BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, August 2014. Vol. **11**(2), 907-910

Computerized Breast Cancer Detection System

R.J. Hemalatha , G. Hari Krishnan, G. Umashankar and Sheeba Abraham

Department of Biomedical Engineering, Sathyabama University, Chennai, India.

doi: http://dx.doi.org/10.13005/bbra/1357

(Received: 15 April 2014; accepted: 05 June 2014)

Breast cancer is a leading cause of death among women, and its incidence is rising. Mammography has been shown to be effective in screening asymptomatic women to detect occult breast cancer and to reduce mortality by as much as 30% in women aged between 50 and 69 years. Our objective is to develop a CAD system to automatically detect, analyze, and to classify the different features in mammographic images through image processing technique. The feasibility of the proposed approach was explored on the images, and the sensitivity rate is 89% and the specificity is 93%.

Key words: Mammography, asymptomatic, breast cancer, CAD.

Breast cancer, i.e., a malignant tumor developed from breast cells, is considered to be one of the major causes for the increase in mortality among women, especially in developed countries. More specifically, breast cancer is the second most common type of cancer and the fifth most common cause of cancer death[1]. According to the World Health Organization, more than 1.2 million people will be diagnosed with breast cancer this year worldwide. Breast cancer was the most common form of cancer and cancer related death in women worldwide. In Asia alone, breast cancer represents 19% of cancer deaths and the 24% of all cancer cases. Nearly 25% of all breast cancer deaths occur in women diagnosed between ages 40 and 49 years. The risk of a woman developing breast cancer during her lifetime is approximately 11%. Early detection of breast cancer is of vital importance to successful of treatment, with the main goal of increasing the probability of survival for patients[2].

Currently, the most reliable and practical method for early detection and screening of breast cancer is mammography. Mammographies are low dose X-ray projections of the breast, and it is the best method for detecting cancer at an early stage. In a recent study, Vacek *et als* work detailed about the proportion of breast tumors that were detected in Vermont (US) by screening mammography increased from 2% during 1974-1984 to 36% during 1995-1999 [3]. Mammography is highly accurate, but like most medical tests, it is not perfect. On average, mammography will detect about 80%-90% of the breast cancers in women without symptoms[4]. When radiologists examine mammograms, they look for specific abnormalities[5]. The most common findings seen on mammography are masses, calcifications, architectural distortion of breast tissue, and asymmetries. Micro calcifications (MCs) can be an important early sign of breast cancer; they appear as bright spots of calcium deposits associated with extra cell activity in breast tissue. MCs are potential primary indicators of malignant types of breast cancer 80% of the MC is benign. The calcifications are small; usually varying from 100 micrometer to 300 micrometer, but in reality may be as large as 2mm[6]. Individual MCs are sometimes difficult to detect because of the surrounding breast tissue and variations in shape, orientation, brightness and diameter. Therefore, their detection can be important in preventing and treating the disease.

* To whom all correspondence should be addressed.
Mob.: +91-9994789165;
E-mail:haris.eee@gmail.com

Recent studies showed that the interpretation of the mammogram by the radiologists give high rates of false positive cases. The images provided by different patients have different dynamics of intensity and present a weak contrast. Moreover the size of the significant details can be very small[7]. In order to improve the accuracy of interpretation, a variety of computer-aided diagnosis (CAD) systems have been proposed.

Several research works have tried to develop tools; they could help the radiologists in the interpretation of the mammograms and could be useful for an accurate diagnosis. A computer aided diagnosis (CAD) system assisting the radiologist could have a tremendous impact by helping to correctly diagnose the missed malignant cases and reduce the number of unnecessary surgical biopsies and other oversights that may result from poor mammographic image quality, radiologist fatigue, or alternative sources[8]. Computer aided detection (CAD) systems in screening mammography serve as a second opinion for radiologists by identifying regions with high suspicious of malignancy[9]. The ultimate goal of CAD is to indicate such locations with great accuracy and reliability. Thus far, most studies support the fact that CAD technology has a positive impact on early breast cancer detection[10-11.]

Methodology

The proposed system follows a hierarchical approach. Initially the CAD system prescreens a mammogram to detect suspicious regions in the breast parenchyma that serve as candidate location for further analysis. In this the first stage is to extract the breast region from the mammogram by suppressing the pectoral muscle region. The second stage is to segment the micro-calcification masses from the breast tissue and analyzing their features for easy discrimination from the normal tissue. The efficiency of the system is based on the following parameters: Sensitivity - Sensitivity (also called recall rate in some fields) measures the proportion of actual positives which are correctly identified as such. Specificity - Specificity measures the proportion of negatives which are correctly identified.

