

Pectoral Muscle Removal in Digital Mammograms Using Region Based Standard Otsu Technique

Jacinta C. Anusionwu¹, Vincent C. Chijindu², Joy N. Eneh³, ThankGod I. Ozue⁴,
Nnabuike Ezukwoke⁵, Mamilus A. Ahaneku⁶, Edward C. Anoliefo^{7,*}, Walter A. Ohagwu⁸

^{1,2,3,4,5,6,7} Department of Electronic & Computer Engineering, University of Nigeria, Nsukka – Nigeria

⁸Department of Electrical & Electronic Engineering, Enugu State University of Science & Technology, Enugu State, – Nigeria

Article Info

Article history:

Received Aug 19, 2022

Revised Apr 6, 2023

Accepted Apr 26, 2023

Keywords:

Digital Mammogram
Medio-Lateral Oblique
Pectoral Muscle
Standard Otsu thresholding
Jaccard Similarity Index
False Negative Rate
False Positive Rate

ABSTRACT

Mammography is usually the first preference of imaging diagnostic modalities used for breast cancer detection in the early stage. Two projections Cranio Caudal (CC) and Medio-Lateral Oblique (MLO), used during digital mammogram acquisition depict different degrees for visualizing breast tissues; the MLO view shows more breast tissue and Pectoral Muscle (PM) area when compared to the CC view. Although the PM is a criterion used for proper positioning, it can result in biased results of mammographic analysis like cancer detection and breast tissue density estimation because the PM area has a similar or even higher intensity than breast tissue and breast lesions if present. This paper proposed a Region Based Standard Otsu thresholding method for the elimination of PM area present in MLO mammograms. The developed algorithm was implemented using 322 digital mammograms from the Mammographic Image Analysis Society (MIAS) database, and the difference between the PM detected and the manually drawn PM region by an expert was evaluated. The results showed an average: Jaccard Similarity Index, False Positive Rate (FPR) and False Negative Rate (FNR) of 93.2%, 3.54%, and 5.68% respectively, and also an acceptable rate of 95.65%.

Copyright © 2023 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Edward C. Anoliefo
Department of Electronic & Computer Engineering,
University of Nigeria, Nsukka,
Enugu State, Nigeria.
Email: edward.anoliefou@unn.edu.ng

1. INTRODUCTION

Cancer is a chronic disease posing serious challenges to health systems globally [1]. The World Health Organization in 2020 recorded a mortality rate of about 10 million globally relating to cancer, and cases of approximately 70% occurred in low and middle-income countries [2]. In Nigeria, about 102,000 thousand new cases of different cancer occur annually, with high number of recorded cases of breast cancer [3], [4]. Breast cancer results from abnormal changes in the genes responsible for regulating cell growth in the breast tissues. It is the most common of cancers in the feminine gender because of the structure, orientation, and presence of developed breast muscles. Breast cancer can begin from any breast part; giving rise to the different progressions of breast cancer that depends on specific breast tissue affected and the degree of spread [5], [6].

The use of medical imaging machines has been an effective method used to reveal the inner structure of the human body mainly for clinical analysis, making diagnosis and treatment less cumbersome. Also, many body parts are not torn open with assumptions in surgery [7], [8]. Imaging tools like Mammography, MRI, Ultrasound, and Thermography [5], [9] are utilized for breast screening for the detection of breast cancer. However, results associated with the medical images acquired depended on the expertise and visual interpretation of the radiologist. Computer-aided system designed for medical image assessment depends on the type of medical examination, part of the body affected, disease type, and desired findings from the image. The

computer-aided system helps resolve shortcomings associated with the manual interpretation of medical images [10], [11]; as the detection process is increased and the error due to human fatigue is drastically reduced. Computer-aided systems are categorized into two schemes Computer-Aided Diagnosis (CADx) and Computer-Aided Detection (CADE). The CADE focused on identifying potential abnormalities in the image and outlining suspicious Regions of Interest (ROI) for further examination by the radiologists. Although detailed characteristics of the tumor detected are unknown, the radiologist's attention is drawn to a tumor that may be unnoticed. CADx is targeted on the diagnosis, and classification of suspected regions as normal or malignant or benign [12], [13].

Mammography is usually the first option in detecting breast lesions because of its high sensitivity in detecting cancer at an early stage. Mammography uses a low dose of X-ray 0.4mSv (millisieverts or mSv is a unit of measure for radiation dose) [5], [9] to produce a mammogram (breast image) usually taken in two views which are: Cranio Caudal (CC) and Medio-Lateral Oblique (MLO). These views allow a degree of visualizing dense breasts to an extent, and the MLO view depicts more breast tissue and Pectoral Muscle (PM) area when compared with the CC view [14], [15]. Thus, mammograms in MLO projection are used for this study.

Although, the presence of the PM is a criterion used to show proper positioning during image acquisition [15] and also, as an index to compare the symmetry between the left and right breast images [16]. The PM contains no relevant information during automatic intensity-based mammographic analysis like breast tissue density estimation and identifying cancer lesions [17]. Thus, the removal of PM is an essential preprocessing step required for most CADx systems to minimize biased results and false positive cases [18], [19]; because the intensity of the pectoral muscle area is similar or even higher in some cases than breast tissue and lesions if present. Recently, different researchers proposed and developed various automated and semi-automated methods for PM removal in MLO view mammogram images.

Fadhil and Dawood [19] developed an algorithm based on Split Orientation Local Thresholding (SLOTH) to remove the PM area. The method was implemented by splitting the pre-processed mammogram image into four quadrants and the upper left or upper right portion with the PM region selected based on image orientation. A set threshold of 150 was used to perform binarization on the selected quadrant, thereby eliminating the PM region. The image without PM was gotten by grouping the selected quadrant without PM with the rest of the quadrants split earlier. The algorithm was implemented using C# and tested using images 100, 110, and 322 from the MIAS database with an accuracy of 98%, 90.9%, and 93.2% obtained respectively. There were two limitations in this approach. First the algorithm applied on the selected quadrant implied that some images with some portions of the PM seen on other quadrants not selected are missed. Secondly, the manually set threshold of 150 used for all images considered in the experiment might not be optimal for some images.

