MLO Mammogram Pectoral Muscle Masking with Adaptive MSER

Ch. Sarada

Sr. Asst. Professor, CVR College of Engineering/CSE Department, Hyderabad, India E-mail: Chintala.sarada@cvr.ac.in

Abstract: **Breast cancer claims the lives of many people every year. Breast cancer diagnosis is a difficult process that requires competent radiologists. Manual detection of breast cancer disease takes a large amount of time, as does manual treatment of disease. As a result, automated detection is required, which aids in early treatment and in some circumstances, saves lives. These technical advancements are beneficial for early treatment due to resource availability and computing capability. A mammogram is a technique for detecting breast cancer masses early. To identify masses from MLO mammograms, several image processing-based computer-aided diagnostic techniques have been developed. In computer aided diagnosis systems, the presence of Pectoral Muscle in MLO Mammograms has a considerable detrimental influence on mass detection from MLO Mammograms. Masking the Pectoral muscle improves mass detection from an MLO mammogram. Locating the Pectoral muscle is tough since the intensity of this tissue is equivalent to that of a malignancy. The primary goal of this research is to develop a Most Stable Extremal Region (MSER) based method to locate and mask Pectoral muscle from the Mediolateral Oblique Mammogram. The empirical analysis suggests that the proposed novel procedure is straightforward and gives promising results in locating and masking Pectoral muscle. The suggested technique enhances accuracy by 96.27% compared to 95% for state-of-the-art solutions. Python and MATLAB are used to create the new system.**

Index Terms: **MLO Mammogram, Pectoral muscle, Most Stable Extremal Region (MSER), Image processing.**

I. INTRODUCTION

Depending on the type of breast tissue being discussed, a woman's breasts can be fatty, fibrous, or glandular. The lobules and ducts of the breast are examples of glandular tissue. From the skin to the chest wall, there is a layer of connective tissue called fibrous tissue. Connective tissue includes the fibrous tissue that makes up ligaments and scar tissue. The transition from glandular to connective tissue is facilitated by the deposition of fatty tissue. Fibroglandular tissue [1] is the term used by medical professionals to describe any type of tissue that is not fatty. Milk is made in the mammary lobes, and then transported to the nipple via ducts. Thick breasts typically have glandular or fibrous tissue as their primary composition. These types of tissues appear white, thick, and denser than fatty tissue. Organs and tissues of the breast are shown in Fig.1. The breast is a dense-tissue organ that needs to be observed.

It is true that cancer is fatal, but early detection increases the likelihood of a long and healthy life. Due to its principal sites of proliferation in the milk glands and ducts, breast cancer is particularly deadly for women.

Fatty tissue Glandular tissue Dense tissue Figure1. Different tissue type breast organs

Mammography is the name for the imaging technology used to create Mammogram pictures of a woman's breast organ. As Mammograms allow radiologists to detect abnormal lumps, they are the primary tool used in the diagnosis of Breast Cancer. Mammograms can be taken from two different angles: the mediolateral oblique (MLO) and the cranial-caudal (CC). The MLO view is the better as it observes the lateral side of the breast, which is typically affected by pathological abnormalities because it captures larger areas of the upper-outer quadrant of the breast. The connection between CC and MLO is depicted in Fig. 2. The CC view is a top-down view. This is a view of the MLO from a very specific angle [1].

Figure 2. (a)CC View (b) MLO View (c) Corresponding projection model

Breast mass is a typical early indicator of breast cancer. The process of detecting masses is being sped up with the use of numerous computer-aided diagnosis approaches based on Digital Image Processing. The presence of Pectoral muscle makes the challenge of developing a computer-aided diagnosis system to detect Mass even more difficult, as its

CVR College of Engineering

properties, such as grey level, are similar to those of Mass. The position of the Pectoral muscle on an MLO mammography is shown in Fig.3.

Figure 3. Mammogram with Pectoral Muscle

The presence of Pectoral muscle raises the likelihood of a false positive diagnosis of breast cancer, even if a breast mass is a frequent sign of the disease. This work seeks to implement a Digital Image Processing system based on Most Stable Extremal Regions (MSER) to detect and mask the Pectoral muscle in MLO Mammograms to help with Mass detection.

