Digital Mammogram Segmentation and Feature Extraction: A Review

P.Shanmugavadivu[1], A.Krishnaveni[2], V. Sivakumar[3], S.G. Lakshmi Narayanan[4]
[1,2,4]Department of Computer Science and Applications,
Gandhigram Rural Institute - Deemed University, Gandhigram.- 624 302
Tamil Nadu, India
[3]School of Computing, Faculty of Computing and Engineering Technology,
Asia Pacific University of Technology and Innovation (APU), Malaysia
[1]psvadivu67@gmail.com, [2]krishnaveniayyavu@gmail.com, [3]sivakumar.vengusamy@gmail.com
[4]sg.lakshminarayanan@gmail.com

***Abstract*—Breast cancer has become a prominent threat, leading to woman mortality. Early detection is the best solution to combat this mortality. Digital mammography is the most reliable technique for screening breast cancer due to its simplicity, portability and cost effectiveness. To achieve, higher accuracy rates in the detection of breast cancer, various clinical decision support systems have been designed in the recent years. These computer assisted diagnostic systems consist of four sequential stages namely preprocessing, segmentation, feature extraction and classification. This article is aimed at presenting a review on the recent literature reported on the image segmentation and feature extraction techniques in the domain of digital mammogram analysis.**

***Keywords*-** Mammogram, segmentation, feature extraction, review, analysis

I. INTRODUCTION

Breast cancer has become a prominent threat, leading to woman mortality. Early detection is the best solution to combat this mortality. Digital mammography is the most reliable technique for screening breast cancer due to its simplicity, portability and cost effectiveness. To achieve, higher accuracy rates in the detection of breast cancer, various clinical decision support systems have been designed in the recent years. These computer assisted diagnostic systems consist of four sequential stages namely preprocessing, segmentation, feature extraction and classification. This article is aimed at presenting a review on the recent literature reported on the image segmentation and feature extraction techniques in the domain of digital mammogram analysis.

II. SEGMENTATION / FEATURE EXTRACTION TECHNIQUES

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.

Shanmugavadivu and Lakshmi Narayanan [2013] proposed a mass segmentation strategy in which the seed pixel selection is guided by automatic selection of threshold. *Moumena Al-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.

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.

Another novel abnormal mass segmentation method for digital mammograms devised by *Shanmugavadivu and Sivakumar* [2014], combines the potential of Fractal Hurst based modified Sobel edge detection with fractal dimension. This method could trace the contour of the masses effectively than the ones identified by traditional Sobel method. *Shrasthta Chauhan et al.* [2014], compared the existing methods for automatic detection of brain tumour through Magnetic Resonance Image (MRI) using Histogram Thresholding and Artificial Neural Network

(ANN) and summarized the performance of the proposed methods. These methods find their application in the detection of contour and geometrical dimension of the masses. A modified ANN by the authors uses quantization for image analysis.

Hadeel N. Abdullah *et al.* [2015] have carried out feature extraction for medical images using multi-stage process, which includes image preprocessing 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 neighbourhood property, as proposed in the article, are applied to segment the abnormal region. The obtained results are found to comply with ground truth realities.

Anuj Kumar Singh *et al.* [2015] implemented a computer aided scheme for the detection and segmentation of mass region in digital mammograms using the primitive image processing techniques, namely averaging and thresholding, along with Max-Mean and Least-Variance techniques. Inam ul Islam Wani *et al.* [2014] reviewed the digital mammogram enhancement techniques, to manipulate mammogram images, which are divided into four main categories viz., traditional, region-based, feature-based and fuzzy-based enhancement techniques. The mechanism of each of these methods is dealt in detail in comparison with a few competitive methods.

Nassir Salman *et al.* [2006] formulated watershed and K-means clustering based segmentation technique, which comprises of new merging procedures with mean intensity that are applied to extract the boundary and segment the respective regions. The edge strength technique is employed to accurately trace the edges of images. Lamia Jaafar Belaid *et al.* [2009] developed mathematical morphology based segmentation using watershed transformation. The problem of over-segmentation is addressed by using topological gradient approach.

Jobin Christ *et al.* [2011] suggested a technique that combines K-means clustering and marker controlled watershed algorithm with Fuzzy C-means clustering to produce segmentation maps which have only fewer partitions. The integration of K-means clustering with the marker controlled watershed algorithm, thereby producing a segmentation map suitable for the several anatomies in the medical images and reduces the amount of over segmentation as well.