Vagssa et al [20] assumed the PM border to be a straight line and developed a Hough transform algorithm for the deletion of the PM area in a mammogram. In this work, left image orientation was used to help simplify the algorithm, such that a right mammogram was flipped to the left. Also, the upper left quarter of the mammogram was used as a region of interest to define the Hough mask obtained using the Canny edge detection method and a Hough transformation equation defined in [21]. Although the deletion accuracy rate obtained was 93.8%, the shape of the PM boundary for some mammograms is not always a straight line.

Yoon et al [16] proposed a nonlinear Random Sample Consensus Algorithm (RANSAC) to remove PM. In the PM segmentation, an oblique kernel was used to show the outline of the enhanced image and binarization was done using Otsu thresholding to detect the edges of the outline. Hough transform was used to connect lines having similar angles in the range of 100 to 170 for left MLO images and 280 to 350 for right MLO images, and the line with the longest length was selected as the corresponding PM outline which was further interpolated using the RANSAC algorithm to obtain an optimal PM outline. They obtained an acceptable rate of 92.2%, an average False Negative (FN) of 5.68% and average False Positive (FP) of 4.51%.

Makandar and Halalli [13] proposed a technique based on thresholding and a modified region-growing technique to remove pectoral muscle. The seed point for the region-growing algorithm was selected automatically by considering the orientation of the mammogram. 100 images from the MIAS database were used to implement the proposed method and an accuracy of 97% was obtained. Also, the wiener, adaptive min-max, and median filter were compared and the PSNR of the Wiener filter was high and the RMSE and image quality index (IQI) were reduced. Thus, for preprocessing mammograms. Wiener filter and CLAHE were used to enhance image quality.

Sreedevi and Sherly [22] developed an algorithm for segmenting and eliminating pectoral muscle. In this work, global thresholding is used to identify pectoral muscles, the boundaries are detected using canny edge, and the regions are extracted using CCL. An accuracy of 90.06% was obtained when implemented on 161 images from the MIAS database.

In summary, the bounding box approach based on a defined region within the mammogram image has been widely used in literature as a criterion when developing PM removal algorithms. The main constraint of the works reviewed was the susceptibility of the methods used to eliminate PM from the defined region of the

bounding box and the complexity of the method used in defining the PM boundary. In [19], [20], they tried to improve the bounding box by dividing preprocessed mammograms into four quadrants and considering only the upper quadrants in the vertical edge of the breast profile. Some images however, have a significant portion of PM present in the lower quadrant of the vertical edge of the breast region which was not accounted for. Also, the assumption of the PM boundary as a straight line as seen in [20] is not always the case. In this study, a new segmentation method called Region Based Standard Otsu thresholding is developed for the elimination of PM, and it considered desired regions based on vertical distance since the height of the PM is unknown and unique to different mammograms. Also from research, the choice of threshold value used can be obtained automatically, semi-automatically, and even manually. However, for this research the threshold value used was obtained automatically using Otsu thresholding.

2. RESEARCH METHOD AND MATERIALS

The methods used in this work combine two stages of digital image processing techniques: image enhancement and image segmentation. A detailed description of the methods employed in these stages are described in the sub sections below. Figure 2.1 shows the flowchart for PM removal algorithm used to efficiently implement this work.

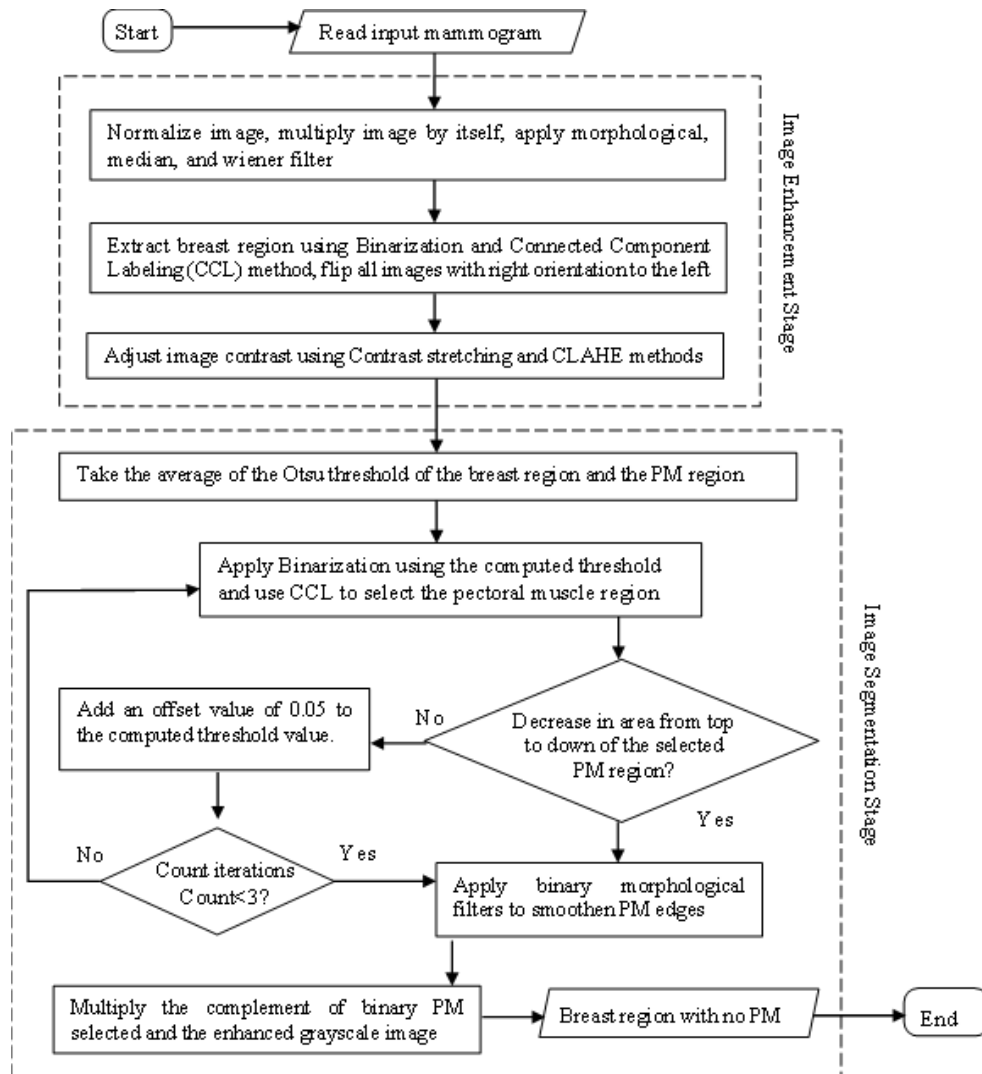


Figure 2.1. Flowchart of PM removal algorithm

2.1. Input Mammogram Image Dataset

Breast images used in this work were gotten from the most widely used MIAS mammogram database [23]. The MIAS database is an online digital mammography data gotten from an organization of UK research

group. It consists of 161 patient's mammogram pairs in MLO views already in grayscale and has the Portable Gray Map (PGM) format with a size of 1024 X 1024.