Section II explains existing Pectoral mask systems, Section III discusses the proposed system, Section IV provides the ensuing analysis, and Section V summarizes the new strategy to detect and mask the Pectoral muscle.

II. LITERATURE SURVEY

S. Marijeta et al. [2] employed mammography contrast improvement and K-means to identify and hide the Pectoral muscle. The K-means technique is used to find the Pectoral muscle area after trimming the original MLO mammography and boosting pixels in the mass. The limits of the Pectoral muscle are determined through cubic multinomial fitting. The Mini-MIAS database analysis of simulated mammograms shows the procedure to be reliable. With 19.25% of instances being deemed acceptable and 12.42 % being deemed poor, this approach is 68.42% accurate.

P. S. Vikhe et al. [3] describe a Pectoral covering technique for detecting the Pectoral area in mammograms using a grey level-based approach. By enlarging the Pectoral region and choosing its boundary points, the method locates the Pectoral muscle's border. With the Least Squares Error method, the initial arbitrary limits are smoothed out. The success percentage of the suggested strategy is 96%.

Wavelets based Pectoral muscle masking strategy is proposed by M. Mughal et. al [4]. In which the Pectoral border line is found by using two-dimensional multilevel wavelets. The approach has a success rate of 76.63%, with 14.59% of results considered good and 7.76% considered poor. However, the suggested solution is only partly automated.

The Pectoral covering approach is proposed by Saeid Asgari Taghanaki et al. [5] using the area-developing strategy and the Shape principles. The separation of the Pectoral muscle is accomplished using a method based on geometrical principles. It doesn't matter whether your pecs are large or rounded for this technique. Finding the Pectoral

muscle's edge is necessary for segmentation, but if the texture is too complicated for edge detection techniques based on colour and brightness, this may be challenging. This approach is based on geometric principles, so it would work with a variety of Pectoral muscle shapes. The Pectoral area is located using a modified region development approach with automated seed finding, which allows for a broad range of muscle forms and orientations. It succeeded in delivering precise results in 67% cases, acceptable results in 22% of the cases and 10% of instances are inappropriate.

Samuel Rahimeto and colleagues [6] employed Ostu's multi-thresholding to locate the bright area that corresponds to the observed Pectoral muscle. This is accomplished by determining the ideal threshold value. Overall, Pectoral muscle area is detected accurately in 93.36 percent of the cases.

The Bounding Box method is developed by Enas Mohammed Hussein Saeed and colleagues [7] based on Region Growing technique with the aim of accurately localising the Pectoral muscle. To solve the problem of Pectoral muscle masking, this study employs a practical solution that combines the Bounding Box (BB) and Region expanding procedures. In this research, pre-processing the mammography pictures is done in two steps. Using a medium filter and a threshold binary picture, first the label and noise are removed. To enhance outcomes while depleting Pectoral muscles, the second phase is combining the bounding box and region growth processes into a unified system.

Pascal Vagssa and his associates proposed a Hough transform-based method to localise the Pectoral muscle [8]. The researcher's conclusions had a 93.8% accuracy rate and a 6.2% error rate. One of the few limitations of this strategy is that the 512x512 pixel region of interest exercised in the upper left corner of the mammography may not be ideal for all Mammogram images.

A strategy for masking Pectoral tissue in mammograms has been developed by G Vaira Suganthi et al. [9], J Sutha, M Parvathy, and C Durga Devi by employing active contour and grey level thresholding methods to define the border between Pectoral and breast tissue. The proposed method accuracy is 92.55 percent. The maximum intensity of each line is used in the suggested approach for determining Pectoral boundaries, which may work for certain images only.

III. PROPOSED SYSTEM

The Pectoral muscle can be seen in MLO Mammogram. The issue with Pectoral muscle is that it has comparable qualities as Mass. As a result, the computer aided diagnosis system could detect the Pectoral region, or a portion of it as a Mass. So, in a computer-aided diagnosis system, it is always recommended to cover the Pectoral region before testing for breast tumours to avoid false positives. In the proposed system, the MSER based approach is used to locate the Pectoral muscle regions. Fig. 4 shows proposed procedure to locate and mask the Pectoral muscle. Cropping the huge image contour to remove artefacts from original MLO Mammogram is the first step in the procedure. In step 2, the resultant Mammogram orientation is set to the right.

The Pectoral area is located, pruned and covered in the third phase of the recommended technique. All phases are detailed in subsections.

