Computer Aided Diagnosis System for Classification of Microcalcifications in Digital Mammograms

Mazen E. Osman[1], Manal Abdel Wahed[1], Ahmed S. Mohamed[1], Yasser. M. Kadah[1,2]

[1]*Biomedical Engineering Department, Cairo University, Egypt, (manalaw@k-space.org)*

[2]*Center for Informatics Science, Nile University, Egypt*

Abstract

Breast cancer is the main cause of death for women between the ages of 35 to 55. Mammogram breast X-ray is considered the most reliable method in early detection of breast cancer. Microcalcifications are among the earliest signs of a breast carcinoma. Actually, as radiologists point out, microcalcifications can be the only mammographic sign of non-palpable breast disease which are often overseen in the mammogram. In this paper a method is proposed to develop a Computer-Aided Diagnostic system for classification of microcalcifications in digital mammograms, it splits into three-step process. The first step is Region of Interest extraction of 32 x 32 pixels size. The second step is the features extraction, where we used a set of 234 features from Region of Interest by employing wavelet decomposition, 1st order statistics from wavelet coefficients algorithms; also, we extracted 1st order statistics, median contrast and local binary partition features. The third step is the classification process where differentiation between normal and abnormal is performed using a Minimum Distance Classifier and K-Nearest Neighbor Classifiers employing the leave-one-out training-testing methodology. The results show acceptable sensitivity and specificity for the proposed system.

1. Introduction

Breast cancer is the main cause of death for women between the ages of 35 to 55 year. Early detection and treatment of breast cancer are the most effective methods of reducing mortality. Microcalcifications (MCs) are among the earliest signs of a breast carcinoma. Actually, as radiologists point out [1], microcalcifications can be the only mammography sign of non-palpable breast disease. Due to the subtle nature of these microcalcifications, these are often overseen in the mammogram. Some retrospective studies state that in up to 40% of the cases unambiguous signs of a cancer were missed by the reader, with in some cases fatal consequences for the patient. Thus the reliable detection and classification of microcalcifications plays a very important role in early breast cancer detection.

Computer Aided Diagnosis (CAD), which is reality today, was a first step to simplify the detection of malignant lesions. Up to now, however, (CAD) systems just put markers on suspicious regions. They do not generate a processed image that might show relevant features more clearly. This restriction is due to two reasons. First, most mammograms are still film-based, and are read using a light box. Thus commercial (CAD) systems digitize the film and present markers on a small display or on a separate printout. Second, for the purpose of detection the image is just decomposed to generate features for a classifier. The task of enhancement is more complex, as it also requires an image reconstruction. Only a very sophisticated reconstruction algorithm can provide images that are still suitable for diagnostic reading.

Many different techniques were used for detection and calcification of microcalcifications. J. Jiang *et al.*[2] Proposed Genetic Algorithm(GA) technique which is characterized by transforming input images into a feature domain, where each pixel is represented by its mean and standard deviation inside a surrounding window of size 9×9 pixel. In the feature domain, chromosomes are constructed to populate the initial generation and further features are extracted to enable the proposed GA to search for optimized classification and detection of microcalcification clusters via regions of 128×128 pixels. Extensive experiments show that the proposed GA design is able to achieve high performances in microcalcification classification and detection, which are measured by ROC curves, sensitivity against specificity, areas under ROC curves and benchmarked by existing representative techniques.

Ping Zhang *et al.* [3] proposed and investigated a neural-genetic algorithm for feature selection in conjunction with neural and statistical classifiers to classify microcalcification patterns in digital mammograms. The obtained results show that the proposed approach is able to find an appropriate feature subset and neural classifier achieves better results than two statistical models. Omara *et al.* [4] used both the wavelet coefficients and the statistical measures of different wavelet detail levels as features that describe effectively any normal and abnormal region. Two Techniques were used for the classification stage the minimum distance classifier and the voting K-Nearest Neighbor classifier. Yacoub *et al.* [5] proposed an algorithm that is divided into three steps.

The first step is region of interest (ROI) extraction of 256 x 256 pixels size. The second step is the feature extraction, where a set of 99 features were used and they found that 83 of these feature are capable of differentiating between normal and cancerous breast tissues. The third step is the classification process. They used the techniques of the minimum distance, the k-Nearest Neighbor (k- NN) and Bayes classifiers to classify between normal and cancerous tissues. Alolfe *et al.* [6] used a set of 88 features found that 78 capable to differentiate between malignant and benign as an input to the classification step using minimum distance classifier and K-NN classifier. Wei et *al.* [7] considered support vector machine (SVM), Kernel Fisher Discriminant (KFD), Relevance Vector Machine (RVM), and committee machines (ensemble averaging and AdaBoost), of which most have been developed recently in statistical learning theory. They formulated differentiation of malignant from benign MCs as a supervised learning problem, and applied these learning methods to develop the classification algorithm. As input, these methods used image features automatically extracted from clustered MCs.

