Early Detection of Breast Cancer using Computer Aided Detection and Diagnosis: Recent Advances

Amith Bhaskar. C
[M.Tech], Microelectronics & Control Systems
Dayananda Sagar College of Engineering
Bangalore, 560078

Meharunnisa S. P
Assoc. Prof, Department of I.T
Dayananda Sagar College of Engineering
Bangalore, 560078

***Abstract*— According to Indian Council of Medical Research (ICMR), it is reported that One in 22 women in India is likely to suffer from breast cancer during her lifetime, while the figure is definitely more in US with One in 8 being a victim of this deadly cancer. In metropolitan cities such as Mumbai, Delhi and Kolkata, breast cancer accounts for 30%, 26.9% & 27.2% respectively. Among women in Bangalore it is almost 26.9%, while the occurances in men have been slightly less.**

Primary prevention seems impossible because the causes of this disease still remain unknown. So Early detection is the key to improving breast cancer prognosis. X-ray mammography is the most common technique used by radiologists in the screening and diagnosis of breast cancer. Although it is seen as the most reliable method for early detection of breast carcinomas (reducing mortality rates by up to 25%) its interpretation is very difficult. About 10%–30% of breast lesions are missed during routine screening due to oversightedness.

Masses and micro-calcifications are two important signatures of breast cancer. Masses which can be obscured or be similar to normal breast parenchyma, is the main objective of this study. Reading mammograms is a demanding job for radiologists, hence a Computer Aided Detection (CAD) System can provide a consistent second opinion to a radiologist and greatly improve the detection accuracy. This paper summarizes the survey on Breast Cancer diagnosis regarding different Image Enhancement techniques, Machine Learning algorithms and Diagnosing methods, which are used to improve the accuracy of predicting cancer. This survey can also help us to know about number of papers that are implemented to diagnose breast cancer and statistically analyze the recent advancements in the field of Digital Mammography.

Keywords—Masses; CAD; Mammography; Breast Cancer; Lesions.

I. INTRODUCTION

Breast cancer has become a significant health problem worldwide. In order to prevent the increase of deaths caused by breast cancer, early diagnosis of the disease has been very effective. Detected early, breast cancer is easier to treat, with fewer risks and reduces mortality by 25%[1]. This early detection can be achieved by subjecting women at risk (mainly postmenopausal women) to a mammography every two years, since it takes about five years for a breast tumor to reach 1 mm, two years longer to reach 5mm and one or two years to measure 2 cm, large enough to detect by palpation. Now adays, X-Ray mammography, ultrasound, and magnetic resonance imaging (MRI) are imaging modalities routinely used to screen for breast cancer. It is well known that there is no technology at present, which is capable of curing cancer. But, it is well known that early detection of cancer can aid in recovery and prolong patient life. A major reason for these errors is due to the fact that radiologists depend on visual inspection. During manual screening of a large number of mammograms, radiologists may get easily worn out, missing out vital clues while studying the scans. As yet, there is no definitive literature which focuses on an elaborate discussion on the feature extraction, feature selection and classification methodologies used in breast cancer detection.

Increasing risk of developing breast cancer includes early menarchy, delayed menopause, obesity, infertile mothers, contraceptives, consuming alcohol and certain inherited genetic mutation. Despite advances in other breast imaging modalities, including ultrasound and magnetic resonance imaging, mammography is still a method of choice. However the detection of small malignancies is especially difficult in younger women who tend to present denser breast tissues. Biopsy is applied for the most accurate diagnosis. However, it is an aggressive and high-cost procedure with some risk and causes inconvenience to the patient.

Mammography is the best available inspection facility to detect the symptoms of breast cancer at the early stage and it can disclose information about abnormality, such as masses, microcalcifications, bilateral asymmetry, and architectural distortion. Mammography is the breast image taken with a special X-ray. It uses a low-dose X-ray, high-contrast, and high-resolution film. For the hundreds of mammographic images scanned by a radiologist, only a few are cancerous. While detecting abnormalities, some of them may be missed, as the detection of suspicious and abnormal images is a recurrent mission that causes fatigue and eyestrain. Using enhanced images and segment the suspicious area , extract features , select more accurate features and then classify them into appropriate category are the most important steps that computer aided detection systems should follow. However, due to low contrast and strong noise, digital mammograms are among the most difficult medical images to analyze.