2.2. Image Enhancement Stage

The enhancement technique has various approaches, which are goal-driven to modify images so that the resulting image is more suitable than the former for specific applications. Digital Mammograms are susceptible to noise and naturally have poor contrast between the densities of different breast tissues and cancer tumors if present. As seen in Figure 2.2, a typical mammogram has various information, of which not all are relevant. This work combined different enhancement techniques to improve image contrast and remove noise, tape artifacts, and labels if present.

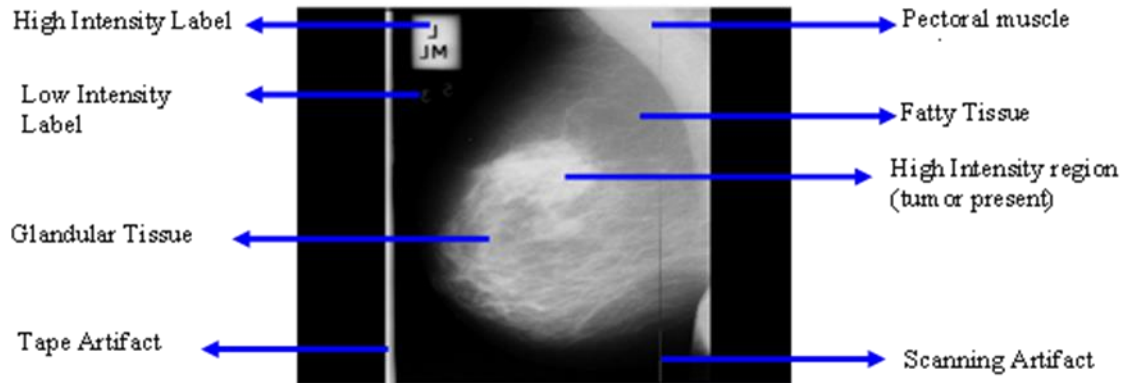


Figure 2.2. The different component in a typical MIAS dataset image (mdb 111)

2.2.1. Thresholding

Thresholding is a segmentation technique that categorizes image pixel values; using a set threshold [7], [24]. This method is described mathematically in Equation 2.1 below.

$$G(x, y) = \begin{cases} \text{foreground} & \text{if } F(x, y) \geq T \\ \text{background} & \text{if } F(x, y) < T \end{cases} \quad (2.1)$$

$F(x, y)$ is the input image, T is the set threshold or range for threshold values, and $G(x, y)$ is the output image.

Binarization thresholding segments a grayscale image into a binary image such that the background is set to '0' (black) and the foreground is set to '1' (white) following a defined threshold [25]. Double thresholding is a modification of binarization such that the threshold criterion is a given range of threshold values.

Multi-thresholding segments object in a grayscale image into a limited number of gray levels defined by more than one threshold value or range of values.

The main parameter for optimal thresholding segmentation is the threshold value used. A threshold value can be obtained manually by random selection, or automatically from an image histogram, or using different statistical methods like mode, mean, variance, etc. There are different approaches in literature for choosing an optimal threshold value.

2.2.1.1. Otsu Thresholding

This method considers the image pixel value Probability Density Function (PDF) as a bimodal distribution, such that the background and foreground are defined clearly in the image histogram. With this assumption, it automatically finds the optimal value for global thresholding [7]; through an extensive search of possible thresholding values by calculating their corresponding weighted between-class variance.

The optimal threshold value is the threshold with the minimum variance within two classes or the maximum weighted between class variance. The intra-class variance or within-class variance can be defined using the equation of weighted variances of each cluster as in Equation 2.2. Also, the between-class variance is calculated using Equation 2.6 or 2.7

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (2.2)$$

Where the weights q_i are the probabilities for each class to be estimated using Equation 2.3.

$$q_1(t) = \sum_{i=1}^t P(i) \quad \text{and} \quad q_2(t) = \sum_{i=t+1}^I P(i) \quad (2.3)$$

And the class mean is given by Equation 2.4

$$\mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad \text{and} \quad \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)} \quad (2.4)$$

The class variance is also given by Equation 2.5

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad \text{and} \quad \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)} \quad (2.5)$$

$$\sigma_b^2(t) = q_1(t)[\mu_1(t) - \mu(t)]^2 + q_2(t)[\mu_2(t) - \mu(t)]^2 \quad (2.6)$$

The variable $\mu(t)$ is the mean of the image histogram. Also, the between-class variance $\sigma_b^2(t)$ can be further simplified in Equation 2.7

$$\sigma_b^2(t) = q_1(t)q_2(t)[\mu_1(t) - \mu_2(t)]^2 \quad (2.7)$$

2.3. Pectoral Muscle Removal

The effective removal of PM is a function of the breast density type, quality of image contrast, size of the breast region, and PM area. The following characteristics were considered when developing the PM removal algorithm:

- The PM is usually at the top corner of the vertical edge of the breast region.
- The PM always intersects the vertical edge of the breast region
- The PM tissue has a higher intensity than the surrounding tissue.
- There is a gradual decrease in the PM width from top to bottom.

The threshold used was obtained by taking the average of the Standard Otsu for the breast region defined by points ABDE and the standard Otsu threshold for the Pectoral Muscle region defined by points ABCF as in Figure 2.3. And this average value obtained was used as the set threshold to perform binarization and CCL to eliminate the pectoral muscle region.

