Segmentation of Pectoral Muscle in Mammograms Using Granular Computing

Divyashree B. V., University of Mysore, India
Amarnath R., University of Mysore, India
Naveen M., University of Mysore, India
Hemantha Kumar G., University of Mysore, India

ABSTRACT

In this paper, pectoral muscle segmentation was performed to study the presence of malignancy in the pectoral muscle region in mammograms. A combined approach involving granular computing and layering was employed to locate the pectoral muscle in mammograms. In most cases, the pectoral muscle is found to be triangular in shape, and hence, the ant colony optimization algorithm is employed to accurately estimate the pectoral muscle boundary. The proposed method works with the left mediolateral oblique (MLO) view of mammograms to avoid artifacts. For the right MLO view, the method automatically mirrors the image to the left MLO view. The performance of this method was evaluated using the standard mini MIAS dataset (Mammographic Image Analysis Society). The algorithm was tested on 322 images, and the overall accuracy of the system was about 97.47%. The method is robust with respect to the view, shape, size, and reduces the processing time. The approach correctly identifies images when the pectoral muscle is completely absent.

KEYWORDS

Ant Colony Optimization, Granular Computing, Malignancy, Mammograms, Pectoral Muscle Segmentation

1. INTRODUCTION

For decades, breast cancer is considered as number one potentially lethal disease reported among women worldwide. According to Union health ministry, government of India, breast cancer occurred is about 25.8 per 100,000 women and mortality reported is 12.7 per 100,000 women (Malvia et al 2017). Early detection of breast cancer is crucial in reducing the mortality rate (WHO, 2017) and increases treatment options. Mammography is the most powerful and successful breast screening tool that is used to detect cancer in the early stage. A breast mammogram contains different parts: mainly the breast region, pectoral muscle, high-density label, low-density label and background (Figure 1(a)). The computer-aided classification of mammograms into normal and abnormal is an emerging research area in the field of medical image processing. So far, the identification of malignancy in the breast region in mammograms, ignoring the pectoral muscle, has been attempted. However, a recent study states that the presence of masses in the pectoral region indicates that it could also be malignant (Figure 1(b, c, d)). Hence, it is important to analyze...
The segmentation of the pectoral muscle in mammograms is a challenging task because the mammogram images have low contrast. Further, the appearance of the pectoral muscle varies in size, shape, and orientation (left and right), having unclear boundaries between the pectoral region and the breast region (Mustra, Grigic & Rangayyan, 2016). In the state-of-the-art, researchers have attempted to segment the breast region by suppressing the pectoral muscle. However, in this paper, the effort is to segment the pectoral muscle for further analysis using granular computing and the ant colonization approach. Granular computing is used by researchers for exploring the masses in mammograms with some different statistics (Roselin & Thangavel, 2012) but in the proposed work, only pixel intensity values and mean statistics are considered for pectoral muscle segmentation. This step helped in eliminating additional enhancement techniques and reduced the processing time by removing the artifacts and labels present in the background region automatically. Ant colony optimization algorithm adopted in the proposed paper is slightly different from the ant colony optimization algorithm used in the state of art (Karnan, Thangavel & Ezhilarasu, 2008) which is discussed in section 3. The proposed method was tested on the benchmark mini MIAS dataset.

The paper is organized as follows: the related research work is presented in section 2; the materials and methods are discussed in section 3; section 4 shows the experimental results and discussion and section 5 draws conclusions with future avenues.
2. RELATED WORKS

Many researchers have proposed methods with regard to pectoral muscle boundary identification and removal. Existing studies on pectoral muscle segmentation focused on different approaches to accurately segment the pectoral muscle, among which the most popular ones are region growing, polynomial fitting, straight-line estimation and thresholding methods.

Rampun et al. (2017) applied the canny edge detection method, extracted edge features, and adopted the contour region growing method based on the edge features to detect the actual boundary of the pectoral muscle. Hazarika & Lipi (2018) presented a novel seeded region growing method to remove pectoral muscle from mammograms. Local region growing algorithms applied within the triangular segmented of the pectoral muscle. Here, region growing method is limited to well-defined triangular area that divides the image into area with pectoral muscle and the area without pectoral muscle. Makandar & Halalli (2016) presented segmentation technique combining morphological thresholding method and region growing method. They also discussed various filtering techniques and contrast adjustment techniques. The region growing methods used in the state-of-the-art ended with errors in segmentation and increased processing time.

