Breast cancer detection in Mammographic Images using Watershed and Thresholding Technique

Kavneet Kaur[1], Ms.Navneet Kaur[2], Mr.Rashwinder Singh[3]

[1,2] Department of Computer Engg., Punjabi University Patiala,India.

[3] ECE Department,CGCCOE,Landran,India.

Email- thind_kavneet@yahoo.com

Abstract— Breast Cancer is the utmost usual cancer among the women and is also leading cause of deaths in women all over the world. However, with timely diagnosis its treatment is possible. So, mammography is the most common method for diagnosis through which we can have knowledge about its abnormalities, symptoms and different modalities that are used by the medical professionals to diagnose abnormal conditions. In this paper, a Computer Aided Design (CAD) system is developed that can detect the abnormalities from the input image. Firstly, the image is cropped to extract the ROI from the input image, and then it is pre-processed to enhance the details of ROI and then the masses are extracted from the input images using watershed transform as well as thresholding technique. The database taken is Mammographic Image Analysis Society (MIAS) that is an organization of UK research groups interested in the understanding of mammograms and has generated a database of digital mammograms.

Keywords— CAD,CADx, MIAS, Mammography, Watershed Transform, Dilation, Erosion, and Thresholding.

1. INTRODUCTION

Breast cancer is a potentially fatal disease that is growing in frequency in developed countries [1] and is becoming a major public health problem among women. Early detection of this disease can aid in decreasing the number of patients dying by 20%to 30% [2] because the earlier the detection is made, the better treatment works. It is very common among women while rare among men [1]. Breast cancer most commonly affects women after the age of 40, however younger women can also be affected especially with genetic predisposition that is a genetic characteristic that affects the development of an organism under the influence of environmental conditions[16].Breast cancer is the result of abnormal cells that spread beyond the ducts or lobules, invading the surrounding tissue and lymph nodes or blood stream. Diagnoses can be done by several types of biopsy: fine needle aspiration cytology (with sensitivity of90% - 95%); excision biopsy; frozen section biopsy; and by ultrasound; and mammography [3].

Mammography is the widely used technique for detection and characterization of breast cancer. It uses an X-ray system of a low-dose to look inside the breasts. Even small tumors and micro calcifications can be detected using mammograms [5]. By providing an independent second opinion, Computer Aided Detection (CAD) or Computer Aided diagnosis (CADx) systems [4] could help radiologists in the early detection of breast cancer. On the other hand, it is important for future studies to distinguish between computer-aided detection and computer-aided diagnosis systems[17]. The latter CAD system could help radiologists to classify the abnormalities as benign or malignant, which would provide specificity [5]. There are basic viewing and image enhancement systems that provide simple tools such as the ability to zoom in on a digital mammogram image, inverting between black and white on the image and increase/decrease the grey shades; so that a radiologist could have a clearer look of certain areas, or regions, of interest[18].

2. DATABASE RESOURCES

The source of the mammograms used in this work is the MIAS database [11]. The Mammography Image Analysis Society (MIAS) is an organization of UK research groups interested in the understanding of mammograms who have produced a digital mammography database for research purposes[19]. The X-ray films in the database have been carefully selected from the United Kingdom National Breast Screening Programme and digitized with a Joyce-Lobel scanning microdensitometer to a resolution of 200 µm × 200 µm, a device linear in the optical density range 0-3.2 and representing each pixel with an 8-bit word. Every image is 1024 X 1024 pixels in size. The database contains left and right breast images for 161 patients. Its quantity consists of 322 images, which belong to three classes such as normal, benign and malignant. There are 208 normal, 63 benign and 51 malignant (cancerous) images [6].

3. PROPOSED METHOD

This section presents the techniques that are used to detect the mass lesions in the mammogram images and the process flow is shown as in Figure1.

