Classical and Novel Diagnosis Techniques for Early Breast Cancer Detection – A Comparative Approach

T.Venkat Narayana Rao
Research Scholar JNTUK and SNIST, Hyderabad INDIA

A.Govardhan
SIT, JNTUH, Hyderabad, INDIA

Abstract: Breast irregularity is defined with broad variety of features and could be easily missed out or misread by the radiologists while reading huge number of mammographic images supplied in screening methods. Breast cancer is a lone and oldest common malignant disease among women. It is very vital to provide patients early diagnose and action. Since this disease has been identified, studied and analysed but yet still this disease is deemed dreadfull of all times. In the present medical science, there are plenty of newly developed methods and algorithms to obtain timely discovery of breast cancer. Most of these techniques make use of highly accpeted technologies such as medical image processing. This paper is a study and an attempt to emphasize on availability of breast cancer detection methods based on image processing. This paper also offers an outline on affordability, feasiblity, reliability and broad details of results of each technique.

Keywords: Diagnosis, mammogram, mass, Micro calcifications, ultrasound, wavelet, MRI.

I. INTRODUCTION

Breast Cancer is hghly predominant in women in today's world it can start in the breast and can speard to other areas of the body in the course of time. This disease is typically obvious in women but men can also get affected from it, based on recent reports. USA as a nation has recorded forty two thousand cases in a year time. The data emreged has shown drastic increase in the cases and so far, plenty of research is being done to detect the cancer at the earliest. Medical research on breast cancer is not new but lack of proper methods for early detection is still a challenge to medical domain and still it is considered as the deadliest ailment. With the advancement in medical field and specifically the contribution of information technology in the medical field has introduced a new diagnosis dimension termed as Medical Image Processing. Research efforts in image processing have gained abundance of drive in the past two decades. Now image-processing techniques have found their way into computer vision, image compression, image security, medical imaging and more. Medical Image Processing has a unique feature of diagnosing not only cancer disease but also in diverse fields such as lungs, chest and brain . By the virtue of image processing techniques, it has become very simple to detect cancerous mass from an infected breast. Early detection can help in appropriate diagnosis and action resulting in minimizing the risk of unnecessary outcome such as death. Medical image inherently have different nature and need hence, an exceptional processing before diagnosis. There are number of sources, which can produce images of the human body parts. Different mechanisms are required to deal with particular type of image, which lead to classification, analysis and process. This study focuses on various breast cancer detection methods along with several influencing factors. It is an attempt to offer the accuracy, performance and feasiablity matrix of the depicted techniques. It is found that every technique is sole in its character and aim at special kind of state and merits as shown in table I. The algorithms used on such images would identify abnormal masses to simplify additional investigation later[12][5].

A. Mammography: X-rays are the key tools used for captering images and is used in the medical systems for mammographic images. These images are specially recorded for processing, examination and analysing the human breast. These mammgraphic images are useful to the radiologists in investigation for drawing out conclusion relating to the possiblity of presence of abnormality. Mammography can detect several types of breast cancer and is widely used, as it is cheaper and less complicated. Detection of abnormal masses is quite probable by mammography and assures detection of 93% - 97% of breast cancers. The breast part image is extracted and processed and is visualised on a film. Radiologist resort for optimization of images for faster investigation. All the observations are subjcet to patients age, heridity and constraints hence, the study differs from patient to patient[1][4].

B. Preprocessing: Raw mammogram images are complex to understand hence, the images are subjected to enhancement and precise extraction for further segmentation, analysis and classification is performed. Preprocessing steps comprise of removal of unnecessary or immaterial areas and to create region of interest by enhancing the contrast. This is accomplished by setting a threshold rate.

C. Post Processing: At this level the pre-processed

mammogram image is segmented into pixels of small matrices of 2x2 , after which all pixels values of the block are examined and the value representing maximum occurrence in the block is assigned to all pixels of the block. Based on the logic this value is disseminated to rest of the pixels of the blocks. This ensures that the entire block consists of same pixel values. The homogenous blocks are combined to form a 4x4 pixel blocks based on the requirement for making the algorithm to cover the entire image. In the similar lines 8x8 pixel block is also constructed.This process of combining the blocks continue till the image pixles exhaust and are fully assigned. Post highlighting the area the segmentation of breast image is done by implementing color quantization method. Each segmented part is coded in a way that represent regions by particular traits. The algorithms are applied to compute vital parameters on the mammogram test the images. The results are used to construct histograms in MATLAB through which a comparison is made between a healthy and a malignant \breast . Based on above parameters, such as sensitivity, specificity and true positives etc. can be computed for locating area of the region or the set of pixels fall under suspision[1][2][8].