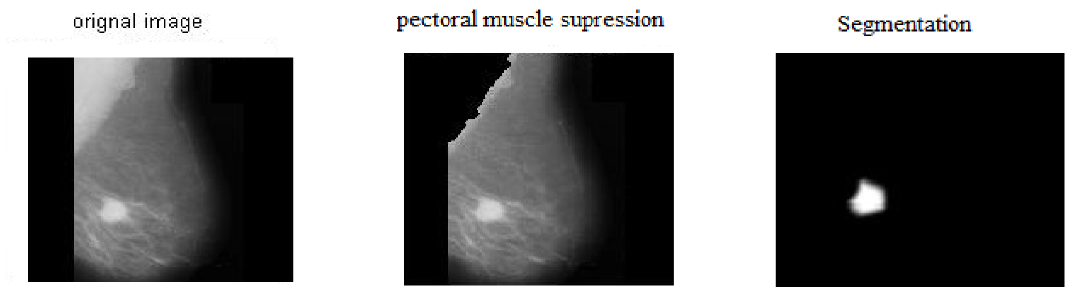
Detection of tumors in mammogram is divided into two main stages. The first step involves a filtering procedure, which is used to remove the noise and enhances the image. Then the removal of pectoral muscles from the mammographic image to extract the breast tissue region. After the mammogram enhancement, segment the tumor area and the features are extracted from the segmented mammogram.

Pre-processing

Mammograms are medical images that are difficult to interpret, thus a preprocessing phase is needed in order to improve the image quality and make the segmentation results more accurate. We have used Gaussian and morphological top hat filtering technique for the better results. After filtering and enhancement the second step is to remove the pectoral muscle from the breast area. There are several approaches for the background segmentation which ranges from histogram thresholding and smoothing, polynomial modeling, active contour approaches, hough transform or gabor filters as edge detectors. In this work a new approach is designed, which involves an automatic thresholding algorithm (separates the area composed of pectoral muscle and breast) along with Connected Component Labeling algorithm to extract the breast portion[12-13].

Segmentation

Segmentation refers to the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. This is



the main stage of a CAD system and by using this we can detect abnormalities such as, masses and microcalcifications in the mammogram. There are three kinds of segmentation methods based on pixel classification: Markov random field (MRF) or Gibbs random field (GRF), region growing, and region clustering. The segmentation technique used for the identification of suspected region is Region Clustering[14]. Region clustering is based on pixel classification method. It searches the region directly without initial seed pixel. The *K*-means algorithm is a well-known clustering procedure. *K*-means clustering algorithm separates the pixels into clusters based on their intensity and their relative location. A clustering algorithm was used for fully automated segmentation[15]. Similar to region growing technique, used a pixel-by-pixel *K*-means clustering method[16,17] for initial mass segmentation. The clustering process separates one or more disjoint objects within the ROIs, which were filled, grown in a local neighborhood, and eroded and dilated by morphological operators.

Feature extraction and classification

The third stage in the CAD (computer aided diagnosis) is the feature extraction and selection. The features are calculated from the ROI based on some characteristics such as the size, shape, density, and smoothness of borders, etc. The feature extraction and selection is a key step in mass detection. Feature selection is the process of selecting an optimum subset of features from the given problem domain after the image segmentation[19].The features are directly calculated from the boundaries and areas of ROIs, Euclidean distance, centroid etc. these features are called shape features. The shape features are also called the morphological or geometric features. In feature classification step masses are classified as benign or malignant using the selected features. Various methods have been used for mass classifications. Some of the most popular techniques are artificial neural networks and linear discriminant analysis.

Artificial neural networks (ANNs): ANNs usually use non-linear mapping functions as decision boundaries. The advantage of ANNs is their capability of self-learning, and often suitable to solve the problems that are too complex to use the conventional techniques, or hard to find algorithmic solutions. It includes an input layer, an output layer and one or more hidden layers between them. Depending on the weight values the inputs are either amplified or weakened to obtain the solution in the best way. There are mainly two types of ANN classifiers for masses: the three-layer back propagation neural network[20, 21] and the radial basis function (RBF) network. As per Varela *et als* work the feature sets are merged into a back propagation neural network (BNN) classifier to reduce the number of false positives. The results yielded a sensitivity of 88% and sensitivity of 94%.

Evaluation of cad performance

In CAD research, the quality of the detection algorithm is usually reported with the miss detection rate, false-positive rate, or similar metrics. The most essential requirement from a radiologist point of view for image processing algorithms is the ability to achieve enhanced visualizations of anatomical structure, while preserving the detail of the structure[22]. There are numerous researches, which worked on the classification and segmentation of glandular tissues. Each classification and segmentation result needs evaluation of its performance. There are three types of performance evaluations. The first type involves qualitative assessment, the second is quantitative assessment involving the ground truth evaluation, and the third is a statistical evaluation. In general, the various terminologies used to determine the performance of a CAD system is defined as follows: *True positive rate (TPR)*: The ratio of the number of malignant cases classified to the total number of malignant case in the tests set. *False positive rate (FPR)*: The ratio of the number of benign cases incorrectly classified to the total number of benign cases in test set. *Sensitivity*: True positive rate. *Specificity*: True negative rate. the sensitivity of the CAD system is 89% and the specificity is 93%.

