

An Efficient System for Cancer Detection using Digital Mammograms

Ashraf Anwar

Faculty of Computers and Information Technology, Taif University, P.O. Box 888, 21974 Taif, Saudi Arabia

Abstract: Breast cancer can be considered one of the most dangerous types of cancer among women. Early detection of breast cancer leads to significant improvements in treatment. Digital mammograms are one of the most effective means for detecting breast cancer in early stages. In this paper, an efficient system based on performing professional pre-processing phase and on applying Discrete Cosine Transform (DCT) for features extraction, and support vector machine has been used for classification into benign and malignant. We have used Mias data set for experimentation purpose. We tune the coefficients of DCT to get the best sensitivity, specificity, positive predictivity, and accuracy results. We reach 100% performance rate in some cases.

Keywords: Breast Cancer, Mammogram, Discrete Cosine Transform, Svm, Mias Data Set.

I. Introduction

Breast cancer can be considered a leading cause of cancer deaths among women in many parts of the world. Since the causes of the disease till now, are completely unknown, it's highly difficult to prevent it or take precautions for it. But an early detection of the disease may help to increase the chances/rate of survival. Breast cancer is an abnormal, un-controlled growth of breast cells where the cells get divided un-conditionally giving rise to tumor (mass). The tumor developed this way can be either Benign or Malignant. The Benign tumors are not cancerous. They just grow larger but do not spread to other parts of the body. Hence they do not cause much harm to the person. But malignant tumors are considered as cancerous. They grow larger and spread to other parts of body. Hence they are treated as very dangerous to the person, in whom they are found [1]. The Digital Mammography has proved itself as the best technique for early detection of breast cancer, but Mammography itself cannot prove that a suspicious area is malignant or benign. With the growth of computer technology, radiologists have a chance to improve their image interpretation using computer capabilities that can improve the image quality of mammogram [2]. The image processing techniques like pre-processing, segmentation, feature-extraction and classifications are used in Digital Mammography that provides clear mammogram images, which helps the radiologists to give the most accurate second opinion. A variety of computer-aided diagnosis (CAD) systems [3-7] have been proposed. CAD plays an important role in diagnosis of breast cancer. Radiologists are trained to differentiate between benign and malignant abnormalities but for computer it needs proper training to select successfully.

In this task, we proposed an efficient system based on a professional preprocessing technique, accurate tuning of the discrete cosine transform coefficients, and smoothing parameter of a SVM classifier, to realize a perfect performance rate. The rest of this paper is split into four sections; section 2 contains material and methods, experimental results are given in section 3, section 4 concluded this paper.

II. Material and method

The database used in this task is taken from Mammographic Image Analysis Society (MIAS) [11]. It contains 322 mammogram; each mammogram is 1024 x 1024 pixels and 200 micron resolution. Out of which 62 are benign, 51 as malignant and 209 are normal mammogram. This task is implemented by using Image Processing Toolbox in Matlab 7.5. The block diagram of the proposed method is shown in Figure 1. It consists of four main steps: pre-processing, features extraction, feature selection, classification, and performance evaluation.

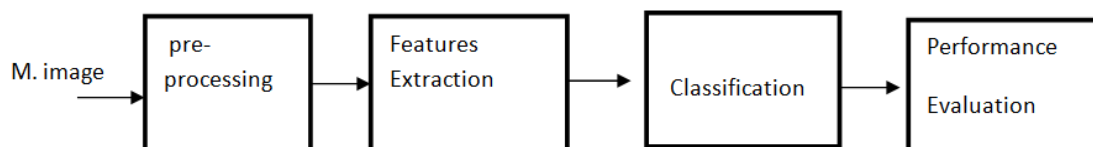


Figure 1: The block diagram of the proposed method

2.1 Pre-Processing

pre-processing stage is very important to minimize the computational cost and also finds the ROI (region of interest). Those parts which are not part of breast can misguide the algorithm for classification and this will affect the performance of the proposed method. The pre-processing stage consists of:

- Noise removal : we can remove noise by using averaging filter and a median filter
- Pectoral muscle removal : we used seeded region growing algorithm for separation the pectoral muscle from the breast [8]
- Extract small elements and detail from images : top-hat transform is an operation that extracts small elements and details from given images [9]
- Contrast enhancement : the mammogram are contrast enhanced by using; Contrast Limiting Adaptive Histogram Equalization (CLAHE) [10]

A Sample of mammogram pre-processing for a benign case(mdb1) and a malignant case (mbd72), is depicted in Figure2

2.3 Feature Extraction

Features have a significant effect for classification stage success in CAD (Computer Aided Diagnostic). For this purpose we have used Discrete Cosine Transform (DCT) in our proposed system.