II. Popular Algorithms on Mammography

*A. Vector Quantization :*The method of quatization and texture analysis is known as vector quantization. This mecahnism uses segmentation of images that implement creation of codebook and training vector by isolating the whole image into minute boxes of same size. LindeBuzo and Grey Algorithm (LBG) has propose this method for vector quantization categorization that transforms full image to training vectors. A centroid is determined and Euclidean distance is calculated between training vectors and with this 8 clusters and a codebook of 128 size are created. The technique has facilities to reconstruct the image for further investigation also.

B.Tumor Identification with 3D Structure Analysis: Now a day's 3-dimensional digital mammography and digital spot imaging following the analysis of abnormal obsevations of screening mammograms have emerged as a novel approach. Digital 3-D imaging i.e. Tuned Aperture Computed Tomography (TACT), Digital Spot Imaging (DSI) examinations was found to satisfactory. All patients were retested with this approach as it is not possible to reject the presence of breast cancer on screening films. The 3-D softcopy reading in all cases was performed with Delta TACT mammography computer unit, while the film images were read using a mammography-specific light box. During the softcopy reading of images by both the methods it is found that the diagnostic image eminence of digital 3-D and digital spot images are superior than that of film images but the cost of proceesing 3-D anaysis is much more than digital mammography method[6][4].

C. Digital mammography: Mammography can be classified as film mammography and digital mammography. Digital mammography has better image superiority when compared to film mammography. Mammography is a valuable and cheaper imaging technique for early detection of masses and microcalcifications. The computer assisted investigative software can be used to study the mammograms and to detect the small calcium deposits. On mammogram, several enhancement techniques can be applied. The Full-field digital mammography (FFDM) has several benefits than screen-film mammography (SFM). Digital technology also provide opportunities for implementation of advanced applications, computer aided detection (CAD) and tomosynthesis. Until very recently, some countries did not even permit the use of digital mammography in breast cancer screening. The reason for this is slow growth in the usage of digital mammography has apprehension about resolution, exposure syndrome and disbelief in digital copy reading. Nevertheless, there is a rapid conversion to digital mammography in breast cancer screening in many western countries.

D. K-Means and Fuzzy logic Means : There are some techniques such as K-means and fuzzy logic, which deal with, both supervised and unsupervised approaches. Both methods use the labeling and clustering of input image. The unsupervised method deal with disconnected regions of the image and labeling them as segments. In supervised K-nn means method Euclidean and standard deviation are applied to segment the image. The unsupervised pixel is marked as zero or 255 depending upon the result of K-nn method. This ensures that the outcome illustrates whether the pixel is tumor tissue or not malignant tissue. Fuzzy logic is well suited to represent and to manipulate data/ knowledge at different levels of the algorithm. The fuzzy system is inspired by human sytem of learning and training. This process consists of rules, functions and parameters . The system operates on well-matched fuzzy members and train by supervisory back propagation method and the parameters and rules are optimized automatically, but if necessary, parameters can be set up manually in every stage. In fuzzy logic based system, firstly the system should be identified. The basic structure includes number of inputs, outputs, fuzzy conditional rules and form of member ship functions. Number of inputs is considered as equal as the dimensions of applied feature vector number of outputs of the system equals 1 in every state. This means that only one neuron is considered in outer layer and the system is trained in a way that the output figure of the neuron equals to the number of desired class. The detection and segmentation of the individual microcalcifications, measurements on the segmented microcalcifications such as shape, contrast, relative localization use measurements as inputs of a learning system which concludes if the current case is malignant or not.

