**Indian Journal of Science and Technology,** *Vol 9(40), DOI: 10.17485/ijst/2016/v9i40/99589, October 2016*
*ISSN (Print) : 0974-6846*
*ISSN (Online) : 0974-5645*

# Mammogram Segmentation using Region based Method with Split and Merge Technique

**G. R. Jothilakshmi[1*], P. Sharmila[2] and Arun Raaza[3]**

[1]Vels University, Pallavaram, Chennai - 43, Tamil Nadu, India; Jothi.se@velsuniv.ac.in
[2]Comm System, Vels University, Pallavaram, Chennai - 43, Tamil Nadu, India; psharmi88@gmail.com
[3]Centre for Advanced Research and Development, Vels University, Pallavaram, Chennai - 43, Tamil Nadu, India; egopileo@gmail.com

## Abstract

**Objectives:** *Now a days*, the major health risk in woman is breast cancer. In order to reduce the death rate, early detection of cancerous region in mammogram images is needed. But finding the lesion part and its spread from the mammogram image is very difficult. To identify the type of cancer, segmentation of lesion region is important. To perform further breast cancer classification, this paper proposes an improved segmentation algorithm using digital mammogram**. Method/Analysis:** An automated method is used to segment the affected mammogram in a effective manner using split and merging technique based on region based segmentation method by identifying a seed point .The proposed algorithm uses morphological operation to remove the noise digitally and region split and merge technique to remove the background and separate the affected region in the image. **Findings:** The efficiency of this proposed algorithm is calculated by measuring five different parameters Mean, Variance, standard deviation, entropy, correlation and the output is compared with existing technique and it is observed that proposed method shows better results than previous threshold decomposition method. **Improvement:** From the segmented output the features are extracted for finding the type of cancer from the mass and Micro calcification region

**Keywords:** Breast Cancer, Microcalcification, Mammogram Images, Region Based Segmentation, Segmentation, Seed Selection

## 1. Introduction

Breast cancer the most deadly diseases in woman. In Recent studies it has been shown that one in every ten women of Europe and one in every eight of the United states are affected by Breast cancer[1]. The abnormalities in digital mammogram are classified into masses and calcifications. Calcifications are the deposits of calcium. The cancer can be categorized as benign or malignant[2]. Benign tumors are not harmful, Whereas malignant tumors are harmful in which the cells grow abnormally.

Calcification can appear with or without an associate lesion and their etiology as well as whether they can be associated with a benign or malignant process. There are two types of calcification: Macro calcification and micro calcification[3]. Microcalcification occurs due to deposit of calcium, may be seen as clusters or as patterns and are associated with extra cell activity in breast tissue. Usually the extra cell growth is not dangerous, but sometimes constricted clusters of micro calcifications can indicate early breast cancer. Distributed microcalcification are usually a sign of benign breast cancer. 80% of microcalcification is benign[4]. Microcalcification in the breast appears as white speckles on the breast X-rays. Segmentation is the process of partitioning mammograms into regions. These regions will contain significant information and easier to analyze. The segmented mammogram can be further used by physicians to take clear cut decision about their patients' health.

The removal and recognition of masses in pectoral muscle using Connected Component Labelling was attempted earlier. The removal of pectoral muscle using

*Author for correspondence

simple thresholding method followed by connected component labelling

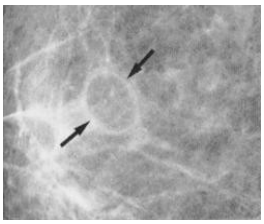


**Figure 1**(a). Benign lesions in digital mammogram.

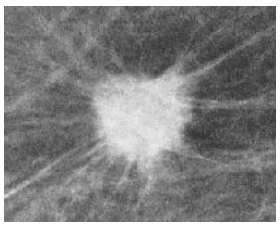


**Figure 1**(b). Malignant lesions in digital mammogram.

and an algorithm to eliminate artifacts in digitized mammograms through morphological open operation followed by reconstruction was achieved[5].

The detection and classification of microcalcification method consists of three stages namely preprocessing, detection and classification. Preprocessing is done by normalization, noise removal and pectoral muscle suppression to improve the appearance of the image and is separated from the breast region. In[6] proposed article Swarm Optimization (PSO) was used to segment the suspicious region from breast region.