Achariya *et al.* [2013] developed Distance Transform based Watershed algorithm to precisely segment the regions of interest in a mammogram. The problem of over-segmentation is corrected by combining the features of watershed algorithm with Laplacian of Gaussian (LoG) edge detection. Niket Amoda *et al.* [2013] used the Region Based Image Retrieval (RBIR) system in combination with Discrete Wavelet Transform (DWT) and Watershed Segmentation, in order to achieve mass segmentation. The concept of Texture Gradient is implemented using Non-Decimated Wavelet Packet Transform and further marker location algorithm is used to detect the significant homogeneous textured or non textured regions. A marker driven Watershed Transform is finally applied to objectively segment the target regions.

Sura Ramzi Shareef *et al.* [2014] explored the diagnostic approaches for breast cancer detection in the ultrasound and X-ray mammography medical images. This method of segmentation uses morphological watershed transform to extract the watershed lines from the topographic representation of input image. Varsha J. Gaikwad *et al.* [2015] demonstrated that Marker Controlled Watershed algorithm exhibits improved segmentation on the homogeneous regions in the digital mammograms. It is proved to be better than region growing methods in view of regions segmented. This method suggests solution for over-segmentation, which is a typical disadvantage of watershed algorithm.

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 *et al.* [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 Heba Al-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.

A technique for preprocessing, segmentation, feature extraction and classification was designed by Meenalosini. S *et al.,* [2012] for the removal of artifacts and pectoral region. The segmentation was performed using Sobel operator, on the contrast enhanced image. The statistical and textural feature extraction using GLCM and local binary pattern method were used by support vector machine.

Belal K. Elfarra *et al.,* [2013] suggested a new method for feature extraction using both shape bound features and computational features to implement Square Centroid Lines Gray level distribution Method (SCLGM). The

Receiver Operating Characteristics (RoC) and Confusion Matrix were used to measure the functional performance of the method.

Ajay Kumar Bansal *et al.* [2013] introduced Content-based image retrieval technique to extract the images on the basis of intensity characteristics namely texture, color, shape and spatial layout. Despite the semantic gap between the user's concepts and low-level features, this method increases the computational complexity. The intensity characteristics such as statistical, texture, color and shape features are accurately extracted using fuzzy logic and neural network.

R. Vanithamani *et al.* [2015] developed a threshold based NeighShrink method for removing the speckle noise for the images in wavelet domain. The topographical watershed segmentation was used and the image features based on shape, texture, gradient and histogram are extracted. The AdaBoost classifier in association with binary logistic classifier is confirmed to minimize the percentage of false positive and false negative classifications.

Snehal A. Mane *et al.* [2015] proposed a classification method for digital mammograms which comprises of three major stages: preprocessing, feature extraction and classification. The first stage involved image enhancement using Wiener filter and histogram based contrast enhancement. The segmentation is obtained using active contour method, while the Gabor Wavelet features are extracted for the purpose of classification.

Neelima Bagri *et al.* [2015] investigated the texture and shape features that are obtained by GLCM and Hu-moments, in combination with tamura texture and shape invariant Hu-moments. This method could achieve high precision and recall, that vouches the merit of the performance of this technique. P. K. Saranya *et al.* [2016] analyzed and suggested thirty three quantitative morphological features and thirty eight textural features of breast masses for prediction and classification of masses either into benign or malignant.

Namita Aggarwal *et al.* [2012] investigated the first and second order statistics of images using characterized textures. Those features are extracted based on statistics of texture. A query-by-example approach used by Peter Howarth *et al.* [2004] extensively uses the three radically different texture features namely statistical, psychological and signal processing. Anupa Maria Sabu *et al.* [2012] devised a technique for feature extraction using the spatial variations of the image, in which the texture analysis was adapted to characterize the spatial variation within the image by extracting information.

S. Deepa *et al.* [2013] presented a technique using the discriminatory capability of contourlet coefficient co-occurrence matrix features for the analysis and classification of mammogram images. Contourlet transform proves its potential in the analysis of mammograms by representing smooth contours and fine geometrical structures. The RoI from the original image is enhanced using histogram equalization and decomposed using contourlet transform. Further, the co-occurrence matrices are generated for four different directions as θ=0°, 45°, 90° and 135° and the distance measure as d=1 pixel. For each co-occurrence matrix, various second order statistical texture features are extracted.

III. CONCLUSION

It is observed from this review that thresholding, watershed segmentation, K-means, neural network and FCM clustering are widely used by all the researchers for the detection phase in digital mammogram analysis. Additionally, region-growing and fractal based algorithms have been also reported for the precise segmentation of masses. In addition to the statistical features and morphology bound texture features, recently fractal dimension bound techniques are also being used in the segmentation / feature extraction of masses / microcalcifications from digital mammograms. Many of the reported techniques form a computer-assisted diagnostic tool that comprises of pre-processing, enhancement, segmentation, feature extraction and classification phases for digital mammogram analysis.

REFERENCES

[1] 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.

[2] 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, 2012.

[3] 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.

[4] 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.

