IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol-6, Special Issue-4, May 2018 E-ISSN: 2347-2693

Breast Cancer Detection in Digital Mammograms using Histogram Bins based Otsu Thresholding

P. Shanmugavadivu[1*], A. Thilshat Barveen[2], Ashish Kumar[3]

[1,2,3] Dept. of Computer Science and Applications, Gandhigram Rural Institute (Deemed to be University), Dindigul, India

[*]*Corresponding Author: psvadivu67@gmail.com, Tel.: +91-94437-36780*

Available online at: www.ijcseonline.org

***Abstract*—** Breast cancer is reported as the second most dangerous diseases in the world. Early detection of breast cancer is recorded to reduce the mortality rate by 30 to 60 percent. Digital mammography is a widely accepted as a non-invasive modality for the early detection of breast cancer, based on the abnormalities in the form of lesions, tumours and micro calcifications. The computer-based breast cancer screening involves segmentation of the abnormalities which are characterized by abrupt change in the pixel intensity, against the neighbourhood pixels. This paper presents a computationally simple and efficient enhancement technique that uses the principle of Otsu thresholding. The thresholds value for segmentation is chosen from the histogram bins of the input mammogram. The insignificant segmented partitions are discarded using morphological operations. The proposed method based on Histogram Bins based Otsu Thresholding (HBOT) is proved to segment the suspicious region accurately, as evidenced in the visual perception.

***Keywords*—** Breast Cancer Detection ,Microcalcifications ,Mammogram Segmentation, Histogram Bins ,Otsu Thresholding

I. INTRODUCTION

The breast cancer has become a jutting threat, leading to middy temporality. Early detection of breast cancer is best solution to deal with this issue. Digital mammography is the commonly used safeguard against breast cancer due to its simplicity, transferable and cost efficacy. To achieve, higher precision rates in the detection of breast cancer, various decision support system have been designed in the current years. This paper presents a brief review on the recent research reported on the image segmentation and feature extraction knack in the domain of digital mammogram analysis for artifacts and presents segmentation technique that works in two phase namely global threshold-based enhancement [1] and an intensity-directed region-growing mechanism for segmentation [2]. Digital Mammograms have emerged as the best means for prognosis and diagnosis of breast cancer. Research studies reveal that the process of mammogram analysis can be accelerated by the intervention of computers [13].

II. REVIEW OF LITERATURE

Shanmugavadivu and Lakshmi Narayanan [2013] proposed a mass segmentation strategy in which the seed pixel selection is guided by automatic selection of threshold. *MoumenaAl-Bayati et al.* [2013] applied valley emphasis method, Neighborhood valley emphasis method, thresholding based on variance intensity contrast and Variance Discrepancy for mammogram image segmentation. The method proposed by them is an extension of Otsu method which is proved to be the most effective method to carry out segmentation, with respect to shape measures and uniformity.

Ireaneus Anna Rejani et al. [2009] presented an algorithm with an objective to assist the radiologists for the early detection of breast cancer. This method is a combination of image negative, thresholding and segmentation, using which the abnormalities in mammograms are detected. The algorithm is tested on Mammographic Image Analysis Society (MIAS) dataset. *Zaheeruddin et al.* [2012] devised a mean-based region-growing segmentation (MRGS) method that automatically selects the seed pixel and optimal threshold value. The process of segmentation is reported to be fast and accurate in culling out the breast masses in mammograms.

Hadeel N. Abdullah *et al.* [2015] have carried out feature extraction for medical images using multi-stage process, which includes image pre-processing by median filtering and segmentation by threshold and watershed algorithm. Shanmugavadivu and Lakshmi Narayanan [2013] proposed an enhancement technique in order to distinguish the mass regions from the background, in digital mammograms. The statistical methods in association with neighborhood property, as proposed in the article, are applied to segment the abnormal region. The obtained results are found to comply with ground truth realities.

Vennila et al. [2014] developed a segmentation method, using region-growing and Otsu threshold. The segmentation is preceded by morphological preprocessing to suppress the artifacts and pectoral muscle.*Shanmugavadivu and Sivakumar* [2014] proposed a new approach based on the principle of fractal thresholding which is proved to be efficient in segmenting the masses from background tissues based on their roughness texture.