2.3.1 Discrete Cosine Transform (DCT) Features.

Discrete cosine transform (DCT) is used for converting the signal into its frequency components. The DCT of an NxN image f(x,y) is defined by:

$$F(u, v) = c(u)c(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{\pi u (2x+1)}{2N}\right) \cos\left(\frac{\pi v (2y+1)}{2N}\right) \tag{1}$$

Where $c(u) = c(v) = \frac{1}{\sqrt{N}}$, for u, v =0

$C(u) = c(v) = \frac{2}{\sqrt{N}}$, for u, v ≠0

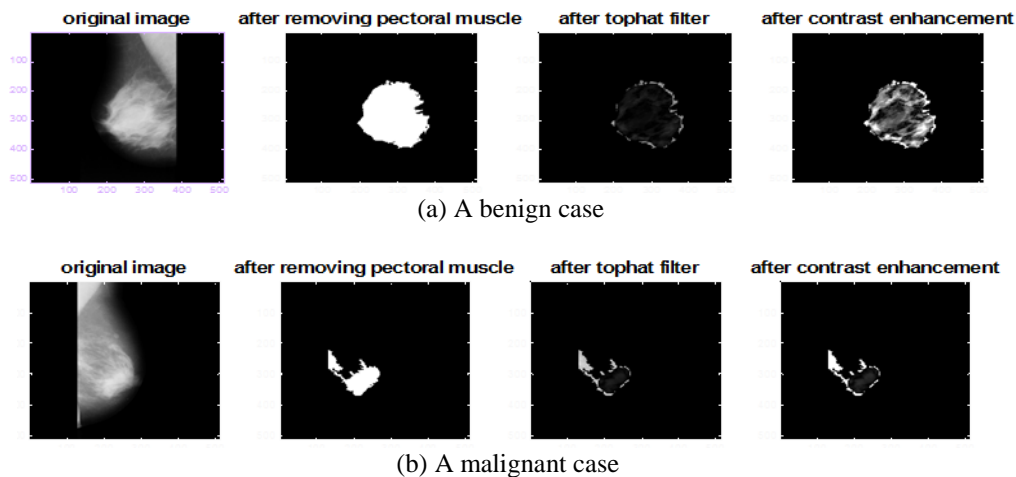


Figure 2: A Sample mammogram pre-processing for a benign case (mdb1) and a malignant case (mbd72), The DCT, when applied to an MxN image, compresses all the information of the image and concentrates it in a few coefficients located in the upper-left corner of the resulting real-valued MxN DCT matrix [7]. Note that the transform image has zero or low level intensities except at the top left corner where the intensities are very high. These low frequency and high intensity coefficients carry most of the information about the original image.

2.3.2 Square Window Method to Extract Features.

This method extracts (kxk =k^2) lowest frequency coefficients in the upper corner of the DCT matrix as shown in Figure 3 To illustrate the scanning scheme of the square window, let c_{mn} designate the coefficient in the DCT matrix located in m^{th} row and n^{th} column. Then a 1x1 window, generates the 1-element vector $W_{1x1}=[c_{11}]$. Similary, 2x2 window generates the vector $W_{2x2}=[c_{11} c_{12} c_{21} c_{22}]$, so on.

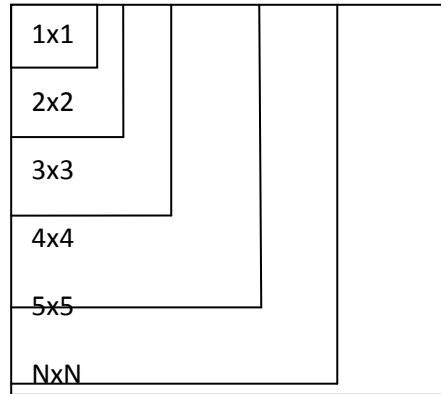


Figure 3: Square window method

2.4 Classification

The SVM is a discriminative classifier formally defined by a separating hyperplane (the plane with maximum margins) between the two classes of the training samples within the feature space by focusing on the training cases placed at the edge of the class descriptors so not only an optimal hyperplane is fitted, but also training samples are effectively used. In that way high classification accuracy is achieved with small training sets [12].

In soft margin classification, the SVM algorithm can be summarized as the following optimization problem: given a training set $(x_i, y_i), i=1,2,\dots,n$

$$\min_{\frac{1}{2}W^TW + C \sum_{i=1}^n \xi_i} \text{ for all } \{(x_i, y_i)\}$$

Subjected to: $y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$ for all i (2)

Where : $\Phi(x)$ is a nonlinear function that maps x into a higher dimensional space.