The straight-line approximation method with the help of shape features and average gradient has also been used by Chakraborty et al. (2012). Esener et al. (2018) proposed a novel multi-stage system for the detection and removal of pectoral muscles in mammograms. In the work, mammographic images were first preprocessed by performing noise reduction, background removal and artifact suppression processes. Then, an algorithm is applied based on positional information of pectoral muscles followed by line fitting considering some geometrical characteristics. The straight-line approximation method did not perform well for mammograms with no pectoral muscle and failed in cases where the pectoral muscle region and breast region were almost the same, the edges were not detected properly in images with a curved boundary of the pectoral muscle.

Vikhe & Thool (2016) proposed a method for pectoral muscle identification using a linear enhancement mask based on the features of the pectoral muscle and selected boundary points of the pectoral muscle with the help of a thresholding technique. A method for identification and segmentation of the pectoral muscle using global thresholding and edge detection methods was suggested by Sreedevi & Sherly (2015). Toz & Erdogmus (2018) proposed a method for the detection of pectoral muscle based on Single-Sided Edge Marking (SSEM) technique. They extracted geometric features and neighborhood relations followed by a linear interpolation method to establish borders of the pectoral muscle. The adopted thresholding methods failed in cases where there was no pectoral muscle or a curved boundary of the pectoral muscle.

Vikhe & Thool (2017) suggested an enhancement filter and thresholding method to choose the pectoral boundary points followed by curve fitting to refine the boundary points. Shen et al. (2018) presented a combination of genetic, morphology selection, and curve fitting algorithms that could be used to detect the pectoral muscle. Recently Shi et al. (2018) presented a hierarchical pipeline for segmentation and detection of calcification, wherein the skin–air boundary is detected using a gradient weight boundary; unsupervised pixel-wise clustering was adapted to remove the pectoral muscle and the calcification was detected using a texture filter.

Yoon et al. (2016) presented morphological and random sampling consensus (RANSAC) techniques to segment the curved shaped pectoral boundaries. The combination of Orthogonal kernel, distance weight and RANSAC methods produced satisfactory results in the detection of pectoral muscle as if failed to segment some cases with complex curves.

The state-of-the-art motivated the authors to accurately detect the pectoral muscle boundary. In this context, in a recent study, Cheddad et al. (2015) mentioned that malignancy could also be witnessed in the pectoral region. Hence, to analyze the pectoral region, the segmentation of the pectoral muscle in mammograms was accomplished in the current study.
3. PROPOSED METHOD

The proposed method focuses on the segmentation of the pectoral muscle region combining the advantages of layering, granular computing and ant colony algorithms.

The architecture of the proposed method is presented in this section. The flow diagram of the architecture is depicted in Figure 2. The input to the system is a mammogram image (in RGB). The output would be the segmented or localized pectoral region in binary image format. The procedure is as follows: First, to reduce the computational time, the grayscale image was used for analysis. Second, granular computing (Bargiela & Pedrycz, 2003) was employed to distinguish the pectoral muscle and breast regions. Third, the ant colony algorithm (Dorigo, Birattari & Stutzle, 2006, Yuanjing et al., 2011) was used to accurately segment the pectoral muscle. Here, the approximation algorithm (Richard & Qiang, 2007) was inculcated to increase the accuracy rate.

3.1 Preprocessing

The input to the system was the RGB mammogram image. Primarily, the RGB image was converted to grayscale to reduce the computational time. In most cases, the image would be re-sized to suit the requirements; here, there would be considerable loss of information. To avoid such losses, the proposed model does not re-size the image. For illustration purposes, the original right and left MLO view images and their corresponding grayscale images with artifacts are shown in Figure 3.