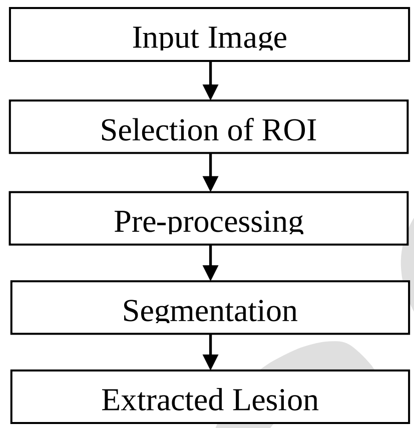

Figure 1: Block Diagram of CAD System

3.1 DETECTION OF REGION INTEREST

The original image size in MIAS database is 1024 × 1024 pixels and in this work, using the locations of abnormalities supplied by the MIAS for each mammogram, a size of 650 × 550 pixels is extracted. A *region of interest* is a portion of an image that you want to filter or perform some other operation on[6]. You define an ROI by creating a *binary mask*, which is a binary image that is the same size as the image you want to process. In the mask image, the pixels that define the ROI are set to 1 and all other pixels set to 0.You can define more than one ROI in an image[15]. The regions can be geographic in nature, such as polygons that encompass contiguous pixels, or defined by a range of intensities. In the latter case, the pixels are not necessarily contiguous. Extraction of the required ROI reduces the calculation overhead and so increases the speed.

3.2 PRE-PROCESSING

The purpose of pre-processing is to improve the quality of the image being processed. There are reasons for the need of image pre-processing:

- improvement of image quality to meet the requirements of physician
- noise reduction
- contrast enhancement
- correction of missing or wrong pixel values
- elimination of acquisition-specific artifacts

Here in preprocessing some morphological operations are performed which are as follows:

3.2.1 DILATION

Dilation is an operation that "grows" or "thickens" objects in a binary image. The specific manner and extent of this thickening is controlled by a shape referred to as structuring element [1]. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller .Suppose A is a 11x11 matrix and B is a 3x3 matrix[3][4].

0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	1	1	1	0
0	1	1	1	1	0	0	1	1	1	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	0
0	1	1	0	0	0	1	1	1	1	0
0	1	1	0	0	0	1	1	1	1	0
0	1	1	0	0	0	1	1	1	1	0
0	1	1	1	1	1	1	1	0	0	0
0	1	1	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Figure 2: A (11x11 matrix)

For each pixel in A, superimpose the centre of B. Each pixel of every superimposed B is included in the dilation of A by B. Dilation can be performed using toolbox function imdilate[3][4].

1	1	1
1	1	1
1	1	1

Figure 3: B (3x3 matrix)

The dilation of A by B is given by C that is 11x11 matrix.

1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	0	0
1	1	1	1	1	1	1	1	1	0	0

Figure 4: C (11x11 matrix)

3.2.2 EROSION

Erosion "shrinks" or "thins" objects in a binary image. As in dilation, the manner and extent of shrinking is controlled by structuring element [1]. The basic effect of the operator on a binary image is to erode away the boundaries of regions of foreground pixels (*i.e.* white pixels, typically).Suppose A is a 13x13 matrix and B is a 3x3 matrix[5][7].

1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	0	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 5: A (13x13 matrix)

Thus areas of foreground pixels shrink in size, and holes within those areas become larger. Erosion can be performed using toolbox function imerode[5][7].

1	1	1
1	1	1
1	1	1

Figure 6: B (3x3 matrix)

The erosion of A by B is given by C that is 13x13 matrix.

0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	0	1	1	1	1	0
0	1	1	1	1	0	0	0	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 7: C (13x13 matrix)

3.3 SEGMENTATION

Segmentation refers to the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify the representation of an image into something that is more meaningful and easier to analyze[19]. Image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image[22]. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as colour, intensity, or texture[21].

3.3.1 THRESHOLDING

In many vision applications, it is useful to be able to separate out the regions of the image corresponding to objects in which we are interested, from the regions of the image that correspond to background. Thresholding often provides an easy and convenient way to perform this segmentation on the basis of the different intensities or colours in the foreground and background regions of an image

[20].In addition, it is often useful to be able to see what areas of an image consist of pixels whose values lie within a specified range, or band of intensities (or colours). Thresholding can be used for this as well. The input to a thresholding operation is typically a gray scale or colour image. In the simplest implementation, the output is a binary image representing the segmentation. Black pixels correspond to background and white pixels correspond to foreground (or *vice versa*). In simple implementations, the segmentation is determined by a single parameter known as the *intensity threshold*. In a single pass, each pixel in the image is compared with this threshold. If the pixel's intensity is higher than the threshold, the pixel is set to, say, white in the output. If it is less than the threshold, it is set to black[8].

