ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

A Survey on Various Procedures Implemented in The Early Detection of Breast Cancer

Sunil Subrahmanyam Yadavalli
M.S (Software Engineering) VIT University, Vellore, Tamil Nadu, India

Abstract
Breast cancer is the main cause of cancer related mortality among women. Mammography is the most reliable method for early diagnosis of breast cancer and is widely recommended with the introduction of several Computer Aided Diagnosis (CAD) techniques. We are presenting some of the pattern recognition techniques that have been most effective in detecting the micro calcifications in digital mammograms. These methods have been developed and implemented in statistical knowledge hypothesis over the past ten years and they gave hopeful results in the tumor diagnosis. Early detection is the best hope for reducing the death rate caused by the invasive disease. Micro-calcifications are usually associated with extra cell activity in the breast tissue. The extra cell activity does not have to be cancerous and it usually is not, but if micro-calcifications are grouped in clusters, it can be referred as a malignant tumour. X-ray mammography is generally considered to be the best for early detection of breast cancer.

Keywords
Microcalcifications, Wavelets, Mammogram, Breast Cancer, Roi (Region Of Interest)

I. Introduction
Breast cancer is the most common and deadly of all cancers. Mammography is a unique type of medical imaging used to screen the breast for identification of the disease. A typical mammogram is an intensity X-ray image with showing levels of contrast inside the breast. It characterizes normal tissue and different calcifications that are identified during the screening process. Currently, X-ray mammography represents the effective method of breast imaging. Other methods like Magnetic Resonance Imaging (MRI) and ultrasound approaches, are as yet either less effective or too expensive for detection purposes. Mammography is a specific kind of imaging method that uses a low-dose X-ray system to examine breasts. It is performed on an outpatient basis. Currently there is consensus on the fact that the mammography of tracking may reduce mortality by breast cancer at least 40%. Therefore a computer system has been developed in assisting radiologists to pre-screen the mammograms.

II. Diagnosis tools
The diagnosis task is classified as a two-class classification task. The features are extracted from Regions of Interest (ROIs - the region containing the masses or micro-calcifications). The ROI containing the abnormality can be identified by detection algorithms. Each ROI is classified using a classification algorithm. The classification method is trained on a set of sample images (whose classification is known) at first is called the training set. The performance of the algorithm is then tested on a separate testing set. The metrics used to report the accuracy of these algorithms are sensitivity and specificity. Sensitivity is defined as a lesion for which the CAD predicts that it is cancerous and it is actually found to be malignant. Specificity is the fraction of benign lesions that are correctly identified by the detection algorithm. A graph between sensitivity versus specificity is called a Receiver Operating Characteristic (ROC) curve and this is used to report the performance of the CAD technique used. Significantly the area under the ROC curve, Az is mainly focused. The higher the value of Az, the better is the performance of the CAD technique used. A good CAD technique has Az values closer to one.

A. Mass detection
A mass is space occupying lesion seen in at least two different angles defined with a several range of features. The mass can indicate a benign change which is not to say that it does'nt represent a malignant change [1]. Masses with round, smooth and limited margins usually indicate benign changes while masses with small points, rough and blurry margins indicate the malignant part. Some researchers have focused mainly on the detection of these masses because of their high likelihood of malignancy [4]. The algorithms for cancer detection in digital mammography generally consist of several steps: segmentation, feature extraction, feature selection and classification.

B. Reason for micro-calcification detection
Mass detection is a more difficult problem when compared to micro-calcifications cluster detection because masses are often of varying size, shape, and density. They exhibit poor contrast, highly connected to the surrounding tissue, particularly for small pointed lesions, surrounded by non-uniform tissue background with similar characteristics. Microcalcifications are tiny in size and early sign for breast cancer. Several computer aided techniques in digital mammography today gives more accurate results in case of detecting the clustered microcalcification [2][3]. Hence due to the above reasons micro-calcification detection is preferred and we investigate the following different methods.

C. Micro-calcification detection algorithms
Calcifications are calcium deposits inside the breast. They are divided in two major groups: macro-calcification and micro-calcifications. Macro-calcifications are as expected, large calcium deposits, whereas micro-calcifications are tiny ones. Macro-calcifications are not usually linked with the development of breast cancer. Hence no special attention is being given to them. On the other hand, detection of micro-calcifications is very important for the early breast cancer detection. Scattered micro-calcifications are usually a part of benign breast tissue. In mammograms calcifications are generally seen as bright dots of different sizes. The exact position and number of micro-calcifications cannot be easily predicted. They can also be grouped in clusters, but more often they found to be alone. Detection of micro-calcifications is