[5] Shrasthta Chauhan and Mrs. Neha Shrma, "Brain Tumor Detection and Segmentation Using Histogram Thresholding and Artificial Neural Network Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 8, August 2014.

[6] Hadeel N. Abdullah and Mustafa A. Habtr, "Brain Tumor Extraction in MRI Using Hybrid Threshold and Watershed Technique", International Journal of Engineering Science and Innovative Technology, Volume 4, Issue 1, January 2015.

[7] Anuj Kumar Singh and Bhupendra Gupta, "A Novel Approach for Breast Cancer Detection and Segmentation in a Mammogram", Procedia Computer Science 54, 676 – 682, 2015.
[8] 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.
[9] Nassir Salman, "Image Segmentation Based on Watershed and Edge Detection Techniques", The International Arab Journal of Information Technology, Volume. 3, No. 2, April 2006.
[10] Lamia Jaafar Belaid and Walid Mourou, "Image Segmentation: A Watershed Transformation Algorithm", Image Anal Stereol 28, 93-102, 2009.
[11] M.C. Jobin Christ and R.M.S. Parvathi, "Segmentation of Medical Image using Clustering and Watershed Algorithms", American Journal of Applied Sciences 8 (12): 1349-1352, ISSN 1546-9239, 2011.
[12] P.P.Acharjya, A. Sinha, S.Sarkar, S.Dey and S.Ghosh, "A New Approach Of Watershed Algorithm Using Distance Transform Applied To Image Segmentation", International Journal of Innovative Research in Computer and Communication Engineering, Volume 1, Issue 2, April 2013.
[13] Niket Amoda, Ramesh K Kulkarni, " Image Segmentation and Detection using Watershed Transform and Region Based Image Retrieval", International Journal of Emerging Trends & Technology in Computer Science, Volume 2, Issue 2, March – April 2013.
[14] Sura Ramzi Shareef, "Breast Cancer Detection Based on Watershed Transformation", International Journal of Computer Science Issues, Volume. 11, Issue 1, No 1, January 2014.
[15] Varsha J. Gaikwad, "Marker-Controlled Watershed Transform in Digital Mammogram Segmentation", International Journal for Research in Applied Science & Engineering Technology, Volume 3 Issue III, ISSN: 2321-9653, March 2015.
[16] Ryszard S. Chora´s, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems", International Journal of Biology and Biomedical Engineering, Issue 1, Volume 1, 2007.
[17] 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.
[18] P.Shanmugavadivu, V.Sivakumar, "Fractal Dimension-Bound Spatio-Temporal Analysis of Digital Mammograms", Eur. Phys. J. Special Topics © EDP Sciences, Springer-Verlag, DOI: 10.1140/epjst/e2016-02613-0, pp. 137-146, Vol.225, Springer, 2016
[19] 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
[20] 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.
[21] Heba Al-Hiary and Basim Alhadidi, "An Implemented Approach For Potentially Breast Cancer Detection Using Extracted Features And Artificial Neural Networks", Computing and Informatics, Volume 31,225–244, 2012.
[22] Meenalosini S, J. Janet and E. Kannan, "A Novel Approach in Malignancy Detection of Computer Aided Diagnosis", American Journal of Applied Sciences 9 (7): 1020-1029, ISSN 1546-9239, 2012.
[23] 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.
[24] Ajay Kumar Bansal and Swati Mathur, "Feature Extraction in Content Based Image Retrieval: A Review", International Journal of Engineering Research and Applications, ISSN: 2248-9622, International Conference on Advancement in Information Technology, 23 February 2013.
[25] Dong ping Tian, "A Review on Image Feature Extraction and Representation Techniques", International Journal of Multimedia and Ubiquitous Engineering Volume 8, No. 4, July, 2013.
[26] Adegoke, B.O., Olarinoye, A.A, Omotayo, M.E. and Olawuni, A, "Medical image Feature Extraction: A Survey", International Journal of Electronics Communication and Computer Technology, Volume 3 Issue 5 2013.
[27] Gaurav Kumar, Pradeep Kumar Bhatia, "A Detailed Review of Feature Extraction in Image Processing Systems", Fourth International Conference on Advanced Computing & Communication Technologies, DOI: 10.1109/ACCT.2014.74, 2014.
[28] R. Vanithamani and R. Dhivya, "Tumor Detection in Breast Ultrasound images", International Journal of Computer Technology and Applications, 8(5), pp. 1881-1885, 2015.
[29] P.Shanmugavadivu, V.Sivakumar, "A Novel Technique for Mammogram Mass Segmentation Using Fractal Adaptive Thresholding", Lecture Notes in Electrical Engineering, Chapter 25, Vol. 291, pp. 213-220, ISSN No: 1876-1100, DOI 10.1007/978-981-4585-42-2_25, Springer, 2014.
[30] P.Shanmugavadivu, V.Sivakumar, "Intensity-Based Detection of Microcalcification Clusters in Digital Mammograms using Fractal Dimension", Advances in Intelligent Systems and Computing, Vol. 236, pp. 1293-1299, ISSN No: 2194-5357, DOI 10.1007/978-81-322-1602-5_135, Springer, 2014.
[31] 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.
[32] Seyyid Ahmed Medjahed, "A Comparative Study of Feature Extraction Methods in Images Classification", I.J. Image, Graphics and Signal Processing, Volume 3, 16-23, 2015.
[33] 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.
[34] S.Selvarajah and S.R. Kodituwakku, "Analysis and Comparison of Texture Features for Content Based Image Retrieval", International Journal of Latest Trends in Computing (E-ISSN: 2045-5364), Volume 2, Issue 1, March 2011.
[35] Namita Aggarwal and R. K. Agrawal, "First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images", Journal of Signal and Information Processing, Volume 3, 146-153, 2012.
[36] Peter Howarth and Stefan Ruger, "Evaluation of Texture Features for Content-Based Image Retrieval", CIVR 2004, LNCS 3115, pp. 326–334, Springer-Verlag Berlin Heidelberg 2004.
[37] Anupa Maria Sabu, D.Narain Ponraj and Dr.Poongodi, "Textural Features Based Breast Cancer Detection: A Survey", Journal of Emerging Trends in Computing and Information Sciences, Volume 3, NO. 9, ISSN 2079-8407, Sep 2012.
[38] S.Deepa and Dr.V.Subbiah Bharathi, "Textural Feature Extraction and Classification of Mammogram Images using CCCM and PNN", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727, Volume 10, Issue 6, PP 07-13, 2013.
[39] 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, 2013.