Figure 4. Top level diagram of proposed system

A.Mammogram Biggest Entity Extraction

Extra black margins and artefacts such as labels would be present in the original MLO Mammogram. The original Mammogram must be cleaned of these artefacts and extra black margins so that the resulting Mammogram is free of artefacts and contain the exact Mammogram part, resulting in faster processing.

The Mammogram biggest entity extraction procedure extracts largest object from the original Mammogram. This is achieved in two steps. Firstly, the Cropping process crops out the breast region in accordance with the Otsu's [10] method. The threshold is in accordance with the Otsu's method. The limit value follows the Otsu formula. Ostu's global thresholding technique finds the threshold by reducing the weighted variance within the class. The image's histogram is analysed, and the components are ranked by minimising the variation between them. Once the cut-off for an Otsu has been determined, morphological techniques like morphology close and open operations are used to filter out insignificant details and partition the breast area from the rest of the body. Cropping of an example MLO Mammogram result is shown in the figure 5.. Detailed steps for extracting largest Object from MLO Mammogram is described in the following algorithm 1.

B.Right Orientation of an Image Step 5: Return the altered Image. Algorithm 1: LargestObjectExtraction(Image) Step 1. Filter the image using a Gaussian filter to get rid of the fuzziness. Step 2. Utilize Otsu's and morphological techniques to identify the greatest contour. Step 3. Find the perimeter of the biggest rectangle. Step 4: Use the rectangle drawn in Step 3 to crop the image.

The Pectoral region-part will be situated in the top left or right corner of the retrieved biggest picture once the largest object has been extracted from the original mammography. The next step in finding the Pectoral area is to rotate the extracted Mammogram to the right. As a result, before applying the Pectoral muscle region identification technique, all left-oriented images are rotated to the right. The total intensity of the top left and top right corners of cropped Mammogram is computed and compared to ensure that the cropped mammography is oriented correctly. The cropped mammogram has to be reoriented if the overall intensity of the left top-left corner is lesser than that of the right topright corner. In such a case, a horizontal mirror image is created from the cropped picture to get the right oriented image. Algorithm 2 shows the steps involved in the setting of orientation.

Sample output of right orientation process is shown in the Fig. 6

 (a) (b) Figure 6. (a) Left-oriented MLO Mammogram (b) After right Orientation

.

(a) (b) (c) Figure 7. (a) Input Image (b) both MSER⁻ & MSER⁺ outcomes for the Input Image shown (a) (C)Only MSER+ outcomes with a smaller number of thresholds

C.Detection of Pectoral Muscle with MSER Technique

Once the raw Mediolateral Oblique Mammogram data has been processed, the resulting picture is clean, properly cropped, and orientated to the right. Here, we present a Maximally Stable Extremal Region (MSER)-based technique for identifying Pectoral muscle regions in preprocessed pictures.

A new approach based on the Maximally Stable Extremal Regions (MSER) technology is created and developed to locate the Pectoral muscle. This system is developed specifically for identifying the Pectoral region. The innovative method consists of two main stages.

The first step, known as adaptive MSER, involves making adjustments to the current MSER, and the second, detecting the greatest Pectoral area based on the first step's output, is known as the detection stage.

In the first stage, the fundamental MSER method [11] finds a set of stable connected components throughout the whole range of grayscale values from 0 to 255, inclusive. In the process, every intermediate step involves thresholding the image to produce a series of monochrome photographs. A single extreme feature will be all white, while the other will be entirely in black color.

Black-to-white (MSER-) and white-to-black (MSER+) MSER operations create redundantly coupled components. A more detailed example is shown in Fig.7. above. The polygon in each line of Fig. 7(b) represents the MSER stable linked components.