Since the calcification have high attenuation properties and small dense tissues similar to bones, they tend to present low contrast. Thus their visual screening is difficult for physicians. As the above difficulties prevail in the existing methodology of mammographic image processing, the improvement of local detail discrimination and removal of noise from the images is a requirement.

Masses and calcifications are two primary signatures of abnormity in mammograms. Existing research results show masses are more difficult to recognize because of their abundant appearances and ambiguous margins than calcifications, and thus, mass detection is a challenging problem.

Therefore, computer-aided diagnosis (CAD) systems have been composed to assist radiologists in the determination of abnormalities in mammography. These systems are developed only as a second opinion and the final decision belongs to the radiologist. The systems have improved the radiologists' performance to achieve the best accuracy of breast cancer detection. However, due to the tremendous amount of samples, adequate speed and accuracy by human is in most cases impossible. Therefore, the use of computer systems and artificial intelligence saves time in the calculations and the ability of specialists to respond to a large number of samples. However, medical centers almost everywhere today are facing a challenge in analyzing the increasingly high volume of mammograms. CAD techniques continues to be designed to accomplish early detection of breast cancer. These systems are built to analyze digital images; although the acquired images can reveal information of the inside of the breast, yet the signs of the cancer are very subtle and differ in form at their early stages; subtle cancerous cells have so far been affecting the performance of CAD systems. Numerous Computer Aided Diagnostic systems were developed in the last two decades for breast cancer.As breast masses detection, tend to be hard to identify, due to its numerous looks and uncertain margins, this paper presents a current study aims that at filling this gap by documenting developments in recent advancements of Computer Aided Detection for breast cancer detection.

II. DETECTION OF SUSPICIOUS REGIONS

First, Masses are defined as space-occupying lesions that are described by their shapes and margin properties. A benign neoplasm is smoothly marginated, whereas a malignancy is characterized by an indistinct border that becomes more spiculated with time. Because of the slight differences in the X-ray attenuation between masses and benign glandular tissue, they appear with low contrast and are often very blurred.