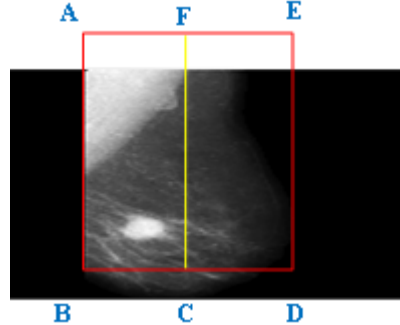


Figure 2.3. Mammogram image showing areas used to obtain threshold value

For some MLO mammograms, the PM boundary is almost identical with the surrounding tissues, especially the lower portion of the PM. Although the optimal threshold obtained does not completely separate the PM for the breast tissue, an iterative threshold selection was used to optimize the segmentation process for such images. The algorithm applied in PM removal is illustrated in Table 2.1.

Table 2.1. Pseudocode for PM removal

1.	Img1 = Grayscale mammogram without noise, labels, and contrast enhanced
2.	Th = Possible values of optimal threshold obtained using multi Otsu thresholding multithresh MATLAB function with gray levels of 5
3.	A1 = Breast region rectangular border ABDE as the area obtained when the minimum value of Th is used to perform binarization
4.	A2 = PM rectangular border ABCF as the area obtained when the median of Th is used to perform binarization
5.	T1= Compute the Otsu threshold for region A1 using the graythresh MATLAB function
6.	T2= Compute the Otsu threshold for region A2 also using the graythresh MATLAB function
7.	Thr=1/2(T1+T2)

-
8. $Img2$ = Binary image of $Img1$ using the **imbinarize** function with the computed value of Thr
 9. $Img3$ = Apply binary morphological filters on $Img2$ using the **imfill** function
 10. Extract PM region
 - i. $Cc1$ = Find the connected objects in $Img3$ using **bwconncomp** function
 - ii. $L1$ = Create label matrix of $Cc1$ using the **labelmatrix** function
 - iii. $Mv1$ = Minimum column coordinate value for pixel value >0 in $Cc1$
 - iv. Mn = Search the $Mv1+10$ columns in the matrix $L1$ for mode from the label values
 - v. $BwPimg$ = select connected object with label Mn as the Pectoral Muscle
 11. Check for optimal PM selection
 - i. $W = 0$
 - ii. Store the maximum column coordinate value for pixel >0 in vector D for $\{(W+1) \times 20\}$ consecutive rows in the $BwPimg$
 - iii. W = Number of elements in vector D
 - iv. For $W < 3$, repeat step 11ii.
End
 - v. Compare the consecutive elements of in vector D for a decrease
 - vi. $Y=0$
 - vii. For step v = False and $Y < 2$
repeat step 8 and increase Thr by 0.05
End
 12. $Pimg=1-BwPimg$
 13. $Fimg$ = Product of the binary image $Pimg$ and $Img1$ using the **immultiply** function
-

2.4. Method of Evaluating the Result of Proposed Algorithm

The area-based metric is a common quantitative method used in literature to evaluate the quality and goodness of segmented regions by comparing the area of the region segmented by an algorithm with the actual area as determined manually by an expert [26]. Three important parameters True Positive area (TP), False Negative area (FN), and False Positive area (FP) are useful in defining some segmentation metrics indicators.



Figure 2.1. Illustration of True Positive, False Positive and False Negative

From Figure 2.4, the TP area shows the segmented region that matches perfectly with the ground truth region determined by experts. For the FN area, the expert reference region is not present in the region segmented by the algorithm. Also, the FP area shows that some region segmented lies outside the expert manual contour [21], [26]. The different metrics used to evaluate the quality and goodness of segmented regions are as follows.

- i. **False Positive Rate (FPR):** is expressed mathematically in Equation 2.8 as the ratio of the extra pixels to the total number of pixels obtained.

$$FPR = \frac{FP}{TP+FP} \quad (2.8)$$

- ii. **False Negative Rate (FNR)** is expressed mathematically as shown in Equation 2.9. It is the ratio of the missing pixels to the total number of pixels obtained

$$FNR = \frac{FN}{TP+FN} \quad (2.9)$$

The FPR and FNR are used to evaluate the missed segmentation rate. Thus, the smaller the FPR and FNR, the better the segmentation algorithm.

- iii. **Jaccard Similarity Coefficient:** It measures the similarity and diversity of two sample images: the ground truth segmentation by experts and the predicted segmentation, by computing the intersection or region overlap between them divided by their union [18].

$$J(I_o, I_s) = \frac{I_o \cap I_s}{I_o \cup I_s} \quad (2.10)$$

Equation 2.10 I_o is the region segmented manually by experts, and I_s is the region segmented by a proposed segmentation method. The value of $J(I_o, I_s)$ lies between 0 and 1. For a perfect match of the two images, the Jaccard similarity coefficient is 1. The Jaccard index can also be expressed in terms of TP, FP, and FN as shown in Equation 2.11.

$$J(I_o, I_s) = \frac{TP}{TP+FP+FN} \quad (2.11)$$

The agreement index based on the Jaccard Similarity coefficient is shown in Table 2.2.

Table 2.2. Agreement Index based on Jaccard values [18]

Range of Jaccard Values	Relationship
≤ 0	No agreement
0.001-0.2	None to slight agreement
0.21-0.4	Fair agreement
0.41-0.6	Moderate agreement
0.61-0.8	Substantial agreement
0.81-0.9	Good agreement
0.91-1.0	Perfect agreement

3. RESULTS AND DISCUSSION

3.1. Experimental Set Up

The algorithms developed for this research were actualized using the MATLAB 2018a software version on a computer with a 2.16Hz-Intel Pentium processor and 4GB of RAM. And the result of the different steps obtained is shown below.

3.2 Results

The 16 image samples in Tables 3.1 were used to illustrate the experimental results obtained using the developed algorithm. A total of 322 images from the MIAS database were also analyzed. The segmented PM results obtained were compared with manually drawn contours by an expert radiologist in mammography. The quantitative metric in terms of False Positive Rate (FPR), False Negative Rate (FNR), and Jaccard index defined in Equation 2.8, 2.9, and 2.11 respectively were computed. Table 3.2 shows the results for each of the 16 images, also Table 3.3 gives a summary of the results for the 322 images in the MIAS dataset.