Microcalcifications appear at a range of varying sizes. Thus it is natural to approach this problem by multiscale techniques. Among the more recently developed techniques, most share a feature based approach based on multiscale filter bank decompositions. The most successful methods apply filter banks that are variations of the standard discrete wavelet transform decomposition. Yoshida *et al.* [8] applied a Discrete Wavelet Transform (DWT) with dyadic scales. They multiply every wavelet scale by a weight factor. Then they reconstruct an image by applying the inverse transform. The weights are determined by supervised learning, using a set of training cases. This approach results in an overall enhancement of edges and structures. There is no coefficient selection scheme in wavelet domain. Strickland *et al.* [9] used the Discrete Wavelet Transform with biorthogonal spline filters. To overcome the restriction of dyadic scales and to adapt the transform better to microcalcifications they abandon the reconstruction property. They computed four dyadic scales plus two additional interpolating scales (voices). On every wavelet scale a binary threshold-operator was applied. The responses of the individual wavelet scales were then combined by the rule of probability summation. The output was used as a feature for detection of microcalcifications. Despite being a very simple algorithm, the detection results of R. Strickland *et al.* demonstrate the power of a wavelet-based approach.

In this work, a software program was prepared to localize the abnormalities using information associated within the data files. Then we were able to present the standardize mammograms with the region of interest highlighted. The 32x32 pixel region of interest was then determined around the center of the abnormalities. Wavelet decomposition [10] was applied over these regions and the statistical features and wavelet coefficients vectors were then extracted. Also, 1st order statistics, median contrast and local binary partition features were calculated from ROI texture. These features were then presented to both the voting k-nearest and minimum distance classifiers. A scaling between (0-1) was made to judge the normality and abnormality of the imaged tissue. The entire procedure of system development is presented in Figure 1. The following gives a detailed description of each step.

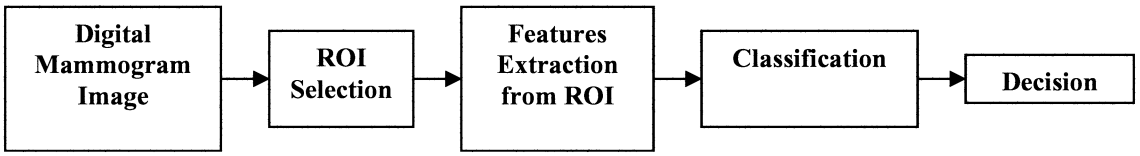

Figure1. A schematic diagram for the CAD system

2. Methodology

The data used in this work was taken from the Mammographic Image Analysis Society (MIAS). The size of all the images is 1024 x 1024 pixels. The existing data in the collection include the coordinates of the center(x, y) and the radius of the microcalcifications [11].

2.1. Extraction of ROI

Using the information supplied by the MIAS for each mammogram, we extracted the ROI of size 32 x 32 pixels with microcalcifications centered in the window. We have used 22 cancerous and 44 normal ROIs.

2.2. Features extraction

2.2.1 Wavelet decomposition coefficients

In this stage, the wavelet decomposition was applied on the region of interest using both the wavelet Daubechues (db4) and the function *wmaxlev* provided by Matlab® [12] to determine the maximum wavelet decomposition scale N, it helps to avoid unreasonable maximum scale values according to number of scales that contain irredundant information. The output of wavelet analysis are the decomposition vector C and corresponding book keeping matrix S, The vector C consists of horizontal, vertical, and diagonal detail coefficients and one approximation. The horizontal, vertical and diagonal detail was extracted from the wavelet decomposition structure [C, S]. These vectors were extracted at each scale from scale one to N+1. The coefficients vectors [H, V and D] were then normalized by dividing each vector by its maximum value. As a result all vectors values become less than or equal one. Then the energy for each vector was computed by squaring every element in the vector.

Since high number of coefficients was produced (about 43800), this number was reduced by summing a predefined number (200 by trial and error) of energy values together in a single number. The 219 produced values are then considered as features for the classification stage.

2.2.2 Statistical features

Wavelet theory provides a powerful framework for multiresolution analysis, and it can be used for texture analysis. The wavelet transform is used to map the regions of interest into a series of coefficients, which constitute a multiscale representation of the ROIS [13]. In this work from each wavelet vector 8 statistical features that include mean, standard deviation, kurtosis, skewness, max, min, median absolute deviation and range were estimated. And 4 features include mean, max, min and range that were estimated from each ROI.