CONCLUSION

Breast region extraction is a key solution in removing different types of artifact from the mammogram image. It is achieved by automated thresholding method and Connected Component Labelling algorithm. Our proposed method was evaluated on 40 mammograms. The first stage (breast border extraction) gave a rate of 100% in detecting the correct border. Second stage is the detection of microcalcification using region clustering. The statistical features are extracted from the decomposed image and given as the input to the neural classifiers. In the third stage, the image is classified into normal and abnormal. The proposed CAD system gives the tumor detection rate of about 90% accuracy.

REFRENCES

1. R.M. Nishikawa. "Current status and future directions of computer-aided diagnosis in mammography", *ComputerizedMedical Imaging and Graphics*, 2007; **31**(4-5): 224-235.
2. N Pal, B Bhowmick, S Patel, S Pal, J Das," A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms", *Neurocomputing*. 2008; **71**(13-15): 2625-2634.
3. P.M. Vacek, B.M.Geller, D.L.Weaver, and R.S.Foster. Increased mamography use and its impact on earlier breast cancer detection in Vermont. *Cancer,* 2002; **94**(8): 2160-2168.
4. Michaelson J, Satija S, Moore R, "The pattern of breast cancer screening utilization and its consequences", *Cancer*. 2002; **94**(1): 37-43.
5. Sampat MP, Bovik AC, Whitman GJ, et al. "A model-based framework for the detection of speculated masses on mammography". *Med Phys* 2008; **35**: 2110-2123.
6. L.Shen, R.M. Rangaan, and J.E.L. Desautels, "Application of shape analysis to mammographic classifications," *IEEE Trans. Med. Imag*., 1994; **13**(2): 263-24 .
7. Sterns EE, "Relation between clinical and mammographic diagnosis of breast problems and the cancer/ biopsy rate", *Can. J. Surg*., 1996; **39**(2): 128-132.
8. Polakowski, W. E. Cournoyer, D. A. Rogers, S. K.DeSimio, M. P. Ruck, D. W. Hoffmeister, J. W. and Raines, R. A., "Computer-aided breast cancer detection and dagnosis of masses using difference of Gaussians and derivative-based feature saliency", *IEEE Trans. Med. Imag.,* 1997; **16**(6): 811-819.]
9. S.M.Astley, "Computer–based detection and prompting of mammographic abnormalities", *Br.J.Radiol*, 2004; **77**: S194- S200.
10. L.J.W. Burhenne, "potential contribution of computer aided detection to the sensitivity of screening mammography", *Radiology*, 2000; **215**: 554-562.
11. T.W.Freer and M.J.Ulissey, "Screening mammography with computer aided detection: prospective study of 2860 patients in a community breast cancer", *Radiology*, 2001; **220**: 781-786.
12. M Wirth, M Fraschini, J Lyon, "Contrast enhancement of microcalcifications in mammograms using morphological enhancement and non-flat structuring elements", in 17th IEEE *Symposium on Computer-Based Medical System*, 2004; 134-139.
13. T Stoji, B Reljin, "Enhancement of microcalcifications in digitized mammograms: Multifractal and mathematical morphology approach". *FME Trans.* 2010; **38**: 1-9.
14. B. Sahiner, H.P. Chan, N. Petrick, M.A. Helvie, M.M. Goodsitt, "Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness", *Acad. Radiol*. 2000; **7**: 1077-1084.
15. B. Sahiner, N. Petrick, H.P. Chan, "Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization", IEEE Trans. Med. Imaging 2001; **20**(12): 1275–1284.
16. S.L. Ng, W.F. Bischof, "Automated detection and classification of breast tumors", *Comput. Biomed. Res.* 1992; **25**: 218–237.
17. F.N. Che, M.C. Fairhurst, C.P. Wells, M. Hanson, "Evaluation of a two-stage model for detection of abnormalities in digital mammograms", in: Proceedings of the 1996 IEE Colloquium on Digital Mammography, No. 072, London, UK, 1996.
18. H.D. Cheng, X.J. Shi, R. Min, L.M. Hu, X.P. Cai, H.N. Du," Approaches for automated detection and classification of masses in mammograms", *Pattern Recognition* 2006; **39**: 646-668
19. H. Li, Y. Wang, K.J.R. Liu, S.B. Lo, M.T. Freedman, "Computerized radiographic mass detection C part I: lesion site selection by morphological enhancement and contextual segmentation", *IEEE Trans. Med. Imaging* 2001; **20**(4): 289–301.
20. B. Zheng, Y.H. Chang, X.H. Wang, W.F. Good, "Comparison of artificial neural network and Bayesian belief network in a computerassisted diagnosis scheme for mammography", IEEE International Conference on Neural Networks, 1999; 4181–4185.
21. M.A. Kupinski, M.L. Giger, "Investigation of regularized neural networks for the computerized detection of mass lesions in digital mammograms", in: Proceeding of the 19th International Conference of IEEE/EMBS, 1997.
22. L.H. Li, Y. Zheng, L. Zhang, R.A. Clark, "False-positive reduction in CAD mass detection using a competitive classification strategy", *Med. Phys*. 2001; **28**: 250-258.