E. Threshold and Multi Wavelet techniques : A well technique such as multi wavelet with hard threshold is also successful which proposes to make better ontrast of the digital mammographic images and to de-noise or to remove blur in the images. This would help the physician to do better

diagnosis on the digital mammograms. The multi wavelet method is used as it offers flexibility in computing parameters with a facilty of image de-noising. All radiological images contain fluctuations, which make it tough to detect small structures and abnormalities. Several mechanisms have been proposed for the computation of the threshold in image de-noising yhe bands in wavelet transform. The area of multi wavelet transform, which has high frequency, undergoes de-noising. Original image is taken out by inverse multi wavelet transform lastly. For the evaluation of technique, the errors are computed between the original and the transformed image. It is found that multi wavelet technique has produced efficient results for physicians to diagnosis breast tumor at the initial stage. In addition, this method can produce best peak signal to noise ratio (PSNR) value.

*F. Thermal Infrared Images :*Earlier literatures had concentrated on mammography imaging technique for the breast tumor detection instead of X-rays as these rays are harmful for the body especially for women of age above 40 years and the excessive exposure of body parts to radioactive rays may result in dangerous health problems. Contrray to this some aouthors have proposed an advanced way of detecting breast cancer in which the body is not exposed to radiations and cancer can be diagnosed by means of thermal indicators at infrared images. On these images, image-processing techniques along with computer artificial tools are implemented. The phenomenon behind using thermal indicators or infrared image and metabolic activity increases in cancerous cells instead of normal cells. All the objects in the universe emit infrared radiations, as a function of temperature. This kinde of thermal imaging technique is known as thermography. Thermal infrared camera is used in Thermograph. The camera takes a picture of the cancer-infected region, which has to be considered for the investigation. The literature existing has shown that this technique by analyzing the results of experiment involve two stages. The first stage uses Asymmetry method and f thermal image development.

*Asymmetry: T*his method is known as asymmetry because the left and right breasts are asymmetric to each other. First, the primary or RGB image is transformed to grey image by means of the function rgb2grey in MATLAB. After transformation is done, the next task is to segment the breast area for which the parabolic Hough transforms is applied. To perform this grey image is converted to edge detected image by applying the edge operator.

Thermal Image Development: The pixels in thermogram image represent thermal radiation of temperature of the human body. Various tissues and organs from which one can identify the abnormality produce dissimilar amounts of radiation. For this the pixel intensity of thermal image technique is applied and later K-means and C-means use the logic of fuzzy methos are applied to further examine the image[14].

*G. Automated Image Segmentation :*In this approach, the top body edges are manually removed. After that a canny edge detectors series is applied. The main step is to lower the breast boundaries detection. To accomplish this canny edge operator is applied repeatedly and checks the effects by varying threshold so that only bottom breast boundary is left. Threshold is selected by means of an algorithm known as automatic threshold selected filters. After the edges are detected, the left and right edges are separated and Hough transform mechanism is computed so that lower breast boundary can be extracted.

*H. Analysis through Artificial Neural Network:*Previously, due to deficiency of image processing tools and image acquisition equipment, results varied and were found to be inconsistent. Now automated tools are available for breast analysis. Like every algorithm, the proposed methodology consists of phases namely: i. Image acquisition. ii. Image preprocessing. iii. Image analysis with Artificial neural networks (ANN). In this approach, the images are acquired and preprocessed by identifying exterior boundaries of the breast. Edges are identified by applying the canny edge operator clips these edges further. The breast image is applied with the morphological operators by considering the breast shape of two convex or a smallest ellipse. The segmented breast is further divided in to four quadrants according to proposed division[10][13]. The idea behind this observation is to display progesterone and estrogen receptors existence in the cells infected by tumor. This demonstration process is carried out by applying the concepts of neural networks. The steps are followed to achieve the task i.e. i. Preprocessing, ii. Image segmentation, iii. Segmentation of cancer cells, iv. Object classification and, v. Feature extraction. The cells after applying to the neural network are classified into two classes on the basis of size and size i.e., blue (N) and brown (P). Brown color means positive, that is, cancer cells are present and blue color represents a negative result. Pixel values for background, P, or N regions respectively are -1, 0 and 1 from the neural network. The local adaptive threshholding and morphological operators for the breast tumor cells66 designate segmentation. Lastly the removal of spike noise is accomplished by morphological erosion and opening assuming a structuring element of disk shape.