In[7] used the morphology mathematical operations in digital image processing method. The investigators used the erosion operator by reconstruction with isotropic structuring element, followed by the segmentation based on contrast enhancement to find new algorithm.

Four steps were used[8] for breast mass detection and classification. The first step involved using homomorphic filtering enhancement of mammogram image. To segment the enhanced image using local seed region growing algorithm is second step. In third step features are extracted using shearlet transform. Benign mass or malignant mass classification using support vector machine is the final step. In hybrid system, ROIs are segmented using local seed region growing algorithm and it provides optimal multi resolution and yields malignant/benign classification

In Digital mammograms, The segmentation of RoI and detection of masses is achieved by morphological preprocessing algorithm. It used to eliminate digitization noises and isolate background region from the region of RoI[9].

## 2. Segmentation

The main motive of image segmentation is to get various features of the images. These extracted image features can be merged or split to build objects of interest. The built objects can be analyzed and interpreted further. Image segmentation mentions to dividing an image into groups of pixels which are homogenous with respect to some criterion. The outcome of segmentation is the splitting up of the image into connected areas. Thus segment is concerned with partitioning an image into meaningful regions. Since medical images has complex intensity distribution, thresholding, region growing, statistics models, active contour models, clustering have been used for segmenting such images. Segmentation is generally carried out based on two basic properties: Discontinuities and similarities. Discontinuities is an approach to divide an image based on sudden changes in gray-scale levels. The prime approaches are detection of isolated points, lines and edges in an image and its boundary estimation using edge detection. Similarities are based on thresholding, region growing and region splitting and merging, its region based segmentation[10,11].

## 3. Seeded Region Growing Algorithm (SRGA)

Since digital mammogram image consisting of different intensity value, the segmentation process performed on the edge map differentiates various regions on the breast. Each region has a dissimilar intensity value. The various parts of breast image like fatty tissues, glands, lobules and the ducts show different intensity values and thus can be separated into different regions. A defect such as a mass, tumors or calcifications may be present within any one of the breast parts has definitely higher intensity values than the normal tissues of the breast. So it is needed to categorize all the obtained closed structures based on their intensity values. Within each segmented region, the

pixels intensities' distribution also vary but the majority of the pixels have similar intensity values. So for each region in the original mammogram, it is necessary to find the arithmetic Mode value for the available intensities and replace those pixels in the region with the computed mode values[12].

Each closed structure is necessarily to be identified and to be isolated and then coloring of each closed structure is completed with their respective Mode values. we proposed a new algorithm by altering the seeded region growing method for coloring bounded objects. Within the mammogram, each region is bounded by a single pixel boundary. Single pixel boundary is got during the edge detection process followed by the anatomical segmentation process. During this process after performing the Anatomical Segmentation of Breast (ASB), segmented image is scanned to locate region. The needed region is yet to be colored. Based on the characteristics of the region, one of the pixels within the region is considered as seed. After seed selection for a region, the process of coloring the region is started. This could be done by first comparing the pixel intensity of that pixel location with the pixel intensity of neighboring pixels. Each pixel has four boundary pixels placed in north, east, west and south of the pixel . The pixels around the seed in all four direction is also verified to check whether they are colored or they form the boundary pixel.

If the pixels are not identified as boundary pixels and not confining within colored region, they again form the seed for further searching. A stack is used to give information about the seeds to be considered. A List has been used to collect the pixels of the regions that have been included to the region. All the pixel positions within the List are then searched on the original image to get their intensity values for calculating the Mode value. The pixel locations of each region are then replaced by the computed Mode value intensity.

## 4. Region Splitting and Merging Algorithm

The basic idea of region splitting is to break the image into a set of disjoint regions which are coherent within themselves. The flow for the region splitting and Merging is as follows.