2.2.3 Local binary partition

It is very simple and useful texture measure. For each pixel P in the image, the eight neighbors are examined to see if their intensity is greater than that of P. The results from the eight neighbors are used to construct an eight-digit binary number b1b2b3b4b5b6b7b8. Where bi = 0 if the intensity of the i th neighbor is less than or equal to that of P and 1 otherwise. A histogram of these numbers is used to represent the texture of the image [14] by calculating the mean and standard deviation.

2.2.4 Median contrast

$$C(i,j) = P(i.j) - \underset{l,m \in window}{\text{Median}}\left(y(l,m)\right) \quad , \tag{1}$$

where *P(i, j)* is the pixel value at position *(i, j),* and window is a 3x3 square area centered at position *(i, j), C (i, j)* is the median contrast at position *(i, j)* [15].

2.3. Classification stage

In order to assess the discriminative power of the extracted features and classifiers, two classification schemes were applied against the verified diagnosis for each case, minimum distance classification and K-nearest neighbor classification. An important initial step of classification is to divide the data into two independent subsets, learn and test sets. This step is important to avoid the bias effects in the error estimation phase [16]. In this step we used what is called leave-one-out training testing method [17] because of the small number of the cancerous data base.

2.3.1 The minimum distance classification

This method assumes that the classes are similar in distribution and are linearly separable. Hence, the decision lines are allocated half way between the centers of clusters of different classes. The algorithm works as follow:

1. Group the learn set into two supervised cluster according to their labels (malignant, and normal), representing the two pathologies of interest.
2. Estimate the sample mean for each class by averaging the parameter set of the class.

3. A test sample is classified by assigning it to the class which has the nearest mean vector.
4. Error rate is estimated by the percentage of misclassified samples.

2.3.2 The voting K- Nearest Neighbor (K-NN) classification

K-nearest neighbor (K-NN) classifier distinguishes unknown patterns based on the similarity to known samples. The K-NN algorithms computes the distances from an unknown patterns to every sample and select the K-nearest samples as the base for classification. The unknown pattern is assigned to the class containing the most samples among the K-nearest samples. In this work we used two kinds of distances: Euclidean distance (ED) and Sum of Differences (SOD) distance.

3. Results and discussion

This section presents the results achieved in this work. We measured, quantitatively, the detection performance of the classifiers by computing the sensitivity and specificity on the data. Sensitivity is the conditional probability of detecting cancer while there is really cancer in the image. Specificity is the conditional probability of detecting normal breast while the true state of the breast is normal.

In the terms of the false-negative rate and the false-positive rate:

Sensitivity = 1- false-negative rate.

Specificity = 1- false-positive rate.

False-negative rate: the probability that the classification result indicates a normal breast while the true diagnosis is indeed a breast disease (i.e. positive). This case should be completely avoided since it represents a danger to the patient.

False-positive rate: the probability that the classification result indicates a breast disease while the true diagnosis is indeed a normal breast (i.e. negative). This case can be tolerated, but should be as in frequent as possible. Table 1 shows the results of minimum classifier, the voting K-NN classifier using Euclidean distance and K-NN using Sum of Differences with varying the value of K between 1- 11.

Table1. Classifiers results

Classifier		Learning set		Testing set	
		Sensitivity	Specificity	Sensitivity	Specificity
Minimum distance		90.91%	77.27%	100%	68.18%
K-NN (ED)	K= 1	100%	100%	77.27%	90.91%
	K= 3	90.91%	100%	40.91%	90.91%
	K= 5	90.91%	100%	18.18%	90.91%
	K= 7	90.91%	100%	100%	86.36%
	K= 9	90.91%	100%	100%	86.36%
	K= 11	81.82%	86.36%	100%	81.82%
K-NN (SOD)	K= 1	100%	100%	100%	90.91%
	K= 3	100%	95.45%	100%	86.36%
	K= 5	90.91%	100%	95.45%	81.82%
	K= 7	90.91%	100%	100%	81.82%
	K= 9	90.91%	100%	100%	86.36%
	K= 11	90.91%	86.36%	100%	68.18%

Results show that: For the learning set the K-NN classifier using ED with K= 1 is better than other values of k (sensitivity = 100%, specificity = 100%), while the K-NN classifier using SOD with K= 1 is better than other values of k (sensitivity = 100%, specificity = 100%). For the testing set the K-NN classifier using ED with K= 7, 9 and 11 showed better sensitivity than other values of k (sensitivity = 100%), while those with K= 1, 3 and 5 showed better specificity than other values of k (specificity = 90.91%). For the K-NN classifier using SOD with K= 1,3,7,9 and 11 showed better sensitivity than the other values of k (sensitivity = 100%), while the one with K= 1 showed better specificity than other values of k (specificity = 90.91%). Finally K-NN classifier using SOD and K= 1 shows better results than other classifiers.