3.3 Discussion

In Table 3.1, the first column shows the image identifier as indexed in the MIAS database. From the third column of Table 3.1, although the original image contrast was enhanced, additive noise was also introduced. The fourth column shows that noise was removed as well as other trivial information present in the image: additive noise present in all image samples, scanning artifacts in 'mbd 104 and 117', low-intensity labels in 'mdb 120, 218, 58, and 28' high and intensity labels in 'mdb 117, 75, 104, 206, 170, 58, 05, 184, 195 and 193' were eliminated. For this work, the left orientation was used, so images in the right orientation like 'mdb 117, 75, 25, 213, 05, 195 and 193' were flipped to the left as seen in the fourth column. The fifth column shows a further enhanced image contrast and the last column show the resulting grayscale image after the removal of PM.

From Table 3.2, an average Jaccard index of 98.69% was obtained, this result shows that the concordance of the PM region segmented manually by an expert and by the proposed method was excellent. From Table 3.2 the FNR and FPR fall below 5%, and the average FNR and FPR are 1.24% and 1.03% respectively. These lower values show that the proposed segmentation algorithm was good and effective. Also, visual inspection of the results presented in Table 3.1 are very efficient.

The overall performance of the algorithm for PM removal can be grouped into two categories: acceptable and unacceptable. Images are said to be acceptable if the PM is correctly removed or a majority of the PM area is removed. Also, an image is unacceptable if most or all of the pectoral muscle is not removed.

Acceptable cases for our proposed method have a Jaccard similarity index greater than 81%, FNR less than 12%, and FPR less than 12%. Thus for the MIAS dataset, our proposed method produced an acceptable rate of 95.65%. Also, the average Jaccard similarity index of 93.2%, an average FPR of 3.54%, and an average FNR 5.68% were obtained.

Table 3.1. Results of the Different Steps of the Algorithm Developed

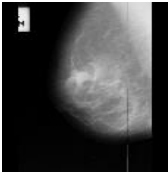
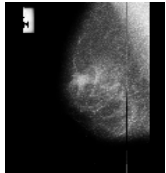
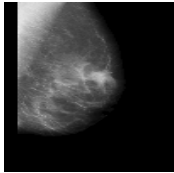
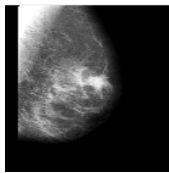
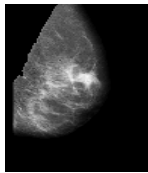
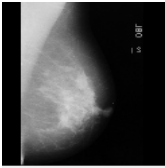
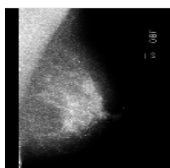
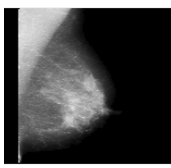
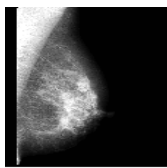
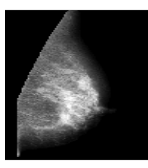
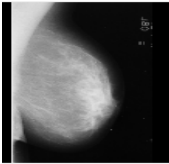
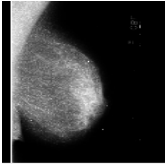
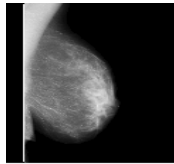
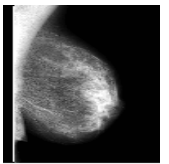
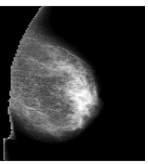
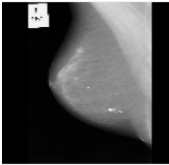
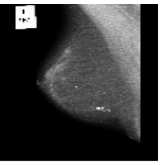
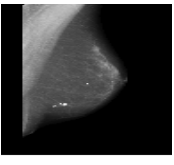
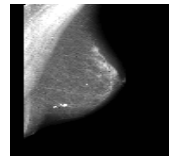
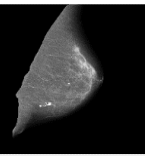
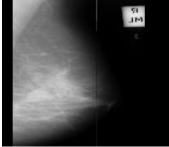
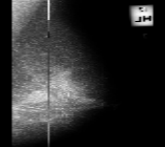
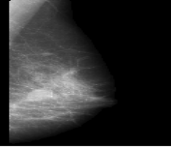
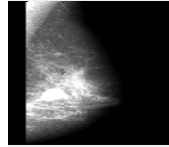
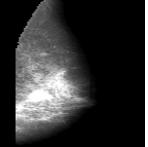
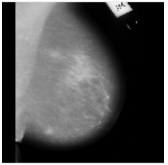
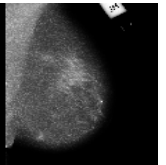
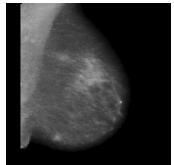
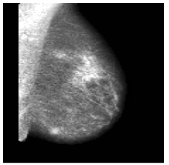
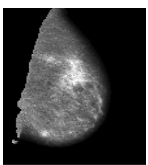
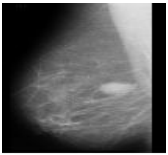
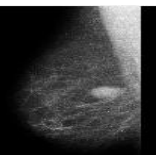
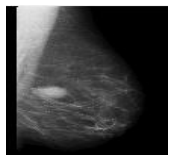
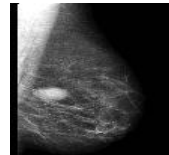
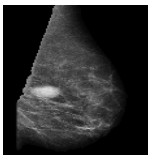
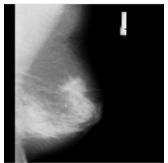
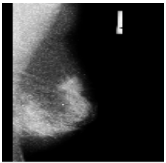
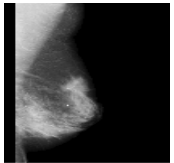
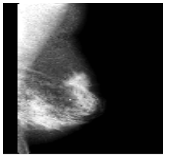
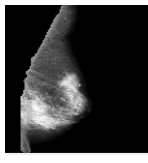
Image No	Original image	Enhanced image	Noise, labels and other artifacts removed	Further enhanced image	Pectoral Muscle removed
mdb 117					
mdb 120					
mdb 218					
mdb 75					
mdb 104					
mdb 206					
mdb 25					
mdb 170					

Table 3.1. (Continued)