- Initially consider the image as a whole to be the area of interest.
- Look at the area of interest and decide if all pixels contained in the region satisfy some similarity constraint.
- If **TRUE** then the area of interest corresponds to a region in the image.
- If **FALSE** split the area of interest (usually into four equal sub-areas) and consider each of the sub-areas as the area of interest in turn.
- This process continues until no further splitting occurs.
- If only a splitting schedule is used then the final segmentation would probably contain many neighboring regions that have identical or similar properties.

Thus, a ***merging*** process is used after each split which compares adjacent regions and merges them if necessary.

The Algorithms can be explained with the following figure 2.

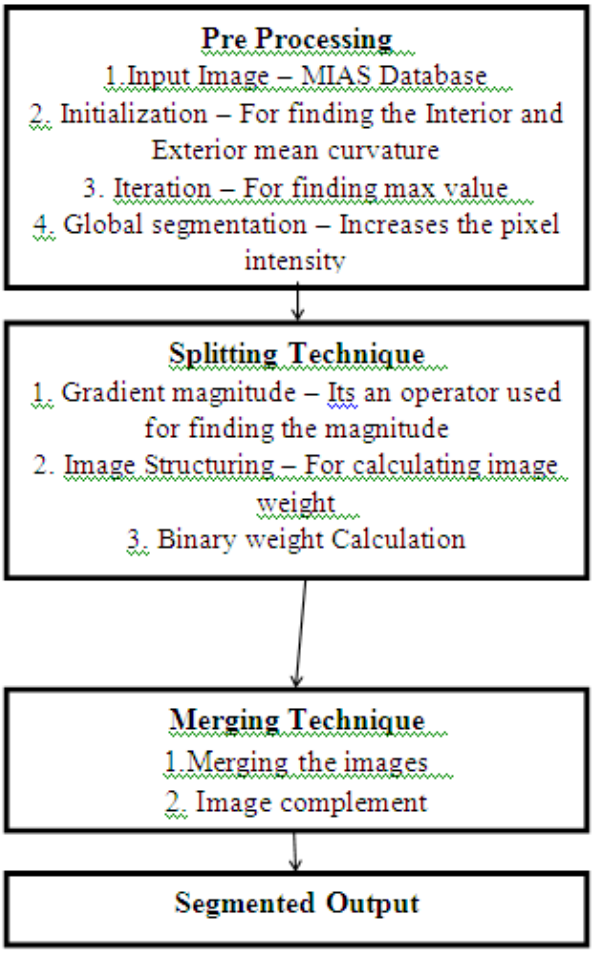



**Figure 2.** Flow chart for Proposed methodology.

## 5. Proposed Methodology

In this paper segmentation of digitized mammogram image is carried out based on region splitting and merging technique. The flow chart for the proposed work is shown in figure 3 and the step by step explanation as follows.

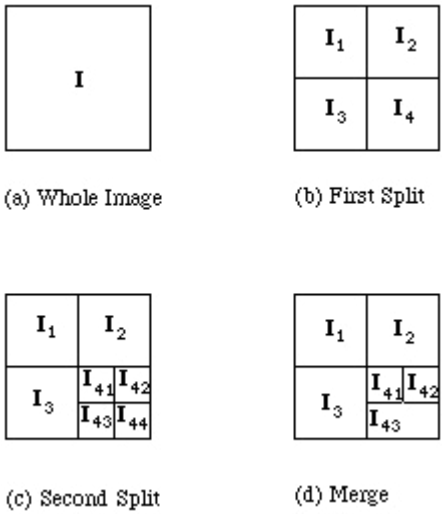



**Figure 3.** Region splitting and Merging diagram.

### 5.1 Preprocessing

The input image is taken from the MIAS database. The Mammographic Image Analysis Society (MIAS) is an UK research group organization interested in the understanding of mammograms and has generated a database of digital mammograms. The images from this database have been digitized to 50 micron pixel edge and representing each pixel with an 8-bit word. Totally 322 digitized films are available and with the report of experienced radiologist. The image size with in this database is 1024x1024. The mammogram image is an unstructured hence for calculating the image weights, we are converting it to structured image using curvature by applying mask. The mammogram interior and exterior curvature mean value are calculated and the image is segmented for initialization. The contour image is iterated and magnified for locating the masses. Sussman level set method is used to set the level of the mammogram image. This method is used for finding the origin of the image and to find the seed point in the mammogram. The image is converted to gray scale for global segmentation to increase the pixel intensity.