In more sophisticated implementations, multiple thresholds can be specified, so that a *band* of intensity values can be set to white while everything else is set to black. For colour or multi-spectral images, it may be possible to set different thresholds for each colour channel, and so select just those pixels within a specified cuboid in RGB space. Another common variant is to set to black all those pixels corresponding to background, but leave foreground pixels at their original colour/intensity (as opposed to forcing them to white), so that that information is not lost[1][2]..

3.3.2 WATERSHED SEGMENTATION

The watershed transform is the method of image segmentation in the field of mathematical morphology. Image segmentation is the process of isolating objects in the image from the background, i.e., partitioning the image into disjoint regions such that each region is homogeneous with respect to some property such as gray value or texture [15].The watershed transform can be classified as a region based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is flooded by water, watersheds being divide lines of the domains of attraction of rain falling over the region. Direct application of watershed transform to a gradient image can result in over segmentation due to noise[11]. Over segmentation means a large number of segmented regions. An approach used to control over segmentation is based on the concept of markers.

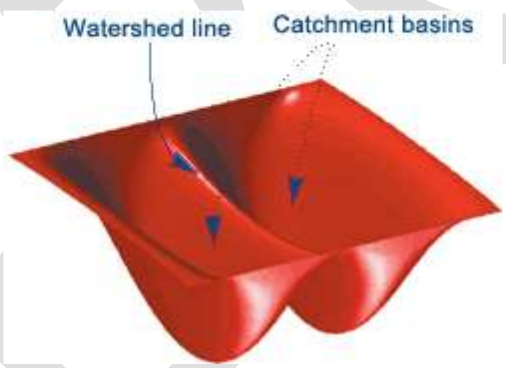

Figure 8: General Concept of Watershed Algorithm

A marker is a connected component belonging to an image. Markers are used to modify the gradient image. Markers are of two types internal and external, internal for object and external for boundary[7]. The marker-controlled watershed segmentation has been shown to be a robust and flexible method for segmentation of objects with closed contours, where the boundaries are expressed as ridges. Markers are placed inside an object of interest; internal markers associate with objects of interest, and external markers associate with the background[13]. After segmentation, the boundaries of the watershed regions are arranged on the desired ridges, thus separating each object from its neighbors [9][11][12]