3.2 Granular Computing

In most cases, distinguishing the pectoral muscle from the breast region is challenging. Figure 4 shows an example image with corresponding histogram where the histogram produces a graph considering

---

**Figure 2. Block diagram for the proposed method**

- **Input Image**
- **RGB to gray**
- **Granular computing**
  - Granulation of image
  - Statistical mean computing
  - Integration of granules
  - Thresholding
  - Rough set Approximation
  - Ant colonization
- **Localization of pectoral muscle**
the pixel value distribution parameter over the full image. The mammographic grayscale image obtained in the previous step possesses low contrast scenario as histogram appears to be narrow (Fig 4). Thus motivated to use granular computing for further analysis.

To solve this problem, the proposed idea is to employ the granular computing technique to find the breast region–pectoral muscle boundary. The four steps to be followed are: First, the gray image is divided into non-overlapping blocks (granules). Second, the statistical mean is computed for individual granules. Third, the computed local statistical mean values are assigned to every pixel value in the corresponding granule. Finally, the granules are integrated for further analysis. Figure 5(a) shows the grayscale image and the zoomed-in portion for better perception. Figure 5(b) shows the resultant image, which is the mean granules-integrated image (mean computed granule portion). Further, the difference between Figure 5(a) and 5(b) can be clearly witnessed in Figure 5(c). This helps to find the boundary between the breast and pectoral regions. The number of blocks is taken

Figure 3. Original right and left MLO view images (a and b); Corresponding grayscale images with artifacts in the background region (c and d)

Figure 4. Gray image with the corresponding histogram
empirically and a block size of $5 \times 5$ is found to be optimal. The experiments that were carried out to choose the block size is discussed in the experimental section.

The procedure for granular computing is as follows:

**Input:** Gray image (G)

**Output:** Filtered Image (B)

**Step 1:** Divide G into $5 \times 5$ blocks (granules)

**Step 2:** For every granule (b), compute

$$\mu = \frac{\sum b(x)}{n}$$

where $n$ represents the number of pixels in b.

Next, assign mean values to every pixel in b:

$$b(x) = \mu$$

**Step 3:** Integrate all the granules $b_1, b_2, \ldots, b_k$ as B. The size of B is equal to the size of G.

### 3.3 Ant Colony Algorithm

In general, the ant colony algorithm, a probabilistic technique, works similarly to real ants that locates good path while solving computational problems. Real ants lay pheromones to direct other ants to form a colony (Dorigo, Birattari & Stutzle, 2006, Yuanjing et al., 2011). In the present paper, ant colony
algorithm is used to provide the path to colonize the boundary of pectoral boundary accurately. Before getting in-depth into this idea, first the mean computed image obtained from the previous step is layered into four regions based on intensity values of the image to distinguish between various tissues. The idea behind this layering process was motivated by the authors’ previous paper (Divyashree et al., 2018). The function of the layering process is given below:

\[
B(i,j) = \begin{cases} 
0, & \text{if } B(i,j) < 50 \\
100, & \text{if } 50 \leq B(i,j) < 100 \\
150, & \text{if } 100 \leq B(i,j) < 150 \\
255, & \text{otherwise}
\end{cases}
\]  

(2)

The layered output image is shown in Figure 6. Here, pixels were witnessed separately as four different regions. The four different regions represent different categories comprising background region (completely black), triangular pectoral muscle and denser parenchyma tissues (completely white), fatty tissues (dark gray) and non-dense parenchyma tissue (light gray).

After the layering process, the next challenge is to identify the views (left and right) in the mammograms. For the purpose of understanding, a sample resultant image showing left and right MLO views are depicted in Figure 7.

To identify the MLO views, the layered image is vertically divided into two equal halves. Here, the notion is to find the composition of white intensity values (255) in each divided portion (left and right). The first ten rows in each portion are considered in the computation. In certain cases, the artifacts (labels) also contribute to the count of white intensity values. The number of rows to be considered is chosen heuristically (as discussed in the experimental section). The portion holding the maximum count of 255 intensity values is identified as the region of interest (ROI). In other words, if the count in the left half is found to be larger than in the right half, then the image is identified as the left MLO view. Otherwise, the image is labeled as right MLO view. Hence, the left and right MLO view images denote the appearance of the pectoral muscle in the respective portions (left and right). This MLO view identification enables segmentation of the pectoral muscle. Here, the idea is to work on the left MLO view image to ignore the artifacts. If the algorithm identifies the image
to be a right MLO view image, a vertical flip (mirror) of the right MLO view image is performed to make it appear like a left MLO view image. A sample flipped image is demonstrated in Figure 8.