Adaptive MSER is a kind of MSER that is optimised to find stable related components and eliminate duplicates more quickly. Instead of handling the whole grayscale range, the adaptive MSER just deals with the range from white to black (MSER+) (MSER) with. The improved MSER version produced the results seen in Fig.7(c) when fed the data from Fig.7(a). Eleven threshold values are used here to classify related parts. By default, we set the threshold for Pectoral muscle intensity at 50 relative to the value of the observation. However, the number of grayscale levels varies according to an image's average intensity. If a Mammogram's typical intensity is 110, for instance, the lowest acceptable intensity would be 90 (110 minus 20). To account for MSER zones whose values are always within 20

of the mean, we subtract the mean from the MSER value. Using these threshold values, we can analyse data and extract stable linked components from a reduced set of images, as compared to standard MSER. The streamlined MSER only generates necessary images without producing duplicates. Using the height of the calculated region, these images are further filtered to retain just the Pectoral region. The Pectoral area is what you'll find in the top left corner. Since the maximum height of the Pectoral region is equal to the maximum height of the Image minus 150, the height of the top left corner of the calculated Pectoral area must be less than the height of the image minus 100.

Second-stage processing involves filtering all related areas detected in stage one down to only those in the Pectoral region. The Pectoral region is then found to include the most interconnected areas.

The algorithm 3 gives steps involved in adaptive MSERbased methods to identify the Pectoral area. Step 2 of this algorithm finds every linked component in a given MLO Mammogram. At step 3, connected components associated with the pectoral muscle are filtered. The variation for each threshold is determined in step 4 and only components with a local minimum are chosen. The component with the greatest area is chosen as the pectoral muscle region from the filtered list of components. The chosen largest area component is then trimmed and masked. Fig. 8 shows output of a sample MLO Mammogram.

Figure 8. (a) Pre-processed Image, (b) Black mask of Pectoral tissue ground truth (c) Black mask of Pectoral tissue computed by MSER Based technique

Key aspects of MSER are maximum area variation, minimum area and maximum area. To ensure that only very small regions are rejected by the adaptive MSER, the minimum area to exclude is set to fifty, while the maximum area at discard is set to. This maximum area function would be helpful in cases when MSER specifies the whole Breast as the region. The Max Area variation has been essential since it determines the maximum intensity fluctuation between extreme places. Dimensions of stable regions are unchanging across a broad range of delta () thresholds. Standard values for delta () are 0.1 and 1.0. The MaxArea variation feature is applied to one of the values determined via experimental research [0.25,0.5, 0.75].

The following are the detailed steps for determining Pectoral regions based on Maxarea variation.

- 1. First, count filtered regions with Maxarea variation 0.25. if filtered regions are discovered. Use Filtered Regions. Otherwise proceed.
- 2. Count filtered areas with Maxarea variation 0.5. If higher than zero, follow the filtered regions; otherwise, continue.
- 3. Step Third, counts filtered areas with Maxarea variation 0.75. If count>0, utilise filtered regions; otherwise, move to the next step.
- 4. No regions are considered if all stages have zero regions.

Fig. 9 shows MaxArea changes cases.

Figure 9. MSER regions with (a) change of 0.75 (b) change of 0.5 (c) change of 0.25

One problem with the direct MSER based approach is that if any Mass region is very adjacent to the Pectoral region, and then it includes that Mass part also as the part of the computed Pectoral region. The following figure shows the concerned sample.

(a) (b) (c) Figure 10. (a) input Image (b)Actual Pectoral (white mask) (c) Computed Pectoral with MSER based approach(white)

This problem is resolved by pruning the detected Pectoral region using the Pectoral width-height triangle. Here, Pectoral width and height of detected Pectoral region is computed. Using this data, a triangle is constructed and pruned. The detected Pectoral region for the sample given in the above figure after pruning and corresponding Mammogram with Pectoral region mask is shown below.

Figure 11. (a) Detected pectoral part before trimming(white) (b) Pectoral region after trimming(white mask)

The MSER-based technique is very effective at detecting the Pectoral area. It is able to identify the pectoral area, which may be of varying shapes.

IV. PERFORMANCE METRICS AND RESULTS ANALYSIS

The recommended technique is validated using the Mini-MIAS database. There are 322 MLO Mammography images overall in this database. The database's mammogram IDs are mdb001, mdb002, through mdb322. While some Pectorals are parabolic in shape, the majority are triangular. According to the background tissue, there are three different kinds of images in Mini-MIAS: fatty, fatty-glandular, and dense-glandular. In dense tissue, Pectoral muscle detection

CVR College of Engineering

is difficult. Every picture captured by Mini-MIAS is analysed. Classification performance is evaluated using a variety of metrics, including Global Pixel Accuracy (GPA), Following is the formula used.