Jun Liu *et al.* developed a mass segmentation algorithm, using automatic marker controlled watershed transform for the segmentation of mass regions, which is further refined with a level set. The combined mechanism of the watershed based segmentation and level set method pays off in accurate segmentation. An image classifier to classify the mammogram images as normal, benign or malignant developed by M. Vasantha *etal.* [2010], extracts the histogram intensity features and GLCM features of the input mammograms. This hybrid approach for feature selection reduces the redundant features, thereby obtaining optimum features. The classification part of this method is accomplished by a decision tree.

Pradeep N *et al.,* [2012] developed a method for mass / tumor classification based on the texture, statistical and structural features of RoI in digital mammograms. The classification of the images was performed using pattern recognition. A four-step approach to segment and detect breast cancer is proposed and evaluated by HebaAl-Hiary *et al.* This method sequentially performs features extraction and detection on the digital mammogram. This technique is reported with an ability to increase the quality of prognosis.

III. METHODOLOGY

The proposed method reads the input image to compute the histogram with size for 16 and 64 and 256 bins to find the optimal bin size. Image Pre-Processing is used to find the suspicious artifacts and to produce the accurate tumours image. Digital mammogram images contain artifacts in the form of labels, markers and wedges in the background region the major problem with the precise segmentation of the breast region is the existence of such artifacts, which may cause trivial segmentation algorithms to fail. The histogram represents a frequency distribution by statistical information that uses rectangles whose widths represents class intervals of the equal size. Histogram bins used to divide the range of intensity values into a serious of intervals, then counting the how many values fall into the each interval. The bins are usually finding non-Overlapping in the suspicious region. In this paper three different bin size are explored such as 16, 64 and 256 bin size for finding the intensity of interest in a suspicious cancer images. The threshold values are obtained namely T_{min} and T_{max} to segment suspicious cancer cells from the probe image. The algorithm in this paper uses one of the threshold based segmentation techniques called Otsu's N thresholding. The Otsu method is famous for its simplicity and effectiveness and selection of the optimal threshold automatically and stably. Thresholding is used to extract the portions of the probe image whose pixel intensity value is greater than the given threshold value. Then the resultant image is subjected to the segmentation process [5]. This technique succeeded in extraction of an object from complex background in an image on attempt to achieve smaller intra class variance, and higher inter class variance [10]. Binarization is a process to convert the all pixels of the probe image into either pure white (1) or pure black (0) based on the principle adapted for the segmentation. Here pure white (1) indicates the fore ground pixel (Suspected cancer cells) and pure black (0) indicates the background. Morphological operation are used to remove the unwanted artifacts in the probe image. Here close and erode operation are used to close the small openings and removes the unwanted artifacts respectively. Resultant foreground is considered as segmented RoI.

The objective of proposed method is achieved in two phases as described below.

Phase 1: Algorithm to find Optimum Threshold Values

Input:	Pre-processed Digital Mammograms
Output:	Threshold Values (T1, T2)

Step 1: Read input image
Step 2: Compute the Histogram of MI
Step 3: Generate Histogram with bin size 16 and 64 and 256
Step 4: Find the Optimal Bins
Step 5: Find Bin with Intensity of Interest in (MI)
Step 6: Find threshold values T1 and T2.

In Phase I , three bin size 16, 64 and 256 bins are selected for gray scale images and the respective histograms are generated for whole database. Based on the ground truth image (GT images) of mammograms it was found that in histogram of bin size 16 , bin no 12 and 13 contains most of the suspected cancerous cells. Based on further intensity analysis the minimum intensity threshold T1 and maximum intensity T2 for highest cancerous cell probability region is obtained. These threshold values T1 and T2 works as input argument for phase II algorithm. (Fig 1).

Phase 2: Algorithm for Segmentation of ROI

Input:	Threshold Values (T1, T2) of MI image.
Output:	Extraction for Region of Interest

Step 1: Apply Otsu thresholding on MI using T1, T2
Step 2: Generate the binary image
Step 3: Remove the artifacts using morphological operation
Step 4: Output Segmented Region of Interest.

In phase II, Binarization of probe image if achieved through Otsu thresholding using T1 and T2 thresholds. All unwanted artifacts are removed using morphological operation such as

close and erode. The Resultant morphological operation yield segmented ROI (Fig 2).