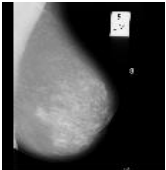
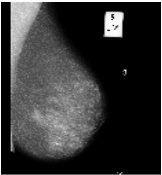
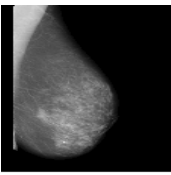
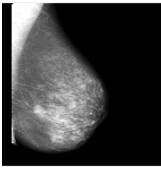
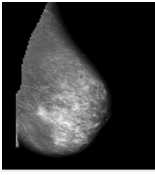
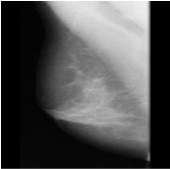
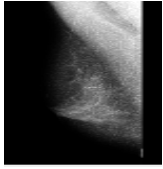
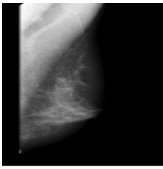
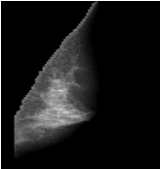
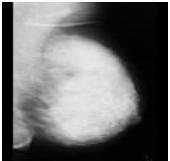
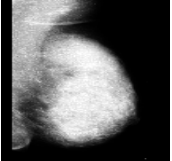
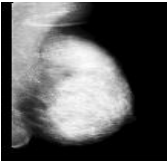
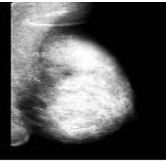
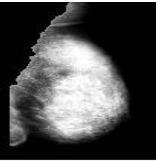
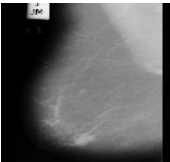
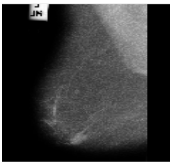
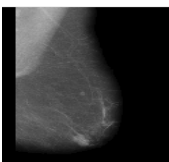
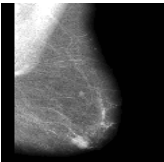
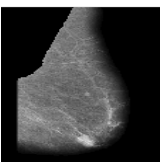
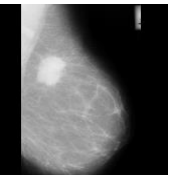
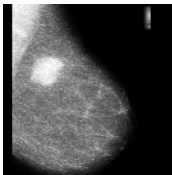
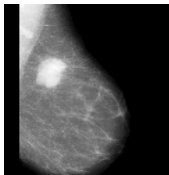
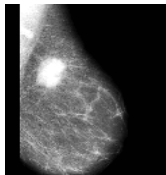
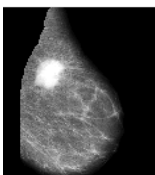
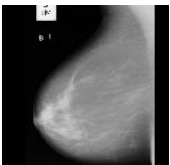
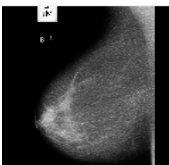
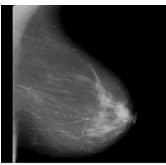
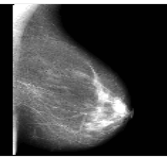
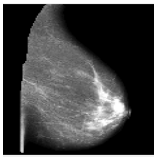
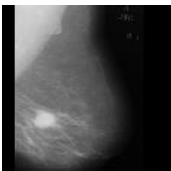
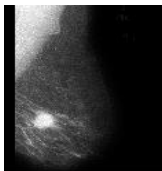
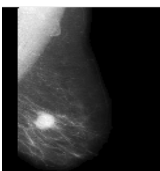
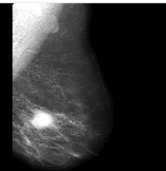
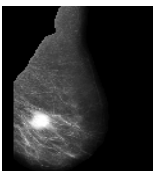
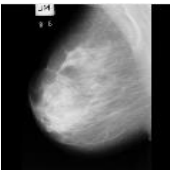
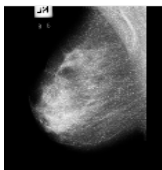
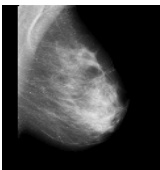
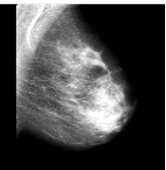
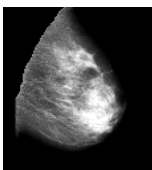
Image No	Original image	Enhanced image	Noise, labels and other artifacts removed	Further enhanced image	Pectoral Muscle removed
mdb 58					
mdb 213					
mdb 240					
mdb 05					
mdb 184					
mdb 195					
mdb 28					
mdb 193					

Table 3.2 Performance Evaluation of PM for 16 Sample Images

S/No.	Image No	FNR	FPR	Jaccard Index
1	mdb 117	0.0098	0.0061	0.9862
2	mdb 120	0.0012	0.0007	0.9988
3	mdb 218	0.0039	0.0404	0.9851
4	mdb 75	0.0011	0.0415	0.9896
5	mdb 104	0.0032	0.0006	0.9965
6	mdb 206	0.0009	0.0119	0.9973
7	mdb 25	0.0418	0.0013	0.9687
8	mdb 170	0.0498	0.0037	0.9465
9	mdb 58	0.0024	0.0067	0.9916
10	mdb 213	0.0327	0.0018	0.9789
11	mdb 240	0.0010	0.0004	0.9989
12	mdb 05	0.0275	0.0035	0.9859
13	mdb 184	0.0014	0.0429	0.9804
14	mdb 195	0.0003	0.0013	0.9987
15	mdb 28	0.0005	0.0010	0.9985
16	mdb 193	0.0215	0.0004	0.9890
Average		0.0124	0.0103	0.9869

Table 3.3 Summary of Performance Evaluation of PM for 322 Sample Images

Range of Jaccard Values	No of images	Range of FNR values	Range of FPR values
≤ 0	3 (0.93%)	$0.8 < \text{FNR} \leq 1$	$0.8 < \text{FPR} \leq 1$
0.001-0.2	2 (0.62%)	$0.6 < \text{FNR} \leq 0.8$	$0.6 < \text{FPR} \leq 0.8$
0.21-0.4	--	$0.5 \leq \text{FNR} < 0.6$	$0.5 < \text{FPR} \leq 0.6$
0.41-0.6	6 (1.86%)	$0.25 \leq \text{FNR} < 0.5$	$0.25 < \text{FPR} \leq 0.5$
0.61-0.8	3 (0.93%)	$0.12 \leq \text{FNR} < 0.25$	$0.12 < \text{FPR} \leq 0.25$
0.81-0.9	34 (10.56%)	$0.06 \leq \text{FNR} < 0.12$	$0.06 < \text{FPR} \leq 0.12$
0.91-1.0	274 (85.09%)	$0 \leq \text{FNR} < 0.06$	$0 < \text{FPR} \leq 0.06$