The ant colonization approach is used to approximately localize the pectoral region. The pectoral region possesses a white intensity value (255). In most cases, the regions contouring the pectoral muscles are found to be non-white regions. With this idea, the transitions from the non-white region to white region (N-to-W) and vice-versa (W-to-N) are spotted in the mammograms. Here, the top-down approach is followed. Algorithmically, the image is scanned from the left to right direction and each occurrence of transition is noted. Here, the first transitions (N-to-W and W-to-N) for each row are sufficient to locate the pectoral muscle. This process takes advantage of flipping (mirroring) the right MLO to appear like a left MLO.

Most pectoral regions are found to be triangular in shape. Therefore, the transition (W-to-N) for adjacent rows should follow a monotonically-decreasing order. Here, checkpoints along the transition positions (W-to-N) are employed to ensure that the region holds a monotonically-decreasing function. These checkpoints are tracked with an offset value, that is, if a checkpoint is positioned at row ‘i’, the next checkpoint would be positioned at row ‘i+c’ (here the constant ‘c’ represents the offset and is empirically chosen. The flow of checkpoints is traced all along the pectoral region. The process stops only when the transitions (W-to-N and N-to-W) are found to be same. In other words, the process halts when the condition of monotonically-decreasing function is violated. The entire process was adapted from a soft-computing technique (ant colony optimization). The checkpoints are depicted as red points along the transition (W-to-N) for the zoomed-in portion of the image as shown in Figure 9.

After localization, if the image is mirrored and the image is found to be a right MLO view, then the resultant image is flipped back to its original format. The localized pectoral muscle is shown in Figure 10.
4. RESULTS

For the validation of the proposed algorithm, a mini MIAS dataset was used. The dataset contains 322 mammographic images including fatty, fatty-glandular and dense glandular tissues. The ground truths were provided with respect to abnormalities in the datasets. But, ground truths were not provided for the pectoral muscle boundary and hence, the ground truths were marked manually in consultation with the expert radiologist. For the study, both left MLO and right MLO images were considered and performance of the model is reported. Irrespective of the left or right MLO views, the presented algorithm could extract the pectoral muscle efficiently. Also, the proposed algorithm was able to automatically remove artifacts in the images (Figure 10). For visualizing the error (difference) in segmentation of pectoral muscle, the segmented pectoral images obtained were matched with the manually marked ground truth images by performing XOR operation as depicted in Figure 11.

The performance comparison between the proposed method and the state-of-the-art methods for pectoral muscle segmentation was done considering parameters such as accuracy rate (ACC), true positive rate (TPR) and false-positive rate (FPR). The accuracy is the sum of the pixels that were correctly segmented as pectoral muscle points and non-pectoral muscle pixels correctly identified as non-pectoral muscle points, divided by the number of pixels in total:

\[
ACC = \frac{TP + TN}{TP + FP + TN + FN}
\]
The true positive rate is the ratio between the number of true positive pixels and the number of true pectoral muscle pixels in total:

\[ TPR = \frac{TP}{(TP + FN)} \]  

(4)

The false-positive rate is the ratio between the number of false-positive pixels and the number of pixels marked as non-pectoral muscle in the true background:

\[ FPR = \frac{FP}{(FP + TN)} \]  

(5)

The accuracy rate, TPR and FPR for the dataset are tabulated in Table 1.
To efficiently segment the pectoral muscle in mammograms, three parameters were used in the proposed model, namely, block size (n), number of layers (m) and offset (c). The parameters n and m were referred from the previous work (Divyashree et al., 2018). A comparison of the accuracy rate by varying the parameters is given in Table 2. Empirically, the values for the block size, number of layers and offset of 5 x 5, 4 and 10 respectively provides a better accuracy rate for the mini MIAS dataset.