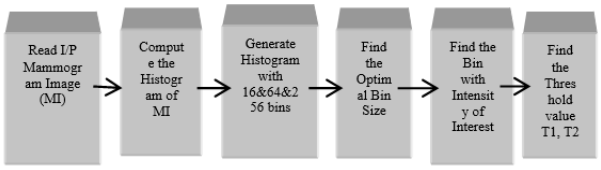


Fig.1: **Phase 1: Computation of Optimum Threshold Values**

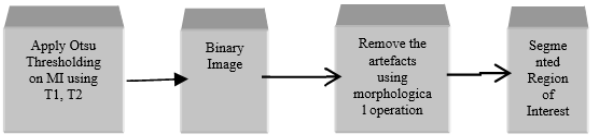
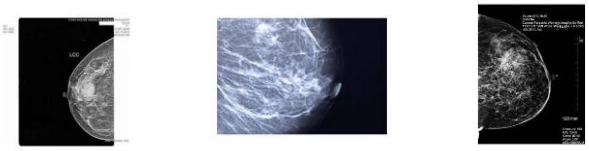


Fig.2: **Phase 2: Segmentation of RoI**

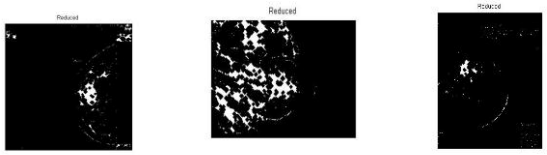
IV . RESULTS AND DISCUSSION

The algorithm presented in this paper is developed in MATLAB 6.5. The performance of the algorithm presented in this paper is evaluated on more than twenty sample images obtained from Mammographic Image Analysis Society (MIAS) database. The affected mammogram images are process to produce the suspected cancerous cells detection in breast cancer. The obtained results are qualitatively validated by comparing them with the ground reality descriptions provided in the MIAS dataset. For illustrative purpose, the results obtained for three of those images [mamo1, mamo2 and mamo3] are presented here. Figure 3: (a1) to (a3) are original mammogram images. Figure 4: (b1) to (b3) are segmented region using Otsu's thresholding with thresholds T1 and T2, and the Region of Interest (RoI) obtained after morphological operations are shown in Figure 5: (c1) to (c3).



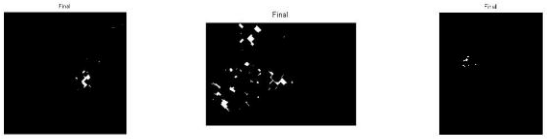
(a1) (a2) (a3)

Fig 3: (a1)-(a3): Original Mammogram Images



(b1) (b2) (b3)

Fig 4: (b1)-(b3): Segmented Images of (a1)-(a3) using Otsu thresholding



(c1) (c2) (c3)

Fig 5: Artifacts removed images of (b1)-(b3) using HBOT

V. CONCLUSION

In this paper, Histogram Bin based Otsu Thresholding (HBOT) based breast cancer detection in mammogram images is presented. The proposed approach HBOT works in two phases. In phase I, optimal threshold values T_1 and T_2 are obtained using histograms of bin size 16. Phase II, uses Otsu N thresholding with thresholds obtained in phase I. Morphological operations are used to eliminate the artifacts. HBOT competes with the state of the art methods for mammogram cancer detection.

REFERENCES

[1] Zaheeruddin,Z.A.Jaffery and Laxman Singh, "Detection and Shape Feature Extraction of Breast Tumor in Mammograms" , Proceedings of the World Congress on Engineering 2012 Volume 2 WCE 2012, July 4 - 6,pp 231-236,2012.

[2] .P.Shanmugavadivu, Lakshmi Narayanan, S.G., "Detection of Microcalcifications in Mammogram using Statistical Measures based Region Growing", SPIE, Proceedings of the International Conference on Communication and Electronics System Design ICESD 2013,vol. 8760, p.no.87601M-1 -6, 2013.

[3] P.Shanmugavadivu, P., Lakshmi Narayanan, S.G., "Segmentation of Microcalcifications in Mammogram Images using Intensity - Directed Region Growing", Proceedings of the International Conference on Computer Communication and Informatics ICCCI 2013.

[4] Shanmugavadivu, P., Lakshmi Narayanan, S.G.: Segmentation of Micro calcification Regions in Digital Mammograms using Self-Guided Region Growing. In: Proceedings of the International Conference on Emerging Trends in Science Engineering and Technology INCOSET 2012, pp. 274–279 (2012).

[5] Y. Ireaneus Anna Rejani and Dr.S.Thamarai Selvi, "Breast Cancer Detection Using Multilevel Thresholding", International Journal of Computer Science and Information Security, Volume. 6, No.1, 2009.