ISSN 2091-2730

4. EXPERIMENTAL RESULTS

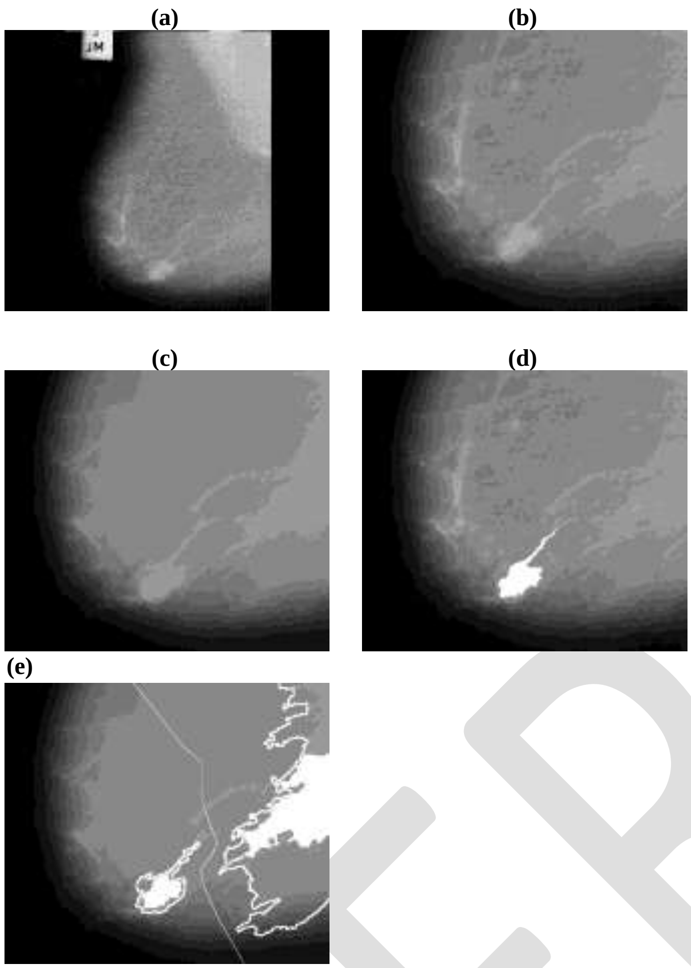

Figure 9: (a) Original Image (mdb005) taken from MIAS database (1024 × 1024 pixels). (b) Cropped Image (650 × 550 pixels). (c) Pre-processed image after Morphological operations. (d) Thresholding Output Image. (e) Final result of Marker based Watershed Algorithm

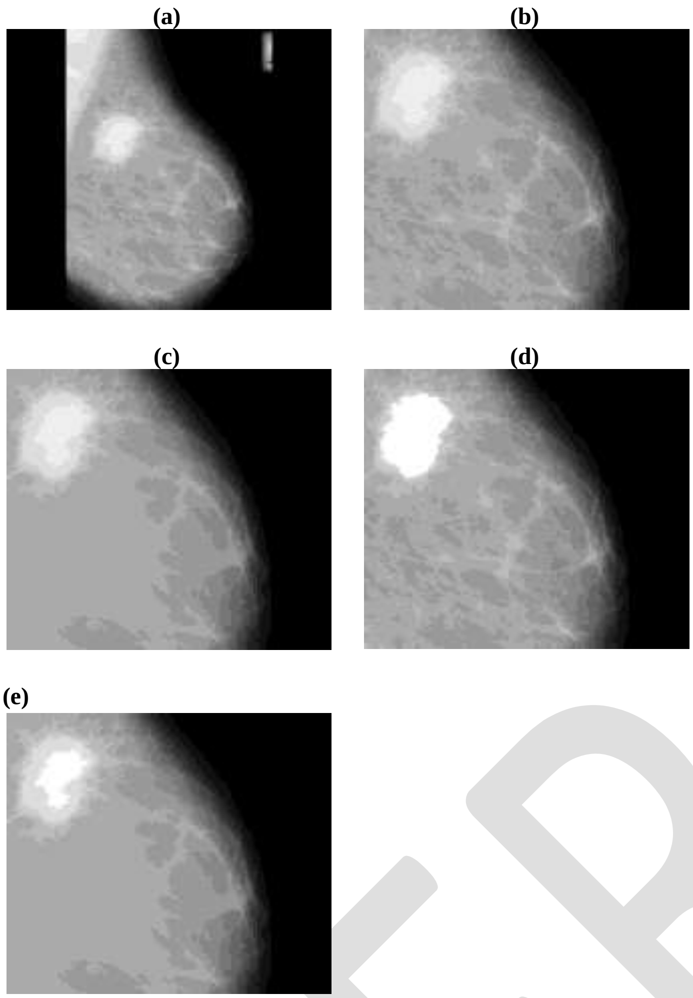

Figure 10: (a) Original Image (mdb184) taken from MIAS database (1024 × 1024 pixels). (b) Cropped Image (650 × 550 pixels). (c) Pre-processed image after Morphological operations. (d) Thresholding Output Image. (e) Final result of Marker based Watershed Algorithm

5. CONCLUSION

In this paper, a novel method for detecting the mass lesions in the mammogram images is presented. The proposed method is designed using three main stages, detection region of interest, pre-processing, and segmentation. Here, thresholding technique gives optimal result but the only issue with this technique is the selection of threshold value. Selection of correct threshold value is utmost important for optimal image segmentation. While watershed transform is an efficient segmentation tool with marker based approach as it automatically detects the mass in the image. Also it draws ridge lines if more than one region is detected in an image as shown in the above results. These proposed methods are successfully able to segment the image effectively and mass lesions were clearly visible.