$W, b,$ and ξ are the weight vector, bias, and slack variable respectively. C is a constant determined a priori. Parameter C can be viewed as a way to control over fitting. Most “important” training points are support vectors; they define the hyperplane. Quadratic optimization algorithms can identify which training points x_i are support vectors with non-zero Lagrangian multipliers α_i . By constructing a Lagrangian and transforming it into a dual maximization of the function $Q(\alpha)$, defined as follows:

$$\max Q(\alpha) = \sum \alpha_i - \frac{1}{2} \sum \sum \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

Subject to: $\sum_{i=1}^n \alpha_i y_i = 0; 0 \leq \alpha_i \leq C, \text{ for } i=1, 2, \dots, n$ (3)

Where $K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$ is the kernel function and α_i is vector of non negative Lagrange multipliers. The kernel function plays the role of the dot product in the feature space.

Suppose that the optimum values of the Lagrange multipliers are denoted α_0 , it is then to determine the corresponding optimum value of the linear weight vector w_0 and the optimal hyperplane as in (4) and (5), respectively:

$$w_0 = \sum_{i=1}^n \alpha_{0,i} y_i \phi(x_i) \tag{4}$$

$$\sum_{i=1}^n \alpha_{0,i} y_i K(x_i, x_j) + b \tag{5}$$

The solution is

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_{0,i} y_i K(x_i, x_j) + b) \tag{6}$$

Kernel functions may be one of the following types:

Linear: $K(x_i, x_j) = x_i^T x_j$ (7)

Polynomial of power p : $K(x_i, x_j) = (1 + x_i^T x_j)^p$ (8)

Gaussian (radial-basis function network): $K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$ (9)

Sigmoid: $K(x_i, x_j) = \tanh(\beta_0 x_i^T x_j + \beta_1)$

- (10)

In this task, we used radial basis function (RBF) as kernel function where:

σ (kernel width) : is the distance between closest points with different classifications

C, σ were experimentally defined to achieve the best classification result

III. Experiments and results

To evaluate the performance of the proposed method, the training set and testing set for Benign and Malignant extracted from MIAS data base is shown in Table 1

Table 1

Four criterias are used as follows:

| Training set | | Testing set | |
|--------------|-----------|-------------|-----------|
| Benign | malignant | Benign | malignant |
| 46 | 38 | 16 | 13 |
| 84 | | 29 | |

$$\text{Sensitivity} = TP / (TP + FN) \tag{11}$$

$$\text{Specificity} = TN / (FP + TN) \tag{12}$$

$$\text{Positive Predictivity} = TP / (TP + FP) \tag{13}$$

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{14}$$

Where:

TP: True Positive; is the number of malignant that are correctly classified.

TN: True Negative; is the number of benign that are correctly classified

FN: False Negative; when a malignant is classified as benign incorrectly

FP: False Positive; when a benign is classified as malignant incorrectly

We investigate the optimum number of DCT coefficients to be used as features to realize high performance rate (sensitivity, Specificity, Positive Predictivity, and Accuracy). The results of changing window size against performance rate are shown in Figure 4. The high performance rate is achieved when window of size 3x3 is processed. Also we change the coefficients inside W3x3 against the performance rate and we get perfect performance 100% at coefficient c_{32} and c_{33} inside W3, 3 as shown in Figure 5