### 5.2 Splitting Technique

The gradient magnitude is calculated from the global segmented output using sobel operator, for finding the maximum magnitude. The gradient magnitude is an operator that is used for calculating the magnitude by finding the square root of rows and column. The magnitude value is grown from the seed point in the mammogram image. The unstructured mammogram image weights are calculated. The pixel weights are converted to binary values.

### 5.3 Merging Technique

The seed point is grown till it satisfy the stopping rule. The binary values are reconstructed in an order to form a structured mammogram image. The structured image is completed for finding the seed point and grown points. The segmented output is found.

## 6. Results and Discussion

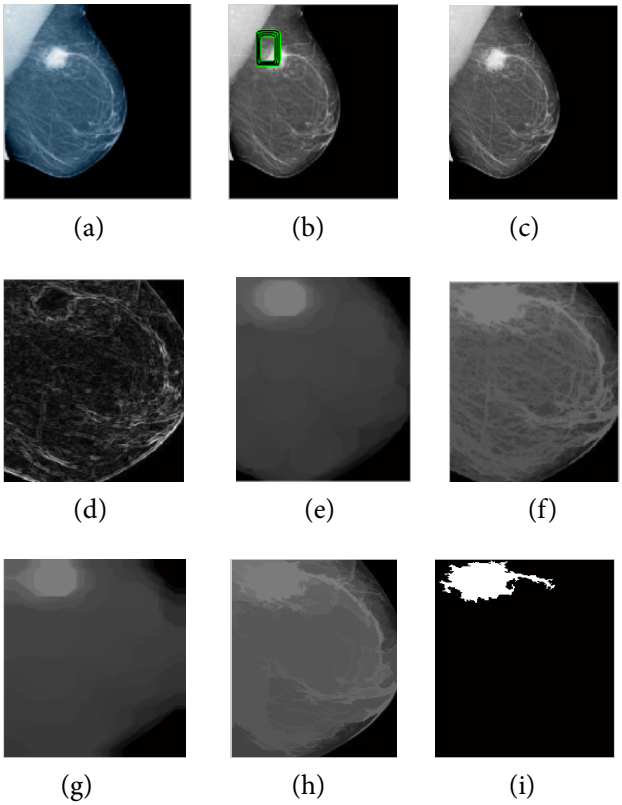
The efficiency of the proposed algorithm is found by calculating five parameters : mean, varience, entropy, standard deviation and correlation by considering a diagnostic digital mammogram. The same five parameter are calculated using existing threshold technique. In thresholding algorithm, the threshold values are compared and the maximum values are found which are useful to segment the mammogram images. The parameters obtained from proposed and existing techniques are compared with standard data base and the output shows the proposed algorithm produces better result than the existing techniques. Figure 4 to 5; table1

**Table 1.** Parameters calculation and comparision

| Parameters | Proposed (SRGA) | Existing (Thresholding) |
|---|---|---|
| Mean | 0.0759 | 0.1899 |
| Variance | 0.0702 | 0.4481 |
| Standard Deviation | 0.2649 | 0.6694 |
| Correlation | 0.7869 | 1 |
| Entropy | 6.521 | 4.7404 |

The Proposed Region Growing Algorithm is compared with the Existing Threshold Segmentation Technique. The existing algorithm used for the comparison is threshold algorithm. In this algorithm the threshold values are

compared and the maximum values that are found and segmented for the image. The parameters that are used for comparison are Mean, Variance, Standard Deviation and Correlation. Which Produces better result than the existing techniques. (Figure 4 to 6 and table1)

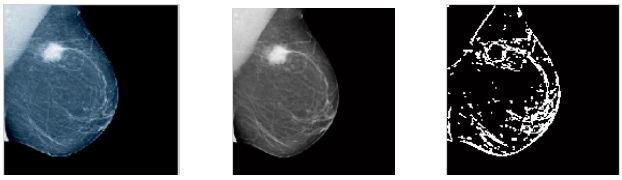


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

**Figure 4.** Segmentation of mammogram (a) Input Image (b) Iterated output (c) Global Segmentation (d) Gradient Magnitude (e) Image structuring (f) Image reset (g) Image ordering (h) Image complement (i) Segmented output.



