



A Survey of Breast Cancer Detection and Classification Based on Texture Feature

Sanjaykumar R. Kinge¹, Dr. B. Sheela Rani², Dr. Mukul S. Sutaone³, Abhijeet Vishwanath Sansare⁴

Research Scholar, Sathyabama University, Chennai and Assistant Professor, Dept of Electronics and Telecommunication, MIT College of Engineering, Pune, India¹

Director Research, Sathyabama University, Chennai, India²

Professor, Dept of Electronics and Telecommunication, College of Engineering Pune, Pune, India³

M.E. Student, Dept of Electronics and Telecommunication, MIT College of Engineering, Pune, India⁴

Abstract: Breast cancer is second most dangerous disease in the world after the lung cancer among women. Because of this reason, breast cancer detection is most focused area by many researchers. Digital mammographic images and advanced techniques are required by many researchers for developing Computer aided detection (CAD) for breast cancer detection. This study describes the recent advances in image processing and machine learning techniques for breast cancer detection. The study shows that Local Binary Pattern method used for feature extraction and Support vector machine for classification as foremost technique used for breast cancer detection. The comparative study of literature work summarizes the effectiveness of different approach used by researchers for breast cancer detection.

Keywords: breast cancer detection, CAD, image segmentation, feature extraction etc.

I. INTRODUCTION

Breast cancer is common and one of the deadly disease found in women. Prevention of the breast cancer in its early stage is difficult as the cause of the disease is still not known. X-ray mammography is very useful technique for screening and early stage detection of cancer. Regular mammography can help discover cancer in its beginning stage, when treatment is best. A mammogram can discover breast changes that could be disease years before physical side effect develop. Previous research shows that women who have regular mammogram can benefits by discovering breast cancer at an early stage. Such women don't need treatments like surgery to remove the breast, which also help them to cure the cancer.

Women who are at high risk for breast cancer based on certain factors should get a Magnetic resonance Imaging (MRI) and a mammogram every year. The American cancer society recommends against MRI screening for women whose lifetime risk of breast cancer is less than 15%. If MRI is used, it should be in addition to, not instead of a screening mammogram. This is because although an MRI is more likely to detect cancer than mammogram, it may still miss some cancer that a mammogram would detect [1].

Identifying masses asymmetries are difficult for radiologist due to inexperience, they may give wrong output because the mammograms are very sensitive to analyse. Computer Aided Detection (CAD) technology can improve the performance of radiologist. The CAD technique can minimize the events of mistakes conferred by the specialists. CAD system can characterize the normal and abnormal, benign and malignant case from digital mammogram by extracting the relevant features.

In computer aided detection segmentation of the region of interest that is breast profile and pictorial muscle is very important preprocessing stage [3]. The pictorial muscle in mammogram biases detection procedure; therefore it is essential to remove the pictorial muscle during preprocessing. Jawad Nagi et al. [3] present an automated technique for mammogram segmentation. The proposed algorithm make the use of morphological preprocessing and seeded region growing (SRG) algorithm to remove digitization noise, suppress radiopaque artifacts, separate background region from the breast profile region and remove the pictorial muscle, for highlighting the breast profile region.

This survey focuses on different schemes used for breast cancer detection and their comparison based on the result obtained through respective research. The Section II describes different methods for development of algorithms used for breast cancer detection, section III evaluates the methodology for development of CAD system and section IV summarizes the study.



**International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering**

ISO 3297:2007 Certified

Vol. 5, Issue 6, June 2017

II. RELATED WORK

R. Guzman Caberrera et al. [2] propose a computer Aided Detection scheme to separate suspicious region containing the masses from background parenchyma using local entropy. The proposed algorithm composed of background and objects texture segmentation and extraction. The texture features of the background and objects are obtained by using local entropy. In the proposed work, window size of 9*9 was used to compute local entropy, after that the properties of the local textures are extracted based on the adjacent neighbours of current pixel. The regions of interests are identified and then extracted using both the original input image and the background texture as a contrast mask. The experiment is evaluated on the mammographic images from DDSM database. However, the study proposed is highly dependent on an adequate value of the reference gray level to achieve a successful segmentation and extraction of the suspicious region.

Pavel kral et al. [4] proposes a method for breast cancer detection at an early stage, using LBP features. The background image and breast image is separated using Otsu method. The Local binary patterns values for breast image are calculated, after which image is divided into a set of the square cells lying on a regular grid. Feature vectors are computed for each cell as a histogram of the LBP values. Every cell is then represented by one feature vector. Support vector machine (SVM) based on uniform LBP histogram is used to classify, whether the corresponding region is normal or cancerous carried out on mammographic images obtained from DDSM and MIAS database. The proposed study achieves the accuracy approximately 84%.