a challenging task for radiologists as well as for computer-aided detection software. Independent diagnostic testing facility (IDTF) refers to a radiologic procedure furnished to a woman with signs or symptoms of breast disease, a personal history of breast cancer or a biopsy-proven benign cancer. Screening mammography refers to a radiologic procedure furnished to a woman without signs or symptoms of breast disease, for the purpose of early detection of breast cancer. The main principle in X-ray mammography is that tumor tissue exhibits micro-calcifications, which affect mammogram formation. Mammogram is 2-D map which represents the intensity of X-ray radiation passed through the compressed breast previously [7]. However, some small contrast between affected and normal tissue for X-rays leads to high false-negative (4–34%) and false-positive (70%) rates. Also, very early stage tumors do not need to exhibit micro-calcifications. One should also notice that ionizing radiation is accumulated over repeated scans. Computer-aided diagnosis system provides a scientific method, which has the advantage to process vast amount of X-ray images in small amount of time, to doctors for valuable comparison and objective double-check. Though number of computer-aided diagnosis systems have been reported, yet how to upgrade the sensitivity and reduce computation burden or time consumption is still considerably required [8]. In order to enhance the pixels of micro-calcifications in the image, partial wavelet's coefficients reconstruction method is applied to suppress the background noise. The Mixed Feature Based Neural Network (MFBNN) was utilized to upgrade the pattern Recognition performance. Although more descriptors can enhance the accuracy of recognition on micro-calcification, a few drawbacks such as heavier computation load and more complexity on the diagnosis algorithm are induced at the same time. Biopsy and magnification are able to confirm the presence of micro-calcifications and can accurately evaluate the rate of false positives. But these tools are not able to determine the number of micro-calcifications and to localize them individually. Interpretation by human beings is highly subjective and the problem of reproducibility is rarely mentioned. In order to eliminate the above complexities of different algorithms, transform techniques are introduced.

KLT (Karhunen-Loeve Transform)

KLT is the unique transform which decorellates its input. To be precise, define the covariance matric CX of a random vector X by

$CX = E((X - \mu x)(X - \mu x))$

If the KLT transform is called K, then the random vector Y = K_X should have uncorrelated components, i.e. CY = K_CXK is a diagonal matrix. This transform gives rise to principal component analysis (PCA). Among linear transforms, when principal components are ranked in decreasing order KLT minimizes the mean square error (MSE). The drawback for the KLT is that we need to recompute the transform each time when the statistics of the source changes. It cannot be separable by its nature.

Fourier Transform (FT)

The Fourier Transform provides the means of transforming a signal defined in the time domain to a defined signal in the frequency domain. When a function is evaluated by numerical procedures, it is always necessary to sample it in some manner. This means that in order to fully evaluate a Fourier transform with digital operations, it is necessary that the time and frequency functions are to be sampled in some form or another. Thus the digital or Discrete Fourier Transform (DFT) is considered to be better than Fourier Transform.

Discrete Fourier Transform (DFT)

Discrete Fourier Transform has an efficient implementation than the FT. The associated basis functions are complex sinusoidal waves that represents to the eigenvectors of a linear time-invariant operator. One drawback of DFT is that the transform works badly when the end points ($x0$ and $xm-1$) are far away from each other. If the full Fourier transform was applied in this case, many higher Fourier components would be introduced to compensate for this. Fourier transform gives only frequency content but cannot localize objects of certain frequency in the image.
In medical image processing applications, we usually deal with discrete data. We will therefore focus the rest of our discussion on discrete wavelet transform rather than continuous wavelet transform.

Discrete Wavelet Transform (DWT)

Discrete Wavelet transform is used to construct time-frequency representation of a certain dataset. As we know that Fourier transform gives only frequency content but cannot localize objects of certain frequency in the image, discrete wavelet transform is therefore better because it gives both frequency content and exact position of the object in the image. Discrete wavelet transform (DWT) is commonly used in image processing for combining of multiple images [9][10]. If several lighting sources are combined, the result can be obtained by combining a set of basis images. The combination can be done very fast in the wavelet domain. Transform may also be applied to the full image or only partially. DWT has an efficient implementation, both in lossy and lossless case. Wavelet based processing algorithms comprises good ability in identifying different frequencies [5]. Wavelets have the capability to zoom in and zoom out.

D. Other recently proposed methods

Besides the use of wavelet contrast enhancement methods with noise estimation, other approaches are used to detect micro-calcifications. Sankar and Thomas proposed the method which uses fractal modelling of mammograms based on mean and variance to detect micro-calcifications. This method was tested on 28 mammograms from the MIAS database and produced the following results: TPR=82% with an average of 0.214 negative clusters per image. The results of the fully automatic and the semiautomatic segmentation on images from the DDSM base are evaluated using areas under the ROC curve. This method gave results of A=0.80 while the semiautomatic gave significantly higher results of A=0.84. Another micro-calcification detection method presented in 2009 uses a different approach [13]. The first stage of the process is extraction of zones that correspond to micro-calcifications by analysing the distribution of brightness over the mammogram. The second stage is identification of clusters as ROIs. The final stage is retrieving the information that might lose in the previous stages. This method is also tested using DDSM database.

