Literature Review on Detection of Breast Cancer by Mammogram Interpretation Using CAD System

Grace John M\textsuperscript{1}, Amresh Kumar\textsuperscript{2}

\textsuperscript{1}Assistant Professor, Vimal Jyothy Engineering College, Chemperi, Kerala, INDIA, gracejohn@vjec.ac.in
\textsuperscript{2}Assistant Professor, G.H. Raisoni College of Engineering, Nagpur, Maharashtra, INDIA,

Abstract: Breast cancer is one of the major causes of non-accidental death in women. Early diagnosis of the disease allows physician to administer suitable treatment, and can improve the patient’s survival rate. Traditional diagnosis involves trained physicians to visually examine the medical images of breast for any signs of tumor development in the region. However due to the large scale of the medical image data, this manual review by the radiologist and for long-term storage. The mammogram currently the most effective tool for the screening tool. An x-ray (radiograph) is a medical test that helps physicians diagnose and treat medical conditions. Imaging with x-rays involves exposing a part of the body to a small dose of ionizing radiation to produce pictures of the inside of the body. X-rays are the oldest and most frequently used form of medical imaging. Mammography can be used to check for breast cancer in women who have no signs or symptoms of the disease. It can also be used if you have a lump or other sign of breast cancer. The goal of mammography is the early detection of breast cancer, typically through detection of characteristic masses and micro calcifications. Digital mammography, also called full-field digital mammography (FFDM), is a mammography system in which the x-ray film is replaced by electronics that convert x-rays into mammographic pictures of the breast. These images of the breast are transferred to a computer for review by the radiologist and for long-term storage. The patient's experience during a digital mammogram is similar to having a conventional film mammogram. The mammogram conﬁrms whether the changes are benign (noncancerous) and no treatment is needed, or whether the changes indicate breast cancer and further tests and treatment will be required.

Radiologists visually search specific abnormalities present in the mammogram images. Some of the important signs of breast cancer that radiologists look for are clusters of micro calcifications, masses, and architectural distortions. However, the analysis of mammograms by radiologists results in high rates of false positive cases. This is possibly due to several reasons. Detection of suspicious abnormalities is a repetitive and fatiguing task since there will be a huge database resulting after a screening test. For every thousand cases analyzed by a radiologist, only 3 to
4 are cancerous and thus an abnormality may be overlooked. As a result, radiologists fail to detect 10-30% of cancers. Approximately two-thirds of these false-negative results are due to missed lesions that are evident retrospectively. Due to the considerable amount of overlap in the appearance of malignant and benign abnormalities, mammography has a positive predictive value (PPV) of less than 35%, where the PPV is defined as the percentage of lesions subjected to biopsy that were found to be cancer. Thus, a high proportion of biopsies are performed on benign lesions. Avoiding benign biopsies would spare women anxiety, discomfort, and expense. Dense breast tissue can look white or light gray on a mammogram. Dense tissues obscuring the cancer and the fact that the appearance of cancer on mammograms has a large overlap with the appearance of normal tissues. This can make mammograms harder to interpret in younger women, who tend to have denser breasts.

To overcome these problems regarding with the interpretation and thus diagnosis of the breast cancer, human start to depend machines for better result. Generally, computers are employed to process and effectively evaluate biomedical images for more accurate diagnosis. Humans are prone to mistakes and their analysis is normally subjective and qualitative. Objective and quantitative biomedical image analysis is made possible with the use of computers results in a more accurate diagnostic assessment by the physician. CAD system is designed and developed to alert a radiologist to abnormal areas of a mammogram. Typically, CAD tool processes digitized version of a mammogram and marks it with „prompts” to highlight mammographic features that the reader should examine. The design goals for CAD include detection of features in a mammogram that might indicate cancer but that they may otherwise miss.

II. Literature Review

Chun-Chu Jen and Shyr-Shen Yu (2014)[1] used mean filter for the image denoising. The study first investigates image preprocessing techniques for obtaining more accurate breast segmentation prior to mass detection, including global equalization transformation, denoising, binarization, breast orientation determination and the pectoral muscle suppression. After performing gray level quantization on the breast images segmented, the presented feature difference matrices could be created by five features, such as mean, standard deviation of the total pixel intensities, the number of different layer heights in which the pixel intensities are equivalent on a layer of the same height, the mean gradient and the total gradient extracted from a suspicious ROI. Subsequently, principal component analysis (PCA) is applied to aid the determination of feature weights. The mean filter is used to eliminate short-tailed noise such as uniform and Gaussian-type noise from the image. ADC classifier is a two-stage classifier for automatic detection of abnormal mammograms based on the features extracted from the ROI. In the first stage, PCA technique is used to compute eigen values for the first defined abnormal and normal feature difference matrices, the initial feature weights would be obtained. After that, the ADC calculates respective centroid coordinates for the two feature difference matrices using the proposed feature weight adjustment algorithm and then generates two rows of normalization denominators for further calculation. In the second stage, by measuring two different Euclidean distances between the tested image that is adjusted by multiplying respective weight factors then divided by the normalization denominators and the corresponding centroid coordinates, the ADC can automatically detect if an abnormality exists in a tested image for the entire mammographic dataset. The experimental results show that applying the algorithm of ADC accompanied with the feature weight adjustments to detect abnormal mammograms has yielded prominent sensitivities of 88% and 86% on the two respective datasets. Comparing other automated mass detection systems, this study proposes a new method for fully developing a high-performance, CAD system that can automatically detect abnormal mammograms in screening programs, especially when an entire database is tested.

J. Dheeba et al (2014)[2] introduced a classification approach for the detection of breast abnormalities in digital mammograms using Particle Swarm Optimized Wavelet Neural Network (PSOWNN). Global thresholding techniques and intensity histogram methods were used in the preprocessing stage. The proposed abnormality detection algorithm is based on extracting Laws Texture Energy Measures from the mammograms and classifying the suspicious regions by applying a pattern classifier. The method was applied to real clinical database of 216 mammograms collected from mammogram screening centers. The detection performance of the CAD system was analyzed using Receiver Operating Characteristic (ROC) curve. This curve indicates the trade-offs between sensitivity and specificity that is available from a diagnostic system, and thus describes the inherent discrimination capacity of the proposed system. The result showed that the area under the ROC curve of the proposed algorithm is 0.96853 with a sensitivity of 94.167% and specificity of 92.105%.
S.D. Tzikopoulou et al. (2010) [3] proposed a fully automated segmentation and classification scheme for mammograms, based on breast density estimation and detection of asymmetry. In the preprocessing stage thresholding and median filtering were employed. In the segmentation stage, interface combining algorithm for breast boundary detection, straight line estimation and validation for pectoral muscle detection and thresholding for nipple detection are used. SVM classifier is used for the classification of mammogram image which detect breast asymmetry.

R. Llobeta et al. (2014) [4] presented a semi-automated and a fully automated tools to assess breast density from full-field digitized mammograms. The semi automated tool is based on a supervise dinteractive thres holding procedure for segmenting dense from fatty tissue. The fully automated method presented combines a classification scheme based on local features and thresholding operations that improve the performance of the classifier. Histogram Stretching are used for contrast & brightness correction and thresholding is used for segmentation of semi automated system. The segmentation obtained is used as a method of supervised pixel labeling, which is used to train a fully automated classifier. Classfification stage of fully automated system used KNN classifier.

In the paper by Rahimeh Rouhi et al. (2014) [5], two automated methods to diagnose mass types of benign and malignant in mammograms are proposed. In the first method, segmentation is done using an automated region growing whose threshold is obtained by a trained artificial neural network (ANN) and in second method, segmentation is performed for medhya cellular neural network (CNN) whose parameters are determined by a genetic algorithm (GA). Local area histogram equalization and then the median filtering is applied to suppress noise. In the histogram equalization stage, the intensity of image pixels is stretched to extend the contrast. Median filtering is used on linear operation often used in image processing to reduce ‘salt and pepper’ and speckle noise. Intensity, textural, and shape features are extracted from segmented tumors. GA is used to select appropriate features from the set of extracted features. In the next stage, ANNs are used to classify the mammograms as benign or malignant. The obtained sensitivity, specificity, and accuracy rates are 96.87%, 95.94%, and 96.47%, respectively.

A good primer to ‘Image denoising based on wavelet transforms need wavelets and smoothing functions’, is introduced by J. Scharcanski and C. R. Jung (2006) [6]. At each resolution, coefficients associated with noise are modelled by Gaussian random variables; coefficients associated with edges are modelled by Generalized Laplacian random variables, and a shrinkage function is assembled based on posterior probabilities. The shrinkage functions at consecutive scales are combined, and then applied to the wavelets coefficients. The image de-noising process is adaptive which does not require further parameter adjustments, and the selection of a gain factor provides the desired detailenhancement.

Shen-Chuan Tai et al. (2014) [7] presents an automatic CAD system that uses local and discrete texture features for mammographic mass detection. The Otsu thresholding method applied to the image to find the breast region and then gamma expansion enhances the pectoral muscle. The proposed system used two types of morphological filters, the opening filter which smooths bright object contours and removes small dark spots and closing filter which smooths the dark object contours and removes small bright spots. The such template matching method is applied to the filtered breast region to identify suspicious mass regions. Two feature extraction approaches were introduced, by combining co-occurrence matrix texture features and optical density features. One is a combination of GLCM (Gray-Level Co-occurrence Matrix) features and optical density features, describes both the gray-level characteristics of local textures and photometric discrete textures based on the global optical density. The other method combines ODCM (Optical Density Co-occurrence Matrix) features with optical density features which characterizes local textures in an optical density image instead of the gray-level image.

Vibha Bora Bafna et al. (2015) [8] adopted pattern based radon transform to detect the orientation of mammograms. The proposed orthogonal radon transform detects the orientation of chest wall and pectoral muscle, and is capable of correcting vertical as well as upside down orientation present in the mammogram. After the application of the method, mammogram is oriented with chest wall in left and pectoral muscle in upper corner to increase the accuracy of post processing steps in CAD. Median filtering and global thresholding were employed in the preprocessing stage. The algorithm was tested on 240 images in which 100 images from mini MIAS, 100 computed radiography images and 40 full
field digital mammogram images. For 98.75% of position invariant cases, orientation was detected and corrected. Due to its robustness, simplicity and fast speed, the method may be used for automatically detecting and correcting orientation of mammograms in CAD analysis.

Jawad Nagi et al. (2010) [9] proposed an automated technique for segmentation of ROI in digital mammogram images. The presence of pectoral muscle in mammograms biases detection procedures, which recommends removing the pectoral muscle during mammogram pre-processing. The proposed algorithm uses morphological pre-processing and seeded region growing (SRG) algorithm for remove digitization noises, suppress radiopaque artifacts, separate background region from the breast profile region, and remove the pectoral muscle, for accentuating the breast profile region. They adopted two-dimensional (2D) Median Filtering approach in a 3-by-3 neighborhood connection for the noise removal and thresholding method for artifact suppression and background separation. To demonstrate the capability of proposed approach, digital mammograms from two separate sources are tested using Ground Truth (GT) images for evaluation of performance characteristics. Experimental results obtained indicate that the breast regions extracted accurately correspond to the respective GT images.

Hitiksha shah (2015) [10] presented a CAD system for automatic classification of breast masses in digital mammograms. The digital mammogram is pre-processed by 2D-median filter, connected component labelling method, and morphological functions for breast extraction. Wavelet transform is used for enhancement of mammogram and triangular mask is used for pectoral muscle suppression. Morphological functions such as opening, closing, erosion, dilation and reconstruction are used for the segmentation of mammogram to extract ROI. From ROI, intensity histogram based texture features are extracted. Extracted features are classified by neural network which is applied for two levels. In the first level, neural network classify the segmented ROI into normal (without tumor) and abnormal (with tumor) ROI. Second level neural network classify abnormal ROI into malignant and benign masses.

Zaheeruddin, Z. A. Jaffery and Laxman Singh (2012) [11] proposed a method for detection and shape feature extraction of breast tumor in mammograms. In this work, the noise and uneven illumination are removed using a Weiner filter, then, applied the histogram equalization method to enhance the image contrast that scales the gray level of each voxel by a defined factor. Segmentation of breast tumor in mammograms presents many challenges related to selection of optimal threshold in various segmentation techniques. In the proposed paper, a mean based region growing segmentation (MRGS) is presented that automatically finds the seed pixel and optimal threshold value and thus makes the segmentation process very fast and accurate. Basic region growing method is known to be the most effective tool for performing the quantitative analysis of anatomical structures in medical images. But, it does not lead to the accurate detection of an object, when directly applied on raw input images containing noise and poor contrast. So, the proposed algorithm could be applied on the mammographic images effectively, when the noise and other local irregularities are removed from the input images. The experimental results were compared with the findings of expert radiologist and marker controlled watershed segmentation approach. A set of 3 mammogram images was used to demonstrate the effectiveness of the segmentation methods. Numerical validation of the results was also provided.

Abdelali Elmoufidietal. (2016) [12] introduced a method to generate and select the features of suspicious lesions in mammograms and classifying them by using support vector machine, in order to build a CADx system to discriminate between malignant and benign parenchyma. The proposed method is divided into two major blocks such as extraction and selection of technical features for each region of interest, and classification of ROIs extracted to benign or malignant parenchyma. Eighteen features, such as mean value, standard deviation, entropy, skewness, kurtosis, uniformity, sum entropy, sum average, difference variance, difference entropy, inverse difference moments, area, perimeter, convexity, compactness, aspect ratio, area to background percentage and perimeter ratio were selected and extracted for the classification. SVM classifier was used to diagnose breast cancer. The method has been verified with the well-known Mammographic Image Analysis Society (MIAS) database and they have used the Receiver Operating Characteristics (ROC) to measure the performance of the method. The experimental results produced a classification accuracy of 96.36%, with 96.77% sensitivity and 95.83% specificity in the training phase and achieved an overall classification accuracy of 94.29%, with 94.11% sensitivity and 94.44% specificity in the testing phase.
Luqman Mahmood Mina and Nor Ashidi Mat Isa (2015)[13] presented a fully automated breast separation for mammographic images. The main contribution of this algorithm is applying the combined of thresholding technique and morphological preprocessing to segregate background region from the breast profile and remove radiopaque artifacts and labels. To show the validity of this segmentation system, it is extensively tested using over all mammographic images from the MIAS database. The MIAS database comprises 322 images with high intensity rectangular labels. Bright scanning artifacts were found to be present in majority of the database images. All square high intensity labels, apart from three were removed at a rate of 99.06%. The qualitative assessment of experimental results indicates that the method can accurately segment the breast region in a large range of digitized mammograms, covering all density classes.

Moh'd Rasoul A. Al-Hadidi etal. (2012)[14] proposed a method for solving mammography problems of breast cancer detection using artificial neural networks and image processing techniques. They presented a complementary technique of breast cancer diagnosis that covers five stages of breast cancer detection based on mammography, which solves many of the problems. Many methods and techniques were successfully merged in order to obtain useful results for human use, which includes scaling of the image, removing small objects, smoothing, extracting features, ROI extraction and many image processing techniques. Besides, neural networks are used to train the system to detect cancer according to the dataset. This combination of multiple techniques can solve problems of breast cancer detection with a high degree of accuracy. Adaptive histogram equalization is used to improve contrast in the mammogram images. Morphological closing was performed in order to smooth out the edges and to make sure that no irregularities occur at the edge of the image.

Pelin Kus and Irfan Karagoz (2012)[15] presented a fully auto-mated gradient based breast boundary detection for digitized x-ray mammograms. Median filtering method is used for the noise removal of the acquired mammogram. Then multidirectional scanning is applied to the resultant image using a moving window of size 15 X 1. By using the intensity value and maximum gradient value of the window, the border pixels are detected. The breast boundary is identified from the detected pixels filtered using an averaging filter. The results obtained by using the gradient based border estimation method introduced are close to the boundary drawn by the expert radiologist. No user guidance is required and all steps are carried out automatically. Selecting the starting point or obtaining initial borders, which are the main disadvantages of the other methods, are not encountered in the presented method. Moreover, the method is not as complicated as other methods and it does not include iterative calculations. Future studies will focus on detecting objects inside the breast region. The method performs well in three directions but for further mass detection application needs vectorial applications of this algorithm must be considered. The segmentation accuracy on a dataset of 84 mammograms from the MIAS data base is 99%.

M. Radhaand Dr. S. Adaekalavan (2016)[16] proposed frequency domains smoothing sharpening technique to enhance mammogram images. The presented technique had the advantages of enhancing and sharpening process that aims to highlight sudden changes in the image intensity, it is usually applied to move random noise from digital images. This method also eliminates the drawbacks of each of the two sharpening and smoothing techniques resulting from their individual application in image processing field. Gaussian filter was used to remove move noises presented in the mammogram, and segmentation is done using three holding techniques. K-means algorithm was used to detect the breast cancer in the early stage. The selection of parameters is almost invariant of the type of background issues and severity of the abnormality, giving significantly improved results even for denser mammographic images. The simulated results show that the high potential to advantageously enhance the image contrast hence giving extra aid to radiologists to detect and classify mammograms of breast cancer.

III. Conclusion

In this survey of mammogram image based CAD system for breast cancer detection, we had studied the major methods for detecting the breast cancer. The study includes existing effective computer methods introduced by various researchers in this area for the past ten years. This work will be a milestone for all the researchers working in the medical image processing field. In future we are developing a fully automated CAD system for the early detection of breast cancer stages by comparing these existing techniques.
Literature Review on Detection of Breast Cancer by Mammogram Interpretation Using CAD System

Acknowledgment

We submit the work in divine feet of God Almighty for blessing us with his wisdom for surpassing every difficulty faced during this work. We wish to thank everyone who supported to fulfill the works successfully.

References