(a) (b) (c)

**Figure 5.** Existing Threshold algorithm (a) Input image (b) Gray scale image (c) Threshold output image.

## 7. Conclusion

In this algorithm, the mammogram has been segmented using Region based segmentation – Split and merge technique. This algorithm is a fully automated which provides a breast contour segmentation. The algorithm is applied

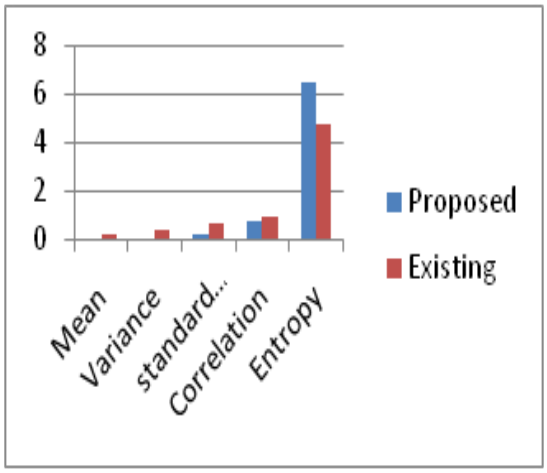



**Figure 6.** Comparison Chart between existing algorithm and proposed algorithm.

to different set of image to provide the segmented output. The advantage of using region growing technique is that in this method we can correctly separate the regions that have the same properties we define. In this method the original images which have clear edges the good segmentation results. We can determine the seed point and the criteria we want to make. The region growing provides results that are sensitive to initial seeds, the accurate seed selection is very important for image segmentation.

## 8. References

1. Engineering C. Review on various techniques of mammogram image segmentation. 2015; 4(5):1456–60.
2. Angayarkanni SP, Kamal DNB, Thavavel DV. Mammogram Mri Image Segmentation and Classification. 2013; 57(2):174–83.
3. Ettabaa KS. A new Mammography segmentation technique based on Watershed, Wavelet and Curvelet transform. 48–51.
4. Bose JSC. Detection of Microcalcification in Mammograms using Soft Computing Techniques. 2012; 86(1):103–22.
5. Subashini TS, Ramalingam V, Palanivel S. Pectoral Muscle removal and Detection of masses in Digital Mammogram using CCL. Int J Comput Appl. 2010; 1(6):71–6.
6. Udaya GG. Efficient Segmentation and Classification of Micro calcification using PSO and SVM in Mammographic Images. 2013; 58–62.
7. Lagzouli M, Elkettani Y. A New Morphology Algorithm for Microcalcifications Detection in Fuzzy Mammograms Images. 2014; 3(1):729–33.
8. Journal I, Engineering OF. A Hybrid Scheme for Mass Detection and Classification in Mammogram. 2014; 3(3):1–3.

9. Sahakyan A. Segmentation of the Breast Region in Digital Mammograms and Detection of Masses. 2012; 3(2):102–5.
10. Saini S, Arora K. A Study Analysis on the Different Image Segmentation. 2014; 4(14):1445–52.
11. Ponraj D, Jenifer M. A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer. J Emerg. 2011; 2(12):656–64.
12. Samir P, Bandyopadhyay K. Survey on Segmentation Methods for Locating Masses in a Mammogram Image. 2010; 9(11):25–8.
13. Khan AM. Image Segmentation Methods: A Comparative Study. Int J Soft Comput Eng. 2013; 3(4):84–92.
14. Kamdi S, Krishna RK. Image Segmentation and Region Growing Algorithm. Int J Comput Technol Electron Eng. 2011; 2(1):103–7.
15. Melouah A, Amirouche R. Comparative study of automatic seed selection methods for medical image segmentation by region growing technique. Recent Adv Biol Biomed Bioeng. 91–7.
16. Rajkumar K, Raju G. Automated Mammogram Segmentation Using Seed Point Identification and Modified Region Growing Algorithm. Br J Appl Sci Technol. 2015; 6(4):378–85.