, O. Carvalho, W. Sampaio, A. Correa, & A. Paiva (2013) Automatic segmentation of masses in digital mammograms using particle swarm optimization and graph clustering, in *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2015, pp. 109–112.
- W. B. Sampaio, E. M. Diniz, A. C. Silva, A. C. de Paiva, & M. Gattass (2011) Detection of masses in mammogram images using CNN, geostatistic functions and SVM, *Comput. Biol. Med.*, vol. 41, no. 8, pp. 653–664, 2011.
- A. Makandar & B. Halalli, (2016) Threshold based segmentation technique for mass detection in mammography, *J Comput*, 11(6), 472-478.
- B.A. Mohamed, N.M. Salem, M.M. Hadhoud, & A.F. Seddik (2016) Automatic segmentation and classification of masses from digital mammograms, *Advances in Image and Video Processing*, vol. 4, no. 4, p. 17.
- A. Rampun, P. J. Morrow, B. W. Scotney, & J. Winder (2017) Fully automated breast boundary and pectoral muscle segmentation in mammograms, *Artif. Intell. Med.*, vol. 79, pp. 28–41, 2017.
- S. Punitha, A. Amuthan, & K. S. Joseph (2018) Benign and malignant breast cancer segmentation using optimized region growing technique, *Futur. Comput. Inform. J.*, vol. 3, no. 2, pp. 348–358, 2018.

- P. Sahni & N. Mittal (2019) Breast cancer detection using image processing techniques, in Lecture Notes in Mechanical Engineering, Singapore: Springer Singapore, 2019, pp. 813–823.
- P. B. Chanda & S. K. Sarkar (2020) Detection and classification of breast cancer in mammographic images using efficient image segmentation technique, in Lecture Notes in Electrical Engineering, Singapore: Springer Singapore, 2020, pp. 107–117.
- R. S. Patil & N. Biradar (2020) Improved region growing segmentation for breast cancer detection: progression of optimized fuzzy classifier, *Int. J. Intell. Comput. Cybern.*, vol. 13, no. 2, pp. 181–205, 2020.
- N. Shrivastava & J. Bharti (2020) Breast tumor detection in digital mammogram based on efficient seed region growing segmentation, *IETE J. Res.*, pp. 1–13, 2020.
- Cordeiro, Lima, Silva-Filho, & Santos (2012) Segmentation of mammography by applying extreme learning machine in tumor detection, in *Intelligent Data Engineering and Automated Learning - IDEAL 2012*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 92–100.
- M. S. Hossain (2019) Micro calcification segmentation using modified U-net segmentation network from mammogram images, *J. King Saud Univ. - Comput. Inf. Sci.*, 2019.
- R. Rajalakshmi, Vidhyapriya, Elango, & N. Ramesh (2021) Deeply supervised U-Net for mass segmentation in digital mammograms,” *Int. J. Imaging Syst. Technol.*, vol. 31, no. 1, pp. 59–71, 2021.
- V. K. Singh et al. (2020) Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network, *Expert Syst. Appl.*, vol. 139, no. 112855, p. 112855, 2020.
- W. Liu et al. (2016) SSD: Single Shot MultiBox Detector,” in *Computer Vision – ECCV 2016*, Cham: Springer International Publishing, 2016, pp. 21–37.
- N. Alam and M. J. Islam (2014) Pectoral muscle elimination on mammogram using K-means clustering approach. Online. [Accessed: 29-Jan-2022].
- N. Alam and M. J. Islam (2015) Pectoral muscle elimination on mammogram using K-means clustering approach. Online. [Accessed: 29-Jan-2022].
- J. E. Ball & L. M. Bruce (2007) Digital mammographic computer aided diagnosis (CAD) using adaptive level set segmentation, *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2007, pp. 4973–4978, 2007.
- A. Oliver et al. (2015) Breast density analysis using an automatic density segmentation algorithm, *J. Digit. Imaging*, vol. 28, no. 5, pp. 604–612, 2015.
- S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, & M. T. Rickard (2015) Automatic pectoral muscle segmentation on mediolateral oblique view mammograms, *IEEE Trans. Med. Imaging*, vol. 23, no. 9, pp. 1129–1140, 2004.
- Meritxell Tortajada, Arnau Oliver, Robert Martí, Mariona Vilagran, Sergi Ganau, Lidia tortajada Gimenez, Melcior Sentís & Jordi Freixenet (2012) Adapting Breast Density Classification from Digitized to Full-Field Digital Mammograms, In *Breast Imaging: 11th International Workshop, IWDM 2012*, Philadelphia, PA, USA, July 8-11, 2012. *Proceedings 11*, pp. 561-568, Springer Berlin Heidelberg.
- A. M. Salih & M. Y. Kami (2018) Mammography image segmentation based on fuzzy morphological operations, in *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, 2018, pp. 40–44.
- P. Das & A. Das (2019) A fast and automated segmentation method for detection of masses using folded kernel based fuzzy c-means clustering algorithm, *Appl. Soft Comput.*, vol. 85, no. 105775, p. 105775, 2019.
- N. Dhungel, G. Carneiro, & A. P. Bradley (2015) Tree RE-weighted belief propagation using deep learning potentials for mass segmentation from mammograms, in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015, pp. 760–763.