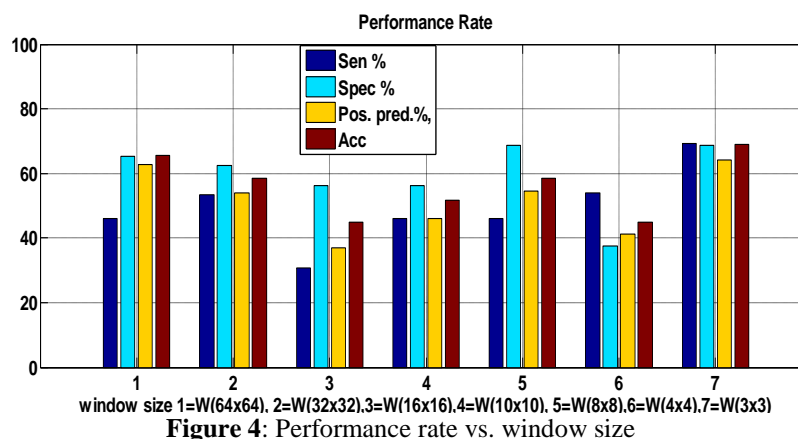


Figure 4: Performance rate vs. window size

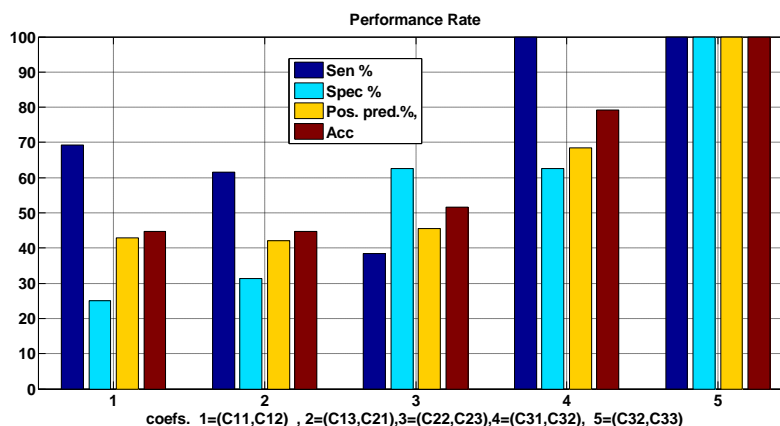


Figure 5: Performance rate vs. window(3x3) coefficients

The results of training phase and testing phase for Benign (B) and malignant (M) resulting from SVM classifier are depicted in Figure 6.

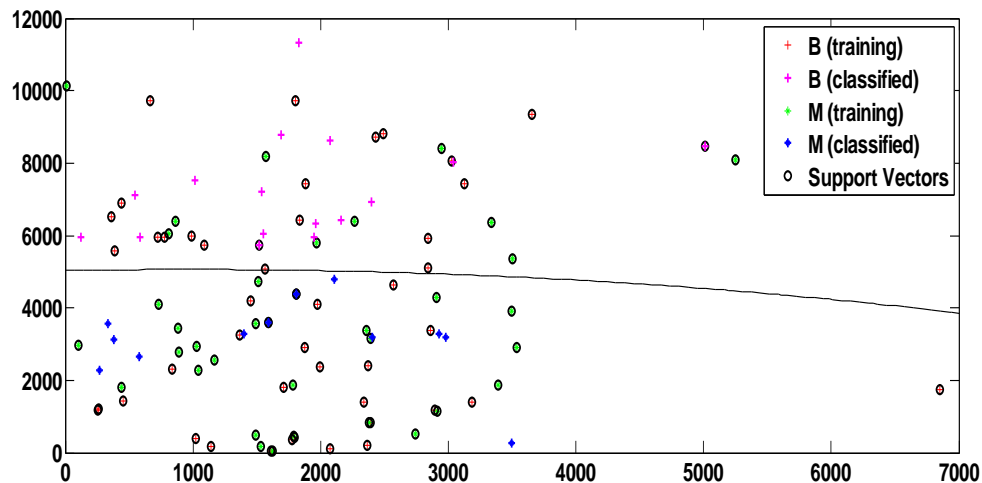


Figure 6: SVM training and testing results

IV. Conclusion

An efficient system to detect benign and malignant breast cancer using mammogram images is developed in this research. A professional preprocessing stage is applied to remove noise and pectoral muscle, to extract small elements and details from given images, and to enhance contrast of the images, then by tuning coefficients of the window in discrete cosine transform to extract the distinguished features, this lead to perfect classification done by support vector machine.

References

- [1]. Priyanka, Diganbar: Digital mammography: A review on detection of breast cancer, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 1, Jan 2016.
- [2]. Talha, et al. : Classification of breast mammograms into benign and malignant, International Journal of Multimedia and Ubiquitous Engineering, Vol. 7, No. 2, 2012
- [3]. Jaffar, et al.: DCT Features based malignancy and abnormality type detection method for mammograms, International Journal of Innovative Computing, Information and control, Vol. 7, No. 9, 2011.
- [4]. Rashed et al. : Mutiresolution mammogram analysis in multilevel decomposition, Pattern Recognition Letters, Vol. 28,2007.
- [5]. Talha, : Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features, Biomedical Research Vol. 27, No.2, 2016.
- [6]. Eltoukhy et al.,: Breast cancer diagnosis in digital mammogram using multiscale curvelet transform, Computerized Medical Imaging and Graphics Vol. 34, 2010
- [7]. Sarhan.,: Cancer classification based on DNA microarray data using cosine transform and vector quantization, IJCA, Vol. 17, No. 4, 2010.
- [8]. Jaffar, et al.,: Multi domain features based classification of mammogram images using SVM and MLP, Fourth International Conference on Innovative Computing, Information and Control, 2009.
- [9]. Anaswara et al.,: Medical image enhancement techniques by bottom hat and median filtering, International Journal of Electronics Communication and Computer Engineering, Vol. 5, No. 4, 2014
- [