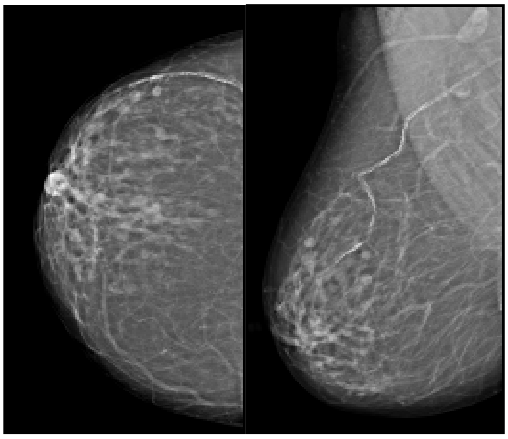
Support vector machines (SVM)

Cancer can either be cancerous (malignant) or non-cancerous (benign). Malignant tumours penetrate and destroy healthy body tissues. Support vector machines have become an increasingly popular tool for machine learning tasks that involves classification.

The SVM algorithm performs excellently in exhibiting high accuracy, sensitivity and specificity. The use of the classifier systems in the medical diagnosis area is gradually growing and there is no doubt that evaluation of data taken from patients and decisions of the experts are the most important factors in diagnosis [12]. Another method demonstrated that the prediction power could be further improved by applying a successive enhancement learning (SEL) procedure, where SVM training is adjusted iteratively by reincorporating misclassified samples.

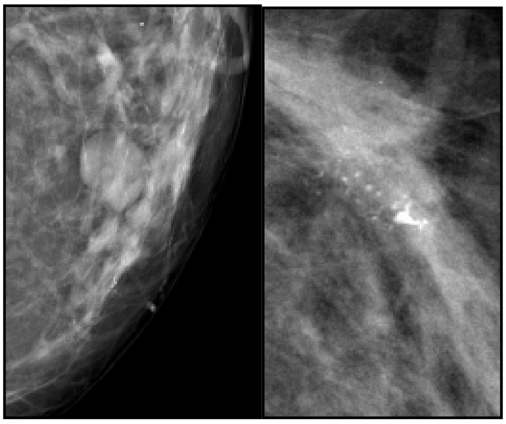
III. Images and photographs

Two basic views of mammographic image: (a) craniocaudal (CC) view, (b) mediolateraloblique (MLO) view.



a b

Examples of abnormalities: (a) mammogram with round mass, (b) mammogram with clustered calcifications



a b

IV. Comparision tables & plots

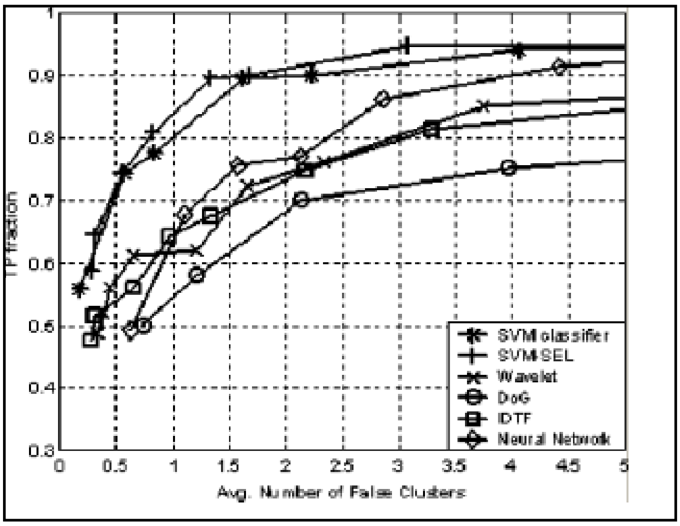
Performance measure of classification of malignant and benign mammogram for DWT feature

Technique	Effici-Ency (%)	Sensi-tivity	Speci-ficity
NN+DWT	93.94	91.3	94.3
Baysian+DWT	95.11	94.1	90.2
KNN+DWT	95.87	92.7	92.7
SVM+DWT	96.13	91.3	91.3

Results with DWT features

Dataset	ANN	KNN	SVM	Baye sian
MIAS	93.94	95.87	95.11	96.13
real mammograms	94.3	96.31	95.52	95.31

FROC comparison of different methods for the detection of MCs [6]:



V. Conclusion

Digital mammography screening programs can enable early detection and diagnose of the breast cancer which reduces the mortality and increases the chances of complete recovery. Screening programs produce a great amount of mammographic images which have to be interpreted by radiologists. Due to the wide range of breast abnormalities features, some abnormalities may be missed or misinterpreted. There is also a number of false positive findings and therefore a lot of unnecessary diagnosis failures. There are a lot of algorithms developed for detection of masses and micro calcifications. Over the years there has been an improvement in the detection algorithms but their performance is still not perfect. The area under the ROC curve is rarely above 90% which means that there are still many false positive outputs. Possible reasons for such a performance are the characteristics of breast abnormalities [11]. Masses and micro calcifications are sometimes hidden in the dense tissue which makes the segmentation of correct regions of interest difficult. Another issue is extracting and selecting appropriate features that will give the best classification results. Further developments in each algorithm step are required to improve the overall performance of computer aided detection and diagnosis algorithms. Semiautomatic method is giving the area under the ROC curve value nearer to 1 which means it is more effective than automatic method.

VI. AcknowledgEment

References

[1] *Prof. Samir Kumar Bandyopadhyay Professor, Dept. of Computer Sc. & Engg, University of Calcutta digital Imaging in Mammography towardsDetection and Analysis of Human Breast Cancer*

[2] *Sankar, D., Thomas, T.: Fractal Modeling of Mammograms based on Mean and Variance for the Detection of*

Microcalcifications. In: Proceedings of the 2007 International Conference on Computational Intelligence and Multimedia Applications, Sivakasi, India, December 2007, pp. 334–338 (2007)

[3] *Acha, B., Rangayyan, R.M., Desautels, J.E.L.: Detection of Microcalcifications in Mammograms. In: Suri, J.S., Rangayyan, R.M. (eds.) Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer. PIE, Bellingham(2006)*

[4] *Varela, C., Tahoces, P.G., Méndez, A.J., Souto, M., Vidal, J.J.: Computerized Detection of Breast Masses in Digitized Mammograms. Computers in Biology and Medicine 37, 214–226 (2007)*

[5] *Enhancement and Identification of microcalcifications in Mammqgram Images Using Wavelets Songyang Yu', Stephen Brown', Yu Xue', and Ling Guan1.1 Department of Electrical Engineering, University of Sydney, NSW 2006, Australia.2 Ultrasonics Laboratory, Division of Radiophysics, CSIRO, NSW 2067, Australia.*

[6] *Rangayyan, R.M., Ayres, F.J., Desautels, J.E.L.: A Review of Computer-Aided Diagnosis of Breast Cancer: Toward the Detection of Subtle Signs. Journal of the Franklin Institute 344(3-4), 312–348 (2007)*

[7] *Shen L, Rangayyan RM, Desautels JEL. Application of shape analysis to mammographic calcifications. IEEE Transactions on Medical Imaging 1994; 13(2):263–74.*

[8] *Zheng B, Qian W, Clarke LP. Digital mammography: mixed feature neural network with spectral entropy decision for detection of microcalcifications. IEEE Transactions on Medical Imaging*

[9] *Nan-Chyuan Tsai, Hong-Wei Chen, Sheng-Liang Hsu. Computer-aided diagnosis for early-stage breast cancer by using Wavelet. National Cheng Kung University. 09 August 2010.*

[10] *Oyvind Ryan. Applications of the wavelet transform in image processing. Department of informatics, University of Oslo. 12 Nov 2011.*

[11] *Long, P.M., Servedio, R.A.: Boosting the Area under the ROC Curve. In: Advances in Neural Information Processing Systems 20, Conference Proceedings (December2007)*

[12] *Nawazish Naveed, Muhammad Arfan Jaffar, Faisal Karim Shaikh. Malignancy and Abnormality Detection of Mammograms using Classifier Ensembling. Mehran University Research Journal of Engineering and Technolgy. Volume 30, no.3, july, 2011.*

[13] *Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography: Recent Advances Jinshan Tang, Senior Member, IEEE, Rangaraj M. Rangayyan, Fellow, IEEE, Jun Xu, Issam El Naqa, Member, IEEE, and Yongyi Yang, Senior Member, IEEE*

Author's Profile

Sunil Subrahmanyam Yadavalli, M.S (se), VIT University. He participated in national seminar on medical thermography (NSMT) organised at SRM University, Chennai on 21st and 22nd Sep, 2011. Had been an active member of CSI - Computer Society of India, VIT Students Branch for two years. He presented a paper entitled Detection of micro-calcifications in digital mammograms for early detection of breast cancer in the 4th international conference on "science, Engineering and Technology (SET)" organized by school of IT and Engineering, VIT University, Vellore on 3rd and 4th may, 2012.