Table 3.4. Comparison of Proposed Algorithm for PM Removal with Previous Works

Author	Method	No. of images	Accept-able	Unaccept-able
Sreedevi and Sherly [22]	Global thresholding	161	90.06%	9.94%
Makandar and Halali [27]	Region growing	100	97%	3%
Yoon et al [16]	Random Sample Consensus Algorithm (RANSAC)	322	92.2%	7.8%
Fadhil and Dawood [19]	Split Orinetation Local Threshoding (SOLTH)	322	93.2%	6.8%
Vagssa et al [20]	Hough transform	322	93.8%	6.2%
Proposed Algorithm	Region Based Standard Otsu	322	95.65%	4.35%

The proposed algorithm for PM removal used in this work achieved better results than previous works as seen in Table 3.4, due to the threshold selection that considered some regions in the mammogram image and took the average as against using the entire image threshold or manual selection of threshold value. Also, the result at the image enhancement stage contributed greatly to the success of the segmentation algorithm.

4. CONCLUSION

In this paper, a Region-Based Standard Otsu thresholding for the removal of PM has been presented. The proposed algorithm was accurate for variations in PM Size, boundary curvature, and tissue density. Also, the results obtained show that the proposed algorithm achieved a higher performance when compared with previous works.

The performance of the algorithm was evaluated using the Jaccard Similarity index, average FNR, and average FPR. Also, the algorithm was compared with other related works reviewed based on the acceptable rate, and it performed better for the wide range of data used. In the future, the image obtained after removing PM can be used for further image analysis using other image processing steps like CADx system design for breast cancer lesions detection in mammograms. Also, other segmentation methods can be applied to the region considered for this work.

ACKNOWLEDGEMENT

The authors will like to thank Dr. Ezigbo Ikemefuna Mammography Specialist and Medical Radiographer, at the Alex Ekwueme Teaching Hospital Abakaliki, Ebonyi State, for his support in the expert manual segmentation and detection of the Pectoral Muscle region of the mammogram images used in this work.

REFERENCES

- [1] P. C. Pearlman *et al.*, "The National Institutes of Health Affordable Cancer Technologies Program : Improving Access to Resource-Appropriate Technologies for Cancer Detection , Diagnosis , Monitoring , and Treatment in Low- and Middle-Income Countries," *IEEE J. Transl. Eng. Heal. Med.*, vol. 4, no. February, 2016.
- [2] World Health Organization, "Cancer - Key Facts." [Online]. Available: <https://www.who.int/en/news-room/fact-sheets/detail/cancer>. [Accessed: 20-Oct-2020].
- [3] International Cancer Control Partnership, "Nigeria National Cancer Control Plan 2018 – 2022," 2020. [Online]. Available: https://www.iccp-portal.org/system/files/plans/NCCP_Final %5B1%5D.pdf. [Accessed: 20-Aug-2020].
- [4] O. Erhabor *et al.*, *Breast Cancer in Nigeria: Diagnosis, Management and Challenges*. UK: Author House, 2016.
- [5] American Cancer Society, "Breast Cancer Early Detection and Diagnosis." [Online]. Available: <https://www.cancer.org/cancer/breast-cancer.html>. [Accessed: 20-Aug-2020].
- [6] R. A. Al-naggar, *Principles and Practice of Cancer Prevention and Control*. OMICS Group ebook, 2014.
- [7] A. Norouzi *et al.*, "Medical Image Segmentation Methods, Algorithms, and Applications," *Inst. Electron. Telecommun. Eng. Tech. Rev.*, vol. 31, no. 3, pp. 199–213, 2014.
- [8] P. Chinmayi, L. Ailandeewari, and M. Prabukumar, "Survey of Image Processing Techniques in Medical Image Analysis: Challenges and Methodologies," in *Proceedings of the Eighth International Conference on Soft Computing and Pattern Recognition*, 2018, pp. 460–471.
- [9] S. J. Sushma and S. C. Prasanna Kumar, "Advancement in Research Techniques on Medical Imaging Processing for Breast Cancer Detection," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 2, pp. 717–724, 2016.
- [10] M. Firmino, G. Angelo, H. Morais, M. R. Dantas, and R. Valentim, "Computer - Aided Detection (CADe) and Diagnosis (CADx) System for Lung Cancer with Likelihood of Malignancy," *BioMed Cent. J.*, vol. 15, no. 2, pp. 1–18, 2016.
- [11] M. A. Berbar, "Hybrid Methods for Feature Extraction for Breast Masses Classification," *Egypt. Informatics J.*, vol. 19, no. 1, pp. 63–73, 2018.
- [12] R. Guzman *et al.*, "Digital Image Processing Technique for Breast Cancer Detection," *Int. J. Thermophys.*, pp. 1519–1531, 2013.
- [13] B. Halalli and A. Makandar, "Computer Aided Diagnosis - Medical Image Analysis Techniques," in *Intech open science*, 2018, pp. 85–109.
- [14] F. Elzahra, A. Hateem, M. Mohammad, and M. Tarique, "Fourier Transform Based Early Detection of Breast Cancer by Mammogram Image Processing," *J. Biomed. Eng. Med. Imaging*, vol. 2, no. 4, pp. 17–32, 2015.
- [15] H. Ture and T. Kayikcioglu, "Accurate Detection of Distorted Pectoral Muscle in Mammograms Using Specific Patterned Isocontours," *IEEE Access*, vol. 8, 2020.
- [16] W. B. Yoon, J. E. Oh, E. Y. Chae, H. H. Kim, S. Y. Lee, and K. G. Kim, "Automatic Detection of Pectoral Muscle Region for Computer-Aided Diagnosis Using MIAS Mammograms," vol. 2016, 2016.
- [17] S. Saxena, Y. Singh, and B. Agarwal, "Boundary Detection to Segment the Pectoral Muscle from Digital Mammograms Images," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 3, pp. 1593–1603, 2020.
- [18] S. Gardezi, A. Faouzi, I. Faye, K. Nidal, and M. Eltoukhy, "Segmentation of Pectoral Muscle Using the Adaptive Gamma Corrections," *Multimed. Tools Appl.*, vol. 75, no. 24, 2017.
- [19] S. S. Fadhil, F. Abed, and A. Dawood, "Automatic Pectoral Muscles Detection and Removal in Mammogram Images," *Iraqi J. Sci.*, vol. 62, no. 2, pp. 676–688, 2021.
- [20] P. Vagssa, N. M. Doudou, T. Jolivo, O. Videme, and D. T. Kolyang, "Pectoral Muscle Deletion on a Mammogram to Aid in The Early Diagnosis of Breast Cancer," *Int. J. Eng. Sci. Technol.*, vol. 12, no. 3, pp. 57–65, 2020.
- [21] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère, "Automatic Identification of the Pectoral Muscle in Mammograms," vol. 23, no. 2, pp. 232–245, 2004.
- [22] S. Sreedevi and E. Sherly, "A Novel Approach for Removal of Pectoral Muscles in Digital Mammogram," *Procedia - Procedia Comput. Sci.*, vol. 46, no. Ict 2014, pp. 1724–1731, 2015.
- [23] J. Suckling and *et al.*, "The Mammographic Image Analysis Society Digital Mammogram Database," *International Congress Series 1069*, 1994. [Online]. Available: peipa.essex.ac.uk/info/mias.html. [Accessed: 29-Jun-2019].
- [24] O. Marques, *Practical Image and Video Processing using MATLAB*. Hoboken, New Jersey: John Wiley and Sons, Inc, 2011.
- [25] A. M. Omer and M. Elfadil, "Preprocessing of Digital Mammogram Image Based on Otsu ' s Threshold," *Am. Sci. Res. J. Eng. Technol. Sci.*, vol. 37, no. 1, pp. 220–229, 2017.
- [26] V. C. Chijindu, U. Chidiebele, M. Ahaneku, and E. C. Anoliefo, "DETECTION OF PROSTATE CANCER USING RADIAL / AXIAL SCANNING OF 2D Detection of Prostate Cancer Using Radial / Axial Scanning of 2D Trans-rectal Ultrasound Images," *Bull. Electr. Eng. Infomatics*, vol. 7, no. 2, pp. 222–229, 2018.