$GPA = (TP + TN)/(TP + FP + TN + FN)$

Pectoral regions (both detected and ground-truth) were used in the assessment of performance. The FP value indicates the total number of Pectoral area pixel assignments that should have been made to the background but were instead made to the foreground, whereas the FN value indicates the total number of background pixel assignments made to the Pectoral area. In this scenario, TP is the aggregate of all reliable forecasts for pectoral muscle pels. The TN value is the estimated total number of non-Pectoral pels with correct classifications.

In 275 pictures, the system found Pectoral muscle area with an accuracy score $> 97\%$ or above, which is deemed accurately segmented. There are 32 pictures where computed Pectoral area accuracy is nearly good that were declared as acceptable since the false positives and false negative rates were less than four percent. In 12 photos, Pectoral area finding is incorrect. So, the total number of instances with good Pectoral area finding is $307 (275 + 32)$. In the three photos, there is no pectoral muscle. So, the total number of images where Pectoral detected and covered correctly are 310. The mean values of the metric measures of the 298 Images is 99.128 \pm 0.654. The Pectoral masking analysis summary of Mini-MIAS is shown in Fig.12. The accuracy achieved is (310/322) 96.2% which is a good improvement over many surveyed approaches. The suggested method is compared to the current literature in the table below. There is evidence that the suggested approach is an improvement above the state-of-the -art.

TABLE I.

Figure 13. Bar diagram of comparative analysis of proposed approach for detection and covering of Pectoral muscle with existing literature.

The Fig.13 clearly shows us that the proposed approach can detect and cover the Pectoral muscle region with very good accuracy.

V. CONCLUSIONS

In this work, the design and development of new techniques to identify and cover Pectoral muscle is explored. The suggested solution outperformed the state-of-the-art methods in about 96.2% of Mini-MIAS cases, leading to favourable findings. The proposed approach cannot detect Pectoral region if the Mammogram is dense, so, it can be extended to detect the Pectoral regions in dense Mammograms also.

REFERENCES

- [1] https://www.mskcc.org/cancercare/types/breast/anatomybreast.
- [2] S. Marijeta, A. Gavrovska, M. Milan, I. Reljin and B. Reljin "Breast Region Segmentation and Pectoral Muscle Removal in Mammograms," Telfor Journal, Vol. 8, No. 1, 2016.
- [3] Vikhe, P.S., Thool, V.R. "Detection and Segmentation of Pectoral Muscle on MLO-View Mammogram Using Enhancement Filter," J Med Syst 41, $\overline{190}$ (2017). https://doi.org/10.1007/s10916-017-0839-8.
- [4] Stefanoyiannis AP, Costaridou L, Skiadopoulos S, Panayiotakis G. "A digital equalisation technique improving visualisation of dense mammary gland and breast periphery in mammography," Eur J Radiol. $2003;45(2):139-149$. doi:10.1016/s0720-048x(02)00057-8.
- [5] Raju, Rajeswari & Maul, Tomas & Bargiela, Andrzej. "New image processing pipelines for membrane detection," Journal of the Institute of Industrial Applications Engineers" 3. 15-23. 10.12792/JIIAE.3.15.
- [6] Rahimeto, S., Debelee, T.G., Yohannes, D. et al. "Automatic pectoral muscle removal in mammograms. Evolving Systems," 12, 519±526 (2021).https://doi.org/10.1007/s12530-019-09310- 8.
- [7] Enas Mohammed Hussein Saeed, Hayder Adnan, "Pectoral Muscles Removal in Mammogram Image by Hybrid Bounding Box and Region Growing Algorithm,".
- [8] Vagssa P., Doudou N.M., Jolivo T., Videme O., Kolyang D.T. 2020. "Pectoral muscle deletion on a mammogram to aid in the early diagnosis of breast cancer," International Journal of Engineering, Science and Technology, Vol. 12, No. 3, pp. 57- 65. doi: 10.4314/ijest.v12i3.6.
- [9] Suganthi G. V, Sutha J, Parvathy M, Devi C. D. "Pectoral Muscle Segmentation in Mammograms," Biomed Pharmacol J 2020;13(3).
- [10] https://learnopencv.com/otsu-thresholding-with-opencv/.
- [11] Fredrik Kristensen and W. James MacLean "Real-Time Extraction of Maximally Stable Extremal Regions on an FPGA".