Leonardo de oliveira et al. [5], proposes a method for detection of masses in mammographic images using k-means and support vector machine. Image segmentation is done using k-means algorithm and texture characteristics of the segmented area is obtained through co-occurrence matrix and shape descriptor. Based on the texture characteristics of the segmented region, support vector machine classifies the normal and cancerous tissues. The proposed algorithm is evaluated on the mammographic images extracted from Digital Database for Screening mammography. There are certain limitations over the use of k-means algorithm used in the study, which can't detect the masses due to lack of ability to differentiate between mass region and neighbouring region (dense breast tissue). The accuracy achieved by these work is over 85%.

Anuj kumar et al. [6] proposes noteworthy research which composed of detection and segmentation of region of malignant tissue. The method detects the cancerous area, then segment the respective area. In a detection stage, by applying averaging filter and thresholding operation to an original input image, cancerous region is detected. With the help of rectangular window over the detected cancerous region, max-mean and least variance techniques are applied to find the malignant tissues. Finally, by applying morphological closing operation, tumour patch is detected and the region boundary is detected using image gradient technique. As per authors claim, the proposed algorithm is simple and fast because of the basic image processing technique and successful in segmentation of the malignant tissues of the mammographic images. However, there is a dependency on manual selection of threshold parameter and size of the averaging filter.

S. Naresh et al. [7] proposes an approach for breast cancer detection, which uses completed Local Binary Pattern (CLBP) for texture feature extraction with support vector machine (SVM) for classification. The experiments are performed on MIAS database. In the work, morphological operations are carried out to remove noisy data, enhance and contrast the input image. Author claims that CLBP gives robust result as it consider both sign and magnitude features compare to the LBP, which consider only sign parameter, leads to the loss of other important information.

A significant research of implementing Eigen faces approach for face classification to mass detection is proposed by Arnau Oliver et al. [8]. The author present some similarity between face detection and mass detection algorithm, in which the face from the image is detected and then it is assigned to a person. The eigenfaces approach for mass detection has two limitations, which are the multiple size of RoI's and the variance of the gray-level range. In the present work, a uniform distribution model for equalizing the images to overcome the problem of gray-level range.

The problem of multiple size of RoI's is solved using 3 different ways: 1. collecting the database of RoI's in different group based on their size.

Whenever a new RoI is detected, based on their size, corresponding collection will be referred. 2. by cropping the RoI's by using same size. 3. by cropping the exact size of detected RoI and resize it. With the help of leave-out methodology and ROC analysis, performance of above mentioned approaches was calculated which was different for each one. The best result obtained through this work is $A_z=0.92$.

Eanes torres Peirira et al. [9] proposes noteworthy classification method to detect six different type of breast cancer namely: CALC, CIRC, SPIC, MISC, ARCH, ASYM. The proposed work is based on LBP feature and SVM classifier. LBP has two important parameters: - P (Number of neighbours) and R (Radius of comparison). The selection of this parameter gives the less or more number of LBP features. In proposed experiment, initially gave 66,080 LBP features by using $P= \{4, 8, 16\}$ and $R= \{1, 2\}$, which increased the processing time of SVM classifier. This problem is overcome by adjusting the parameter ($P= 4, R= 1$) and ($P= 8, R= 1$), resulting in 277 LBP features, used in the study.



**International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering**

ISO 3297:2007 Certified

Vol. 5, Issue 6, June 2017

Experiment was evaluated on images form MIAS database. The performance is measured using ROC, which is greater than 0.8.

Arнау Oliver et al. [10] propose an approach focusing on reduction of false positive mass detection using Local Binary Pattern. A group of 1792 uncommon region of interest was obtained from DDSM database. The RoI images divided into 5*5 rectangular regions. Each rectangle in 5*5 rectangle regions is divided into 3*3 regions and LBP features extracted to improve the quality of final RoI descriptor. Textural descriptors are independently obtained with the help of LBP for each region in 5*5 rectangular regions and then combined to form a global descriptor of RoI based on textural information. The classical LBP histogram is thus amplified to a spatially enhanced histogram which encodes both the local region and the spatial relation of the mass. SVM with a polynomial kernel is used to differentiate RoIs between a true mass and normal parenchyma. From 1792 extracted RoIs, 256 characterize a true mass and remaining normal but suspicious tissue. The performance of research is evaluated using leave one-out strategy and receiver operation characteristics with value $A_z = 0.909 \pm 0.043$.

Spandana et al. [11] presented a notable research for breast cancer detection at an early stage with the help of Matlab and Labview. The extraction of the mass with high intensity value is done by obtaining the mask of the mass by applying manual thresholding using histogram of the image. In Labview, region of mass is extracted by performing logical AND operation to the original image and mask. The border of the mass, which are obtained through morphological operator, are superposed on the original image with the help of logical OR operation in labview. Texture, wavelet and geometric features of the RoI of segmented masses are measured to differentiate between malignant and benign. For experiment evaluation, mammographic images of 14 patients were acquired from KIMS, Hyderabad. Most of the calculated features help in classification of benign and malignant masses.