- [27] A. Makandar, "Pre-processing of Mammography Image for Early Detection of Breast Cancer," *Int. J. Comput. Appl.*, vol. 144, no. 3, pp. 11–15, 2016.

BIOGRAPHY OF AUTHORS



Jacinta C. Anusionwu received her B.Eng degree from the University of Port Harcourt in 2016. She obtained M.Eng. Degree in Electronic & Computer Engineering from University of Nigeria Nsukka. Her main research interests includes: data analysis and machine learning.

E-mail: jacinta.anusionwu.pg00412@unn.edu.ng



Vincent C. Chijindu is an Associate Professor at the Department of Electronic & Computer Engineering, University of Nigeria, Nsukka, Nigeria. He obtained B.Eng and M.Eng in Electronic and Computer Engineering from Anambra State University of Technology Enugu Nigeria and Ph.D. in Computer and Control Engineering from Nnamdi Azikiwe University Awka Nigeria. His current research interests include artificial intelligence in medical diagnosis, digital systems design, signal/image processing, renewable energy materials & systems and wireless sensor networks. E-mail: vincent.chijindu@unn.edu.ng



Joy N. Eneh is currently a Senior Lecturer at the Department of Electronic Engineering of University of Nigeria Nsukka (UNN). She holds a Ph.D in Computer and Control Engineering Research Interest Areas Includes: Control System Networks, Robotics and Artificial Intelligence, Multi- Agent Systems, Optimal Control. E-mail: nnenna.eneh@unn.edu.ng



ThankGod I. Ozue obtained B.Eng in Electrical Electronic & Telecommunication Engineering and M.Eng in Electronic and Computer Engineering from Nnamdi Azikiwe University Awka Nigeria. He is lecturer and researcher within University of Nigeria and has been collaborating with indigenous Tech. start-ups. His current research interests include cyber physical systems, microgrids, iot embedded control systems, and cyber security. E-mail: izuchukwu.ozue@unn.edu.ng



Nnabuike Ezukwoke is a Lecturer in the Department of Electronic Engineering University of Nigeria, Nsukka where he obtained both B.Eng and M.Eng in 2010 and 2016 respectively. He obtained Ph.D in Electronic & Computer Engineering in 2022. His research interest is in Renewable Energy and Advanced Energy Materials - Advanced thin films technology and materials. E-mail: nnabuike.ezukwoke@unn.edu.ng



Mamilus A. Ahaneku obtained his B.Eng. Electrical/Electronic Engineering and MSc Communications Engineering from Federal University of Technology, Owerri, Nigeria in 1994 and 2000, respectively. He holds PhD in Communications Engineering from University of Nigeria, Nsukka. He is a Senior Lecturer at the Department of Electronic Engineering, University of Nigeria, Nsukka. His research interests include Wireless Communication, Microwave Systems and Radio Frequency Design. E-mail: mamilus.ahaneku@unn.edu.ng or ahamac2004@yahoo.co.uk



Edward C. Anoliefo received his B. A (Philosophy) in 1993 and B. Th (Theology) in 1998. He also holds a B.Eng., M.Eng. and Ph.D. in Electronic Engineering from University of Nigeria, Nsukka. He is a Senior Lecturer at the Department of Electronic & Computer Engineering, University of Nigeria Nsukka. E-mail: edward.anoliefo@unn.edu.ng



Walter A. Ohagwu is a Lecturer at the Department of Electrical & Electronic Engineering, Enugu State University of Science & Technology, Enugu State. He holds both Bachelor and Masters in Engineering from University of Nigeria Nsukka. His research interest include machine learning, artificial intelligence in medical diagnosis, digital systems design, renewable energy systems and control engineering. E-mail: walterohagwu@gmail.com