[40] Neelima Bagri and Punit Kumar Johari, "A Comparative Study on Feature Extraction using Texture and Shape for Content Based Image Retrieval", International Journal of Advanced Science and Technology,Volume 80, pp.41-52,2015.
[41] Swati Saxena and R.L Yadav, "Hybrid Feature of Tamura Texture Based Image Retrieval System", International Journal of Recent Research and Review, Volume 8, Issue 2, ISSN 2277 – 8322, June 2015.
[42] P. K. Saranya and Dr. E. S. Samundeeswari, "A Study on Morphological and Textural Features for Classifying Breast Lesion inUltrasound Images", International Journal of Innovative Research in Science,Engineering and Technology, Volume 5, Issue 3, March 2016.
[43] 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.

AUTHORS PROFILE

Dr. (Mrs). P. Shanmugavadivu received her M.C.A. degree from Regional Engineering College (now known as National Institute of Technology), Tiruchirapalli, India. She joined the Gandhigram Rural Institute-Deemed University in 1990 and is presently serving as Professor in the Department of Computer Science and Applications. She received her Ph.D. degree on Image Restoration in the year 2008 from the same institution. She has developed several non-linear filters for Image Restoration and has contributed about 200 research articles to International Journals and Conferences at National and International level. She has guided many scholars and 6 doctorates have been awarded under her supervision. Her research areas include Medical Imaging, Image Restoration, Image Enhancement, Image Segmentation, Content Based Image Retrieval and Big Data Analytics. She is a Life Member of the Indian Society for Technical Education and proud recipient of 21st Century Knowledge Initiative Award by UGC. She is serving as reviewer in many reputed journals and delivered several lectures in research forums.

A. Krishnaveni has completed her M.Phil in Computer Applications from Gandhigram Rural Institute – Deemed University. Her M.Phil work is focused on feature extraction of digital mammogram images and she has got two international journal publications and got participation exposure in seminars and conferences.

Dr. V. Sivakumar has completed his Ph.D from Gandhigram Rural Institute – Deemed University on Digital mammogram segmentation using Fractal Techniques. He has received his M.Sc. (Mathematics and Computer Applications) degree in 1998 from Gandhigram Rural Institute – Deemed University, India, Post Graduate Diploma in Applied Operations Research in 2002 from Annamalai University, Chidambaram, India and M.Phil. degree in Computer Science in 2003 from Manonmaniam University, Tirunelveli, India. He has experience as Lecturer in Gandhigram Rural Institute-Deemed University and Assistant Professor in the Universities of Ethiopia and Libya over 10 years and currently he is serving as senior faculty in Asia Pacific University of Technology and Innovation (APU), Malaysia. His areas of research include Medical Imaging, Image Segmentation, Fractal Analysis and Big Data.

Dr. S. G. Lakshmi Narayanan has completed his Ph.D from Gandhigram Rural Institute – Deemed University and has got six years of experience in teaching and research. His areas of specialization include Medical image analysis, mammogram image analysis and classification. He has participated in various workshops and conferences and has got 8 publications in National and International Conferences and Journals.