[6] Snehal A. Mane and Dr. K. V. Kulhalli, " Mammogram Image Features Extraction and Classification for Breast Cancer Detection", International Research Journal of Engineering and Technology, e-ISSN: 2395 -0056 Volume: 02 Issue: 07, P-ISSN: 2395-0072, Oct 2015.

[7] P.Shanmugavadivu, V.Sivakumar, "Segmentation of Masses in Digital Mammograms using Fractal-Bound Computing Technique for Breast Cancer Prognosis", International Journal of Applied Engineering Research, ISSN 0973-4562, pp. 23187 – 23192, Vol. 10 No.31, 2015.

[8] P.Shanmugavadivu, V.Sivakumar, (2013) "Segmentation of pectoral muscle in mammograms using Fractal method", ICCCI-2013, IEEE Xplore, ISBN: 978-1-4673-2906-4, pp. 1-6.

[9] Pradeep N, Girisha H, Sreepathi B and Karibasappa K, "Feature Extraction Of Mammograms", International Journal of Bioinformatics Research, ISSN: 0975–3087 & E-ISSN: 0975–9115, Volume 4, Issue 1, pp.-241-244, 2012.

[10] Moumena Al-Bayati and Ali El-Zaart, "Mammogram Images Thresholding for Breast Cancer Detection Using Different Thresholding Methods", Advances in Breast Cancer Research, 2, 72-77, 2013.

[11] M.Vasantha,Dr.V.Subbiah Bharathi and R.Dhamodharan, "Medical Image Feature, Extraction, Selection and Classification", International Journal of Engineering Science and Technology, Volume 2(6), 2071-2076, 2010.

[12] Anu Appukuttan and Sindhu L., "Breast Cancer-Early Detection and Classification Techniques: A Survey", International Journal of Computer Applications (0975 – 8887) Volume 132 – No.11, December2015.

[13] K.Vennila, k.Sivakami and R.Padmapriya, "Detection of Mass in Digital Mammograms", International Journal of Computer Applications (0975 – 8887) Volume 104 – No.5, October 2014.

[14] Inam ul Islam Wani, M. C Hanumantharaju and M. T Gopalakrishna ,"Review of Mammogram Enhancement Techniques for Detecting Breast Cancer", International Journal of Computer Applications (0975 – 8887), International Conference on Information and Communication Technologies – 2014.

[15] Belal K. Elfarra and Ibrahim S. I. Abuhaiba, "New Feature Extraction Method for Mammogram Computer Aided Diagnosis", International Journal of Signal Processing, Image Processing and Pattern Recognition Volume 6, No. 1, February, 2013.

Authors Profile

Dr. P. Shanmugavadivu pursed Bachelor of Science from K.N. Govt. Arts College, Thanjavur, 1987.She pursed Master of Computer Science and Applications from Regional Engineering College, Trichy, 1990 and Ph.D from The Gandhigram Rural Institute, India, in year 2008. She pursedMaster of Business Administration from IGNOU, 2016. She is currently working as Professor in Department of Computer Science and Applications, Gandhigram Rural Institute, Dindigul, Tamil Nadu, India. She has published more than 22 research papers in reputed international journals and 31 research papers in edited volumes. Her main research work focuses on Medical Image Analysis, Image Restoration, Image Enhancement, Image Segmentation, Content – Based Image Retrieval. She has 24 years of teaching experience and 15 years of Research Experience.

A.Thilshat Barveen pursed Bachelor of Science from N.S Arts and Science College, Theni, 2015.She pursed Master of Information Technology from The Gandhigram Rural Institute, 2017. She is currently pursuing M.Phil. in Department of Computer Science and Applications, Gandhigram University, India in 2018. Her research work focuses on Medical Image Processing.

Mr. Ashish Kumar pursed Bachelor of Science from Chaudhary Charan Singh University, Meerut, India in 2001 and Master of Computer Applications from The Gandhigram Rural Institute – Deemed to be University, India in year 2011. He is currently pursuing Ph.D. and working as Project Fellow in Department of Computer Science and Applications, The Gandhigram Rural Institute – Deemed to be University, India since 2012. He has published several research papers in reputed conferences including IEEE, Springer Series and it's also available online. His main research work focuses on Digital Image Processing, Artificial Intelligence, Machine Learning and Medical Imaging. He has 3 years of teaching experience and 5 years of Research Experience.