*I. Watershed Segmentation:*The algorithm is based on components of discontinuties and Fourier transforms. In the initial phase, segmentation is done through morphological mathematical operations. In the subsequent, phases the feature extraction is done through Fourier descriptors. The parameters for extraction are color and shape. The final phase contributes classification and labelling of tumor cells. Closing and opening operations are performed in the pre-processing phase by applying a structuring element of disk shape. In watershed method the preprocessing is done to remove unwanted and noise areas from the image while conserving the boundaries around the center. The success rate of this algorithm is almost 80%[6][3].

*J. Ultrasound Images:*Ultrasound is the second most common method that is used to detect breast cancer in an early stage. The ultrasound image is taken as an input in the technique explained and then the image is filtered using anisotropic diffusion algorithm for the removal of speckle noise. Edges were enhanced by using un-sharp masking technique. Image segmentation is performed by using normalized cut (N-cut)

technique. This technique is applied to detect the lesion region. The image is cut into multiple sets. After detection of lesions in every set the grey image was converted into binary one and morphological closing operations are applied to close the boundary of lesion. The closed lesion region is then extracted.

K. Support Vector Machine : SVM is a statistical tool and has emerged as a prominent means to deal with classification. This method proposes a classifier, which is based on learning machine and statistical theory. The merits of SVM over other classifiers are that its computations are easier. To develop SVM performance a bootstrap learning technique is performed. A classification task usually engages training and testing of data. Every object in the training set consists of one target value also termed as "class labels" and attributes also termed as "features". A support vector machine is a machine-learning tool that classifies binary classes by finding and using a class boundary the hyperplane maximizing the margin in the known training data. The training data samples along the hyperplanes close to the class boundary are called support vectors. The margin is the distance among the support vectors and the class boundary hyperplanes. The SVM are based on the concept of decision planes that identify decision boundaries.

L. MRI Imaging: To diagnose disease magnetic resonance method utilizes magnetic and radio waves. The patients are subjected to exposure to the MRI machine for atleast half an hour to capture the required images. They are then advanced into the MRI machine, which contains a very high magnetic field. The method consists of inducting contrast-enhancing dye like material into the patient's bloodstream and using magnetic resonance imaging to monitor the way in which this material is taken up and responded by the tumor tissue. This would enable the expert to differentiate between tumor or suspicion fraction/mass in the body. The aptitude to identify a mass in the breast requires that the mass have a dissimilar appearance (or a different contrast) from the normal tissue. The MRI, can be many times informative than X-ray mammography and has capability to showcase over 96% image capturing capabilities. But MRI is not cheaper and most of the population in the developing countries cannot afford to pay and undergo such examination frequently[7][9].

CONCLUSION

This paper has outlined the methods and algorithms proposed for the detection of breast tumor and understanding the stage of the mass, so that proper advise can be offered to the cancer patient. Digital mammography technique is widely used for early stage breast cancer diagnosis and is found to be cheaper. The other methods such as MRI and Artificial Neural Networks etc. are also proposed with their merits and demerits. The detailed comparative statement is depicted in Table I and it is suggested based on the applicability, suitabilty, cost and feasibility a detection method can be chosen. The aim of this article is to offer an updated review of these studies, converse the conflicting findings, and draw some conclusions from the results.

REFERENCES

[1] Indra Kanta Maitra, Sanjay Nag, Prof. Samir Kumar Bandyopadhyay, Identification of Abnormal Masses in Digital Mammography Images, International Journal of Computer Graphics, 2(1), (2011)

[2] Kekre H.B., Sarode Tanuja K. and Gharge Saylee M., Tumor Detection in Mammography Images using Vector Quantization Technique, International Journal of Intelligent Information Technology Application, 2(5), 237-242 (2009)

[3] Ball JE, Digital mammogram speculated mass detection and spicule segmentation using level sets, Proceedings of the 29th Annual International Conference of the IEEE EMBS, 4979-84 (2007)

[4] Khuzi A Mohd, Besar R, Wan Zaki WMD, Ahmad NN, Identification of masses in digital mammogram using gray level co-occurrence matrices, Biomed Imaging IntervJ, 5(3), 1-13 (2009)