In general, it is to be noted that K-NN classifiers are much better than minimum distance classifier, also there is no considerable differences between K-NN classifiers using Euclidean distance (ED) and K-NN classifiers using sum of differences (SOD) with different values of k. Finally K-NN classifier using SOD and K= 1 showed better results than other classifiers.

Comparing the results obtained from the K-NN classifier in this study with the results obtained from previous study [6], with K-NN classifier using SOD and k=1 for the testing set, the sensitivity and specificity results are much better than previous study (sensitivity = 100%, specificity = 100%).

4. Conclusion

Despite recent advances in this field, computerized microcalcifications detection is still far from being perfect. This returns to many reasons. Those come from the great variability in the database mammograms, the use of poor resolution microcalcifications mammograms and small number of the available database. Future improvement to the work could be by using better data base with higher number of images and better resolution. Extraction of more features and application of more classifiers can also improve the system.

References

[1] Bin Zheng, Lara A. Hardesty, William R. Poller, Jules H. Sumkin and Sara Golla, "Mammography with computer – aided detection : reproducibility assessment – initial experience", *Radiology*, vol. 228, no. 1, pp. 58–62, 2003.

[2] J. Jiang, B. Yao, A.M. Wason. "A genetic algorithm design for microcalcification detection and classification in digital mammograms", *computerized medical imaging and graphics*, 31, pp. 49-61, 2007.

[3] Ping Zhang, Brijesh Verma and Kuldeep Kumar , "Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection", *Pattern Recognition Letters*, 26, pp. 909–919, 2005.

[4] Akram I. Omara, Ahmed S. Mohamed, Abo-Bakr M. Youssef, and Yasser M. Kadah, "Computer Aided Diagnosis in Digital Mammography", Proceeding of the 3rd Cairo International Biomedical Engineering Conference (CIBEC'06), Cairo, Dec. 2005.

[5] Michael A. Yacoub, A. S. Mohamed 1, Yasser M. Kadah, " A CAD System for the Detection of Malignant Tumors in Digitized Mammogram Films", Proceeding of the 3rd Cairo International Biomedical Engineering Conference (CIBEC'06), Cairo, Dec. 2005.

[6] Mohamed A. Alolfe, Abou-Bakar M. Youssef, and Yasser M. Kadah, "Development Computer-Aided Diagnosis (CAD) System for Breast Masses Detection in Digital Mammograms", Proceeding of the 3rd Cairo International Biomedical Engineering Conference (CIBEC'06), Cairo, Dec. 2005.

[7] Liyang Wei, Yongyi Yang, Robert M. Nishikawa, and Yulei Jiang, "A study on several machine learning methods for classification of malignant and benign clustered microcalcifications", *IEEE Trans on Medical Imaging*, vol. 24, no. 3, March 2005.

[8] H. Yoshida, K. Doi and R. M. Nishikawa, "Automated detection of clustered microcalcifications in digital mammograms using wavelet transform techniques", *Medical Imaging, SPIE* , pp. 868–886, 1994.

[9] R. N. Strickland and H. L. Hahn, "Wavelet transforms for detecting microcalcifications in mammograms", *IEEE Trans. on Medical Imaging*, 15(2), pp. 218–229, 1996.

[10] Rafayah Mousa, Qutaishat Munib, Abdallah Moussa, "Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural", *Expert systems with applications*, 28, pp.713-723, 2005.

[11] Mammographic Image Analysis Society Website, http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html (cited on 1/3/2008)

[12] MATLAB user guide wavelet tool box 3, The Math Works Inc. 2004. [On line] http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf

[13] H.D. Cheng, Xiaopeng Cai, Xiaowei Chen, Liming Hu, Xueling Lou, "Computer-aided detection and classification of Microcalcifications in mammograms: a survey", *Pattern Recognition*. 36, pp. 2967 – 2991. 2003.

[14] Linda Shapiro and George Stockman, *Computer Vision*, March 2000. ISBN-10:0130307963.

[15] Song yang Yu and Ling Guan "A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films", *IEEE transactions on medical imaging*, Vol. 19, no. 2, February 2000.

[16] Yasser M. Kadah, Aly A. Farag, Jacek M. Zurada, Ahmed M.Badawi, Abou-Bakr M. Mousses, "Classification algorithms for quantitative characterization of diffuse liver disease from ultrasound images", *IEEE Trans on Medical Imaging*, vol. 15, no. 4, Aug 1996.

[17] Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, pp. 32-33, 2006, ISBN-10: 0-387-31073-8.