The advantages of the proposed method for segmentation of pectoral muscle are: (i) irrespective of the size, the views of the pectoral muscle were able to segment the pectoral region; (ii) because of flipping (mirroring) of the right MLO view, the artifacts were automatically ignored; (iii) correct identification is possible even when the mammograms do not have a pectoral region and the well-segmented images show optimal accuracy when compared to the state-of-the-art. The limitation of the proposed method is that it is unable to detect multi-layered pectoral muscles.

The proposed method is a novel approach and was compared with other methods to check the performance. The other approaches, detailed in the literature, were able to find the boundary between the pectoral region and the breast region. Further, these approaches suppressed the pectoral muscle. The comparison is shown in Table 3.

Summarizing the performances of the proposed method with the existing methods, region growing approach (Hazarika & Lipi, 2018) and thresholding approaches (Sreedevi & Sherly, 2015) alone produced comparatively fewer accuracies than other approaches. Combined approaches involving thresholding method and line fitting methods (Vikhe & Thool, 2017) performed better than them. Whereas, our method outperformed all other recently reported methods by achieving maximum accuracy. Some studies focused on selected images for the validation of their algorithms (Toz & Erdogmus, 2018). Despite, proposed method has experimented on all types of images included in the dataset.

5. CONCLUSION

In this study, an attempt was made to segment the pectoral region to investigate the malignancy in it. Since the mammogram images are of low contrast, segmentation of the pectoral muscle, appeared to be more challenging. This paper presents an approach to extract the pectoral muscle alone using

<table>
<thead>
<tr>
<th>Table 1. Segmentation performance of the proposed algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Number of images</td>
</tr>
<tr>
<td>249</td>
</tr>
<tr>
<td>TPR</td>
</tr>
<tr>
<td>FPR</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. Comparison of accuracy rate by varying the parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block size</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>5 x 5</td>
</tr>
<tr>
<td>10 x 10</td>
</tr>
<tr>
<td>20 x 20</td>
</tr>
</tbody>
</table>
granular computing and the ant colonization algorithm. The proposed algorithm was experimented on MIAS dataset and yielded 97.47% of accuracy. The advantage of the proposed method is that it is robust with respect to size and view. Flipping of the right MLO view to the left MLO view automatically ignores the artifacts present in the mammograms which is an added advantage of the present paper. However, the system fails in some situations because of the errors in layering process. This approach initiates a novel idea of studying the presence of malignancy in the pectoral region in future. The accuracy rate could be improved further by solving the layering problem.

**ACKNOWLEDGMENT**

The authors would like to thank Dr Deepashree Basavalingu, Consultant Radiologist, Blackpool Teaching Hospitals, NHS Foundation Trust, the United Kingdom for her certification of ground truths, valuable help and comments in carrying out this work. The first author would like to thank the Ministry of Tribal Affairs, Government of India for awarding the National Fellowship (201718-NFST-KAR-00159) to carry out this research work.

**Conflict of Interest:** No conflict of interest.
REFERENCES


B. V. Divyashree is currently doing PhD in Computer Science under the guidance of Prof Hematha Kumar G. Her research interests are medical image processing, mammogram analysis for the segmentation, detection, and classification.

Amarnath R. completed M Tech in Computer Science and applications from DoS in Computer Science, University of Mysore, Mysore. He has successfully completed his Ph.D. in Computer Science at the Department of Computer Science, University of Mysore. His research interests are Image Processing, Classification, Pattern Recognition, Algorithms, Artificial Intelligence.

Naveen M. completed his Ph.D. in Computer Science at the Department of Computer Science, University of Mysore. His research interests are Image Processing, Machine learning, Pattern Recognition, Artificial Intelligence.

G. Hemantha Kumar did his B.Sc (1980), B.Ed.(1984) M.Sc (1989), Ph.D. (1997) from the University of Mysore. Prof. G. Hemantha Kumar has published more than 448 research papers in National and International peer-reviewed journals and presented more than 150 research papers in national and international conferences. He has over 1446 scientific citations and H-index 24 for his research contributions as of now. He has given invited and plenary lectures in a large number of National/International meetings/Conferences. Prof. G. Hemantha Kumar has successfully guided 20 candidates for Ph.D. degree and currently 08 students are working towards a research degree.