[5] Bartakke P.P., Vaidya S.A. and Sutaone M.S., Refining structural texture synthesis approach, Image Processing IET, 5(2) 184-189 (2011)

[6] Yan Zhang, Xiaoping Cheng, Medical image segmentation based on watershed and graph theory, Image and Signal Processing (CISP), 3, 1419-1422(2010)

[7] Draa A. and Meshoul S., A Quantum Inspired Learning Cellular Automaton for Solving the Travelling Salesman Problem, International Conference on Computer Modeling and Simulation (UKSim), 45-50 (2010)

[8] Renjie Liao, Tao Wan, Zengchang Qin, Classification of Benign and Malignant Breast Tumors in Ultrasound Images Based on Multiple Sonographic and Textural Features, Intelligent Human-Machine Systems and Cybernetics (IHMSC), 71-74 (2011)

[9] Alina Sultana, MihaiCiuc, RodicaStrungaru and Laura Florea., A New Approach in Breast Image Registration, International Conference on Intelligent Computer Communication and Processing, 149-154 (2010)

[10] Waqas Haider, Muhammad Sharif, Mudassar Raza, Achieving accuracy in early stage tumor identification systems based on image segmentation and 3D structure analysis, Computer Engineering and Intelligent Systems, 2(6), (2011)

[11] Punal.M.Arabi, S. Muttan, R.Jenkin Suji, Image enhancement for detection of early breast carcinomaby external irradiation, IEEE International conference on Computing, Communication and Networking Technologies, 1-9 (2010).

[12] Kother Mohideen, Arumuga Perumal, Krishnan and Mohamed Sathik, Image De noising and Enhancement Using Multi wavelet with Hard Threshold In Digital Mammographic Images, International Arab Journal of e-Technology, 2(1),(2011) .

[13] Essafij, R. Doughrij,s. M'hiri k. Ben romdhane and f. Ghorbelm, Segmentation and classification of breast cancer cells in histological images,International Conference of the IEEE Engineering in Medicine and Biology Society, (2010).

[14] Teshnehlab M., Aliyari Shoorehdeli M. and Keyvanfard F., Feature selection and classification of breast MRI lesions based on Multi classifier, International Symposium on Artificial Intelligence and Signal Processing (AISP), 54-58, (2011) .

Table-1 : Comparison between different diagnosis techniques on mammographic images

Technique/ Algorithm	Advantages / Characteristics	Average Accuracy %	Limitations	Feasiblity and applicability
Vector Quantization	Absence of over and under segmentation	76-78%	Lot of segmentation areas and information	Moderate
3D Structure Analysis	Exact size and thickness of tumor can be calculated	Upto 99%	CAD system has to be combined with the segmentation algorithm to get highest success rate	Low
Digital mammography	Calcification, circumscribed, speculated and other ill-defined masses can be diagnosed	93-98%	Frequent intensity changes may not give 100% diagnostic results	High
K-Means and Fuzzy logic Means	Change of intensity is used as a discriminating feature	Upto 85%	Lot of segmentation areas and information	High
Multi Wavelet techniques	Gives good results in case of dense mammograms by noise cancellation	87%	-	High
Thermal Infrared Images	Tumor is diagnosed depending upon the radiation	85%-90%	May give false results in case of flat	Low
Automated Image Segmentation:	Only lower breast part is taken into account and upper body parts are filtered out	Upto 96%	Some details are compromised during study	Low
Artificial Neural Network	Upper body parts are filtered out for a better analysis.	Upto 98%	Higher programming skill needed for implementation	High
Watershed Segmentation	Division of image on the basis of discontinuities can be useful for better analysis	Upto 88%	More accuracy and validation is needed in case of larger histological slides	Moderate
Ultrasound Images	No radiation is used and extremely small pressure is required	90-92%	Difficult to implement	moderate
SVM Algorithm	SVM achieves this advantage by focusing on the training examples that are most difficult to classify.	Upto 89%	Highly programming based	High
MRI	MRI may decrease unnecessary exposure to ionizing radiation in the women	97%-99%	Costly method and high exposure to the radiation	High