REFERENCES:

[1] R. C. Gonzalez and R. E. Woods , " Digital Image Processing" Third EditionErika J. Schneble1, Lindsey J. Graham1, "Current Approaches and Challenges in Early Detection of Breast Cancer Recurrence" Journal of Cancer 2014, Vol. 5 5(4): 281-290

[2] A.K.Jain, "Fundamentals of Digital Image Processing", Prentice-Hall Inc,Englewood Cliffs,1989

[3] Aarti Abasaheb Sawant1, Prof. V. N. Kshirsagar, "Image Analysis Based On Segmentation Algorithms" International Journal of Emerging Technology and Advanced EngineeringVolume 3, Issue 11, November 2013. ISSN 2250-2459

[4] Ayman AbuBaker," Mass Lesion detection using Wavelet transform and support Vector Machine", International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 2, April 2012.

[5] Amandeep Kaur, Aayushi,"Image Segmentation Using Watershed Transform", International Journal of Soft Computing and Engineering, Volume-4, Issue-1, March 2014. ISSN: 2231-2307

[6] Ammari H., "Biomedical Imaging Modalities" An Introduction to Mathematiccs and Emerging Biomedical Imaging 2008, 198p. ISBN: 978-3-540-79552-0

[7] Anuj Kumar Singhand Bhupendra Gupta, Department of Computer Science & Engineering,"A Novel Approach for Breast Cancer Detection and Segmentation in a Mammogram" 2015 (IMCIP-2015)Procedia Computer Science 54 (2015) 676 – 682

[8] Rangaraj M. Rangayyana, Fabio J. Ayresa, J.E. Leo Desautels, Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, Canada, " A review of computer aided diagnosis of breastcancer: Toward the detection of subtle signs"Journal of the Franklin Institute 344 (2007) 312–348

[9] Neslihan Ozmen, Robin Dapp, Michael Zapf, Hartmut Gemmeke, Nicole V. Ruiter, Member, IEEE,and Koen W. A. van Dongen, "Comparing Different Ultrasound ImagingMethods for Breast Cancer Detection"IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 62, no. 4, April 2015

[10] Leonard Fassa, Imperial College Department of Bioengineering, London, UK, "Imaging and cancer: A review" Molecular Oncology (2008) 115–152.

[11] Jose Luis Vercher-Conejero, Laura Pelegrí-Martinez, Diego Lopez-Aznar and María del PuigCózar-Santiago, "Positron Emission Tomography in Breast Cancer"Diagnostics 2015, 5, 61-83.

[12] Gheonea Ioana Andreea, Raluca Pegza, Luana Lascu,Simona Bondari, Zoia Stoica, A. Bondari, "The Role Of Imaging Techniques In Diagnosis Of Breast Cancer" Current Health Sciences Journal Vol. 37, No. 2, 201.

[13] P. Shanmugavadivu, S.G. Lakshmi Narayanan, Department of Computer Science and Applications, "Segmentation of Microcalcifications in MammogramImages using Intensity Directed Region Growing" 2013 International Conference on Computer Communication and Informatics (ICCCI -2013), Jan. 04 – 06, 2013, Coimbatore, INDIA

[14] Sachin Prasad,Dana Houserkova Department of Radiology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic, "The Role of various modalities in Breast Imaging" Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2007, 151(2):209–218.

[15] M. S. Islam , Grand Forks, N. Kaabouch ,W. C. Hu, "A survey of medical imaging techniques used for breast cancer detection" Electro/Information Technology (EIT),IEEE International Conference,2013.

[16] Barbour S. Warren, Ph.D. Research Associate,Carol Devine, Ph.D., R.D.Division of Nutritional Sciences and Education Project Leader, "Breast Cancer in Men" Cornell University Program on Breast Cancer and Environmental Risk Factors in New York State (BCERF)2001

[17] https://en.wikipedia.org/wiki/Erosion_(morphology)

[18] https://en.wikipedia.org/wiki/Dilation_(morphology)

[19] http://homepages.inf.ed.ac.uk/rbf/HIPR2/threshld.html

[20] http://peipa.essex.ac.uk/info/mias.html

[21] http://homepages.inf.ed.ac.uk/rbf/HIPR2/dilate.html

[22] http://homepages.inf.ed.ac.uk/rbf/HIPR2/erode.html