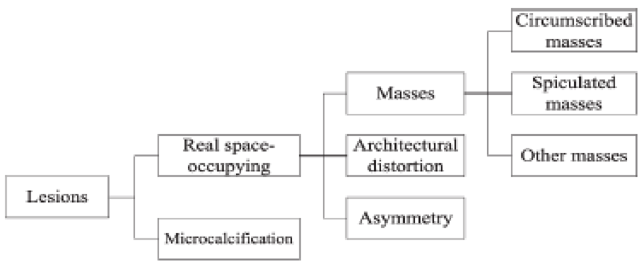

Figure : Classification of Tumors

A. Feature Extraction

The mammogram images are enhanced and the suspicious areas segmented, lesion features are extracted and are more accurately processed. For example, In MENCATTINI et al.[3], we see a technique where the denoising phase is based on a local iterative noise variance estimation. A new segmentation method is developed combining dyadic wavelet information with mathematical morphology.The innovative approach consists of using the same algorithmic core for processing images to detect both microcalcifications and masses. Furthermore, the design of the algorithm goes toward the hardware implementation of the heavy core of the wavelet computation, permitting the realization of a fast real-time processing of the image. In HU et al[4], a novel algorithm is used to detect suspicious lesions in mammograms. The algorithm utilizes the combination of adaptive global and adaptive local thresholding segmentation on a multiresolution representation of the original mammogram. the wavelet transform on the original mammograms removed the singularities and generated the lesion gray-scale information. Subsequently, the wavelet transform on the histograms (PDF curves) removed the fluctuations. Hence, the global local minima can be found as the adaptive global threshold to implement the coarse segmentation. In the meantime, the morphological filter on the transforms images not only removes the background noise and the structure noise inside the suspected mass pattern but also enhances the gray-level feature and shape feature of lesions. Finally, after a convolution between the coarse segmentation and the morphological enhancement filteres the gray-level image to perform the fine segmentation. In Sundarmai et al[5], incorporates both histogram modifications as an optimization technique and Contrast Limited Adaptive Histogram Equalization. It is interesting to note that from the subjective and quantitative measures, this proposed technique provides better contrast enhancement with preserving the local information of the mammogram images as the conventional histogram equalization (HE) usually results in excessive contrast enhancement because of lack of control on the level of enhancement. In Marija et al[6], digitized mammograms were analyzed using basic point operators to highlight particular features and to extract quantitative information. Main focus was on contrast enhancementbetween “suspicious” breast structures and adjacent tissues. Basic algorithms such as normalization, equalization and thresholding were applied and their actions have been shown in both: the image and its histogram. Basic point operators could be used successfully to help radiologists by increasing probability in early detection of breast cancer, even if the original image was not optimally taken. Feature extraction is considered as the most effective step in mammogram classification and can be distinguished in the three following stages[7]:

1. Statistical methods: The extracted features of this class include those obtained from co-occurrence matrices, surface variation measurements (smoothness, coarseness and regularity), and run-length statistics.

2. Model-based methods: The analysis of texture features in this class is based on prior models such as Markov random fields, auto-regressive models, and fractals.

3. Signal processing methods: In this class, texture features are obtained according to either pixel characteristics or image frequency spectrum including Laws energy filtering, Gabor filtering, local binary pattern and multi resolution analysis such as Wavelets, Curvlets, and contourlets.

After feature extraction, selecting the appropriate classifier plays an important role for obtaining good result, there are many classifier have been used such as k-nearest neighbors(KNN), support vector machines(SVM), Neural Network(NN), Euclidean Distance and Bayesian classifier. In this paper two techniques are presented for classification of abnormality in digital mammograms. The Statistical and LBP features as a good tool for features extraction.

III. COMPUTER AIDED DETECTION- DIAGNOSIS

A new system for automated mass detection[8] in mammography images is presented, after optimization of the image and extracting a better picture of the breast tissue from the image and applying log-polar transformation, as being more accurate and valid. Chebyshev moments can be calculated in all areas of breast tissue. Then after extracting effective features in the diagnosis of mammography images, abnormal masses, which are important for the physician and specialists, can be determined with applying the appropriate threshold. When compared the FROC curve with similar systems experts, the high ability of the system was confirmed. The proposed system achieves 100% sensitivity and 2.56 false positive for every image.

Abhijit Nayak et al[9], proposes a systematic method for the detection of suspicious lesions in digital mammograms based on undecimated wavelet transform and adaptive thresholding techniques. Undecimated wavelet transform is used to generate a multiresolution representation of the original mammogram. Adaptive global and local thresholding techniques are then applied to segment possible malignancies. The segmented regions are enhanced by using morphological filtering and seeded region growing. The experimental results show that the proposed method successfully detects 87 of the 92 lesions, performing with a sensitivity of 94.56% at 0.8 false positives per image (FPI), which is better than earlier reported techniques. This shows the effectiveness of the proposed system in detecting breast cancer in early stages.

A computer-aided diagnosis system[10] using the curvelet transform (CT) algorithm is proposed for interpreting mammograms to improve the decision making. The purpose of this study is to develop a method for the characterization of the mammography as both normal and abnormal regions, and to determine its diagnostic performance to differentiate between malignant and benign ones. A support vector machine and the k-nearest neighbor algorithm are used as classifers to build the diagnostic model and are also used for the principal component analysis and linear discriminant analysis for further dimensional reduction and feature selection.