Table I shows the comparative study of all researches discussed in literature review.

TABLE I COMPARATIVE LITERATURE WORK FOR DETECTION OF BREAST CANCER

| Author | Image Database | Feature | Remark |
|----------------------|---|---|---|
| Caberrera et al. [2] | DDSM database | Texture feature | Highly dependent on a reference gray level |
| Pavelkral et al. [4] | DDSM and MIAS database | LBP features | Accuracy achieved – 84% |
| Anujkumar et al. [5] | --- | Gradient | Dependent on threshold parameter selection and size of the averaging filter |
| Oliveira et al. [5] | DDSM database | Texture characteristics | Accuracy achieved – 85% |
| S. Naresh et al. [7] | MIAS database | LBP features | Results of CLBP are better than LBP |
| Oliver et al. [8] | --- | Eigen faces approach | Result is $A_z = 0.92$. |
| Peirrita et al. [9] | MIAS database | LBP features | ROC performance > 0.8 |
| Oliver et al. [10] | DDSM database | LBP features | ROC performance = 0.909 +/-0.043 |
| Spandana et al. [11] | Mammographic images of 14 patients were acquired from KIMS, Hyderabad | Texture, wavelet and geometric features | -- |

III.METHODOLOGY

This section describes about the different stages involved in order to achieve the purpose of each research mentioned in the literature review. The different stages re composed of preprocessing of the input image for extracting the features, which are then applied as an input to the classifier. The output of the classifier differentiates the normal, benign and malignant cases from applied digital mammographic images.

A. Image database.

Digital mammographic images are easily available on internet which can be downloaded from the respective web address. Digital database for screening mammography (DDSM) is one of the databases available from combined efforts of Massachusetts General Hospital, Sandia National Laboratories and the University of South Florida Computer Science and Engineering Department gives approximately 2,500 case studies [13]. The Mammographic Image Analysis Society (MIAS) MiniMammographic Database is another source of digital mammographic images available easily [12]. The mammographic image analysis society is an organization of UK research group has generated a database of digital mammogram.



**International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering**

ISO 3297:2007 Certified

Vol. 5, Issue 6, June 2017

B. Image preprocessing

The purpose image preprocessing is to remove unwanted noise present in the original input image and improve the quality of fine details present in it. This will make it easy for further processing of image in order to achieve the defined aim. The image preprocessing include image enhancement, removal of noise and breast part extraction. Smoothing image pixel using averaging filter [5], Otsu method [4] to separate background from breast region, morphological operations [6], [11], sharpening [2] have been used for preprocessing on digital mammographic images.

C. Feature extraction

Once the original image is preprocessed, it is possible to extract the relevant features specified in literature. Feature extraction can be described as extracting significant fine information from given input while rejecting all other data. LBP texture feature [4], [7], [9], [10], texture [2], [5], intensity value [11], eigenfaces approach [8], shape [6] feature were extracted from digital mammograms.

D. Classification

After extracting the relevant feature, the final stage is to classify the obtained data and assigned it to a particular class. For this purpose, classifier like Support Vector Machine [4], [5], [7], [9], [10], Decision tree [8], KNN [8], thresholding [6] were use

IV. CONCLUDING REMARK

The literature review show that every research discussed here has some dependency on certain parameter in proposed scheme. It can be a reference gray level, threshold selection or size of average filter, which makes the respective algorithm less efficient to achieve defined purpose. Local Binary Pattern method can extract the fine details from image which can be useful to develop CAD system for breast cancer detection in its early stage. From the literature, it is clear that support vector machine is the best option as a classifier for classification between cancerous and non cancerous cases.

REFERENCES

- [1] American Cancer Society, Breast Cancer: Facts and Figures,
- [2] R. Guzman-Cabrera, J.R. Guzman-Supulveda, M. Torres-Cisneros, D.A. May-Arrijoa, J. Ruiz-Pinales, O.G. Ibarra-Manzano, G. Avina-Cervantes, A. Donzalez Parada, "Digital image processing technique for breast cancer detection", *Int J Thermophys*(2013)34:1519-1531
- [3] Jawad Nagi, Sameem Abdul Kareem Farukh Nagi, Syed Khaleel Ahmed, "Automated breast profile segmentation for ROI detection using digital mammograms", 2010 IEE EMBS Conference on biomedical Engineering and Sciences (IECBES 2010), Malaysia, pp.87-92.
- [4] Pavel Kral, Ladislav Lenc, "LBP features for breast cancer detection", *ICIP2016*, pp. 2643-2647.
- [5] Leonardo de Oliveira Martins, Geraldo Braz Junior, Aristifanes Correa Silva, Anselmo Cardoso de Paiva, and Marcelo Gattass, "Detection of masses in digital mammogram using K-means and support vector machine", *Electronic Letters on Computer Vision and Image Analysis*8(2):39-50, 2009.
- [6] Anuj Kumar Singh, Bhupendra Gupta, "A novel approach for breast cancer detection and segmentation in a mammogram", *Eleventh International Multi-Conference on Information Processing-2015(IMCIP-2015)*, *Procedia Computer Science* 54(2015), pp. 676-682.
- [7] S. Naresh, S. Vani Kumari, "Breast Cancer Detection using Local Binary Patterns", *International Journal Of Computer Applications* (0975-8887), Volume 123- No.16, August 2015, pp. 6-9.
- [8] Arnau Oliver, Joan Marti, Robert Marti, Anna Bosch, Jordi Freixenet, "A new Approach to the classification of mammographic masses and normal breast tissue", *The 18th International Conference on Pattern Recognition (ICPR'06)*, pp.1-4.
- [9] Eanes Torres Pereira, Sidney Pimentel Eleuterio, Joao Marques de Carvalho, "Local Binary Patterns Applied to Breast Cancer Classification in Mammographies", *RITA Volume Numero*, pp. 1-15.
- [10] Arnau Oliver, Xavier Llado, "False Positive Reduction in Mammographic Mass Detection Using Local Binary Patterns", *MICCAI2007, Part I*, *LNCS* 4791, pp.286-293, 2007.
- [11] Spandana Paramkusham, K. M. M. Rao, B. V. V. S. N. Prabhakar Rao, "Early Stage Detection of Breast Cancer using Novel Image Processing Technique, Matlab and Labview Implementation", *15th International Conference on Advance Computing Technologies (ICTACT) 2013*, pp. 1-5.
- [12] Mammographic image analysis society (MIAS) MiniMammographic Database, <http://peipa.essex.ac.uk/info/mias.html>
- [13] Digital Database for Sreening mammography, <http://marathon.csee.usf.edu/Mammography/Database.html>

BIOGRAPHIES

Sanjaykumar R. Kinge is Research Scholar, at Sathyabama University, Chennai and Assistant Professor in Electronics and Telecommunication department, MIT, College of Engineering, Pune, INDIA – 411 038. He completed his Bachelor of Engineering in Electronics and Telecommunication in 1992 and subsequently completed his Masters of Engineering in 2008. He is presently pursuing his Ph.D. on the topic "Quantitative Approach to Texture Segmentation" at Sathyabama University, Chennai. His areas of interest and work are Signal and Image Processing, Electronic Design.

**International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering**

ISO 3297:2007 Certified

Vol. 5, Issue 6, June 2017

Dr. B. Sheela Rani was born on November 21, 1966 in Chennai. With a passion towards teaching, she started her career in 1989 as a lecturer and has been with the Institution for over twenty-five years in various academic and leadership positions as Professor, Faculty Head and as Dean (Research). With a strong technical expertise, she acquired M.S (By Research) and Ph.D degree from Anna University, College of Engineering Chennai, in 2001.

By virtue of working for a considerable number of years in the academic and research arena, she reached the position of the Vice Chancellor of Sathyabama University and assumed office in April 2011.

As an active researcher, Dr.B.Sheela Rani is associated with many Research Organizations. She plays a vital role as Research Coordinator of Sathyabama University. She attracted a number of research projects from many scientific funding agencies. She serves as a Project investigator for over 20 sponsored research projects. She delivers invited talks in reputed R&D institutions, chaired technical sessions in Conferences and conducted a large number of short term courses in Image Processing as part of continuing Education programme for the benefit of working Engineers, Scientists and Teachers Prof. B. Sheela Rani is a Life Member of various professional bodies and Member secretary of Indian society of Remote Sensing, Chennai Chapter. She has contributed to the areas of Image Processing, Soft Computing and Intelligent Computing with more than 150 publications in referred research journals, In addition, to the good teacher awards that she has received; she was a recipient of the S&T Award of University for her Contribution to Research. Visited Netherlands for fostering academic and research Collaborations.

Dr. Mukul S. Sutaone is Professor at Electronics and Telecommunication department and Dean Academics at College of Engineering, Pune, INDIA – 411 005. He has Ph.D. in the field of Texture Analysis form University of Pune. His areas of research and work are Multimedia and Multidimensional Signal Processing and Digital Communication.

Abhijeet Vishwanath Sansare is student at department of Electronics and Telecommunication at MIT College of Engineering, Kothrud, Pune. He completed his Diploma in Medical Electronics in 2010 and completed his Bachelor of Engineering in Electronics in 2013. He is presently pursuing his Master of Engineering and his research topic for final year dissertation is breast cancer detection and classification using image processing and machine learning techniques. His areas of research and work are digital image processing and medical product design and development.