In this paper [11], CAD is designed using a new algorithm named Gray level Gradient Buffering (GGB) method. The designed algorithm contains 4 steps: Preprocessing, Segmentation, Feature extraction and Classification. Noise and Artifacts are removed in preprocessing using 2D Median filter and ROI extraction process respectively. Segmentation is carried out using Sobel operator and the proposed GGB algorithm. The features of the segmented image are extracted using Spatial Gray level Dependence Matrix and Local Binary Pattern. Support Vector Machine is used to classify the extracted features as benign and malignant. The efficiency of the proposed method is calculated using Par test which gives 99.3%-Sensitivity and 98.1%-Specificity.

In this paper[12], a Particle Swarm Optimization (PSO) for tuning the enhancement parameter of Contrast Limited Adaptive Histogram Equalization (CLAHE) based on Local Contrast Modification (LCM) is presented. The PSO method of parameter tuning adopted for LCM-CLAHE enhancement for mammogram images achieves very good quality of images compared to other exiting methods. The quality of enhanced image is tested using an efficient objective criteria based on entropy and edge information of the image. the most important property is that it can produce better results with proper tuning of parameter. But in case of Standard Histogram Equalization, Unsharpmasking and Normal CLAHE it produces only one enhanced image for a particular input image.

This paper [13], introduces a computer aided system for detecting and classifying suspicious regions in digital mammograms. The system starts by extracting regions of interest that are suspicious, then all these regions are classified to check whether they are normal or abnormal. For the detection phase, template matching techniques are utilized. As for the classification phase, a 3-step process is applied which enclose feature extraction with wavelet transform such as Haar, Daubechies & Coieflets filter families. Feature selection with statistical techniques and classification with clustering K-Nearest Neighbor classifier. A preliminary result shows a 97.73 % accuracy rate.

In this study [14], we discuss the development of a technology which is able to mark the positions of possible masses, allowing further assessment by radiologists and effectively increasing the rate of correct diagnosis of breast cancer. Because masses in mammograms present themselves as low frequency signals, we have established the following steps for detecting them: Firstly, the original image undergoes wavelet transformation and enhances the mass signals before being inverse-transformed backward to an image; an image with enhanced processes would make masses easier to discern. Second, possible masses are identified and positioned using particle swarm optimization, PSO. Experimental results show that a detection rate of 94.44% or higher can be achieved using this method, hence improved accuracy in breast cancer lesion detection.

In Durgadevi et al [15], a novel hybrid method named M-HBMO (Mammogram based Honey Bees Mating Optimization) algorithm has been proposed to segment the lesion. The cancer profile segmentation is based on texture feature and extraction of the lesion. The M-HBMO is evaluated with conventional ROI (region of interest) Algorithm and the reason of using wiener filters to remove the noise pixel value as well as enhance the quality of the image. It's purely based on random search and neighborhood search method in order to identify the lesion without difficulty. At the Final position of M-HBMO algorithm, a good quality of result is obtained than other optimization techniques.

In Prapti V et al[16] proposes group of pre and post processing techniques with segmentation approach showing better results for medical image processing applications. Watershed transform of gradient magnitude of mammogram image is followed by reconstructive morphological processing with modified regional maxima superimposed on original image. The detection and isolation of accumulated masses from segmented region is very important for extracting useful feature measures which can be input to a classifier which classifies each region into benign or malignant.

In Bhagwati Charan Patel & G. R. Sinha [17] introduces a novel approach for accomplishing mammographic feature analysis through detection of tumor, in terms of their size and shape for early breast tumor detection. The abnormal tumor/tissue inside breast is to be detected using three stages: Preprocessing, Segmentation and post processing stage. By using preprocessing, noise is removed and then segmentation is applied to detect the mass, after that post processing is applied to find out the benign and malignant tissue with the affected area in the cancers breast image. In the proposed work we have designed a new computer aided detection method to detect the mass region in the mammogram. Segmented image contains the suspected region which is given for feature extraction process. The extracted features are classified into normal and abnormal region .The obtained accuracy was 95.6% whereas the sensitivity and specificity were found to be 96.5% and 89% respectively. The proposed system gives fast and accurate classification of breast tumors.

The wavelet analysis[17], is explored for analyzing and identifying strong variations in intensities within the mammographic data which highlights and recognizes the masses effectively. The proposed algorithm, in addition to wavelet transformation, uses morphological preprocessing, region properties and seeded region growing to remove the digitization noises, to remove the pectoral muscle and to suppress radiopaque artifacts, thus segmenting the abnormal masses accurately.

The purpose of this study[18] is to develop an intelligent remote detection and diagnosis system for breast cancer based on cytological images. First, this paper presents a fully automated method for cell nuclei detection and segmentation in breast cytological images. The locations of the cell nuclei in the image were detected with circular Hough transform. The elimination of false-positive (FP) findings (noisy circles and blood cells) was achieved using Otsu's thresholding method and fuzzy c-means clustering technique. The segmentation of the nuclei boundaries was accomplished with the application of the marker-controlled watershed transform. Next, an intelligent breast cancer classification system was developed. Twelve features were presented to several neural network architectures to investigate the most suitable network model for classifying the tumor effectively. Four classification models were used, namely, multilayer perceptron using back-propagation algorithm, probabilistic neural network (PNN), learning vector quantization, and support vector machine (SVM). The classification results were obtained using tenfold cross validation. The performance of the networks was compared based on resulted error rate, correct rate, sensitivity, and specificity. Finally, merged with the proposed computer-aided detection and diagnosis system with the telemedicine platform. This is to provide an intelligent, remote detection, and diagnosis system for breast cancer patients based on theWeb service. The results showed that the predictive ability of PNN and SVM is stronger than the others in all evaluated data sets.

IV. DUAL MODALITY

Multimodality information system provides complementary information to radiologists and has a great benefit for diagnosis and therapy. Radiologists in a normal follow-up face lot of difficulties in interpreting mammograms or ultrasonograms due to which more than half of breast biopsies turn out to be negative. Thus it is desirable to have an alternative approach as a second line of defence. Towards that we have embarked on developing a dual modality model which fuses the features retrieved from mammogram and ultrasound. For example mammogram provides textural and morphological information where as ultrasound images give functional and metabolic information. Hence most of the limitations imposed by unimodal systems can be overcome by including multiple sources of information. Images from different modalities may be taken at different time, different resolutions and may be from different viewpoints so it is very difficult to simply overlay different images from different modalities to fuse the information. Integration of information in multimodality can occur at feature level or at decision level. Feature level methods combine several feature sets into a single fused one which is then used by any conventional classifier. Decision level fusion combines several classifiers resulting in strong final classifier. In addition to that some other discriminative features like denseness texture feature, standard deviation, entropy and homogeneity are also extracted from ROI's of both modalities. Feature level fusion is then achieved by using a simple concatenation rule. Finally classification is done using Support vector machine (SVM) classifiers to classify breast mass as malignant or benign. Receiver operating characteristic curves (ROC) are used to evaluate the performance.

CONCLUSION

MRT, CT, Ultrasonic and Mammography are the available screening methods to detect the cancer cell. The diagnosis rate between histopathology rate and mammography is quite high while comparing with other screening methods. It is inexpensive and works fairly well in all cases. Another advantage of digital mammography is it stores the result as computer code, hence it can be modified for the screening process. Due to the above mentioned reasons Mammography is considered as clinical gold standard for early detection. But due to variety of factors such as poor quality of image, benign appearance of lesions and eye fatigue factor the performance of radiologists varies from 65% to 85%. Wavelet processing is extremely necessary when the analysis of a multiscale feature is important. In fact, when using the wavelet transform, it is possible to detect details that appear at different scales and selectively enhance them within different resolution levels. Hence we conclude by stating that Wavelet analysis could play a major role in processing of a mammogram and intelligent classifiers like SVM and K-nearest neighbour method can help in Computer Aided Detection of Breast Cancer.

REFERENCES

[1] www.icmr.nic.in

[2] http://www.radiologyinfo.org/en/info.cfm?pg=mammo.

[3] Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing, MENCATTINI et al.: MAMMOGRAPHIC IMAGES ENHANCEMENT AND DENOISING FOR BREAST CANCER DETECTION, 0018-9456, 2008 IEEE.

[4] Detection of Suspicious Lesions by Adaptive Thresholding Based on Multiresolution Analysis in Mammograms, Kai Hu, Xieping Gao, and Fei Li: IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 60, NO. 2, FEBRUARY 2011, IEEE.

[5] HISTOGRAM BASED CONTRAST ENHANCEMENT FOR MAMMOGRAM IMAGES, M.Sundarami, K.Ramar2, N.Arumugami, G.Prabini, Proceedings of 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN 2011), 2011 IEEE.

[6] Basic Feature Extractions from Mammograms, Marija Dakovic and Slavoljub Mijovic, Mediterranean Conference on Embedded Computing MECO - 2012.

[7] Breast Mass Classification using Statistical and Local Binary Pattern Features, Mohamed A. Berbar, Yaser. A. Reyad, Mohamed Hussain, 16th International Conference on Information Visualisation, 2012 IEEE

[8] Automated Abnormal Mass Detection in the Mammogram Images Using Chebyshev Moments, Alireza Talebpour, Dooman Arefan and Hamid Mohamadlou, Res. J. Appl. Sci. Eng. Technol., 5(2): 513-518, 2013.

[9] Suspicious Lesion Detection in Mammograms using Undecimated Wavelet Transform and Adaptive Thresholding, Abhijit Nayak, Dipak Kumar Ghosh, Samit Ari, 978-1-4673-2818-0/13/$31.00 ©2013 IEEE

[10] A computer-aided diagnosis system for breast cancer detection by using a curvelet transform, Nebi GEDIK, Ayten ATASOY, Turk J Elec Eng & Comp Sci (2013) 21: 1002 { 1014 TUBITAK

[11] Malignancy Detection in Mammogram using Gray Level Gradient Buffering Method, Meenalosini S, Kannan E, International Journal of Cancer Research, ISSN:2051-784X, Vol.47, Issue.1, March 2013.

[12] Particle Swarm Optimization Based Contrast Limited Enhancement for Mammogram Images, Shelda Mohan and T.R. Mahesh, 978-1-4673-4601-6/13 © 2013 IEEE

[13] A Computer Aided System for Breast Cancer Detection and Diagnosis, Hamada R. H. Al-Absi, Brahim Belhaouari Samir & Suziah Sulaiman

[14] Mass Detection in Digital Mammograms System Based on PSO Algorithm, Ying-Che Kuo, Wei-Chen Lin, Shih-Chang Hsu & An-Chun Cheng, 978-1-4799-5277-9/14

[15] Detection of Mammograms Using Honey Bees Mating Optimization Algorithm (M-HBMO), R.Durgadevi, B.Hemalatha & K.VishnuKumarKaliappan, 978-1-4799-2876-7/13, 2014.

[16] Isolation of Breast Cancer Accumulation in Mammograms for Improving Radiologists Analysis, Prapti V. Patil, Rode Yogesh & K. V. Kale, ©2014 Engineering and Technology Publishing , International Journal of Electrical Energy, Vol. 2, No. 2, June 2014

[17] Mammography Feature Analysis and Mass detection in Breast Cancer Images, Bhagwati Charan Patel & G. R. Sinha, 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies

[18] WAVELET TRANSFORMATION-BASED DETECTION OF MASSES IN DIGITAL MAMMOGRAMS, P.Shanmugavadivu1, V.Sivakumar2, J.Suhanya, IJRET: International Journal of Research in Engineering and Technology eISSN: 2319-1163 | pISSN: 2321-7308, Volume: 03 Issue: 02 | Feb-2014

[19] Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images, Yasmeen Mourice George, Hala Helmy Zayed, Mohamed Ismail Roushdy, and Bassant Mohamed Elbagoury, 10.1109/JSYST.2013.2279415, IEEE SYSTEMS JOURNAL 1932-8184 © 2013 IEEE

[20] Dual Modality: Mammogram and Ultrasound Feature Level Fusion for Characterization of Breast Mass, Minavathi, Murali .S, Dinesh M. S., International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-2 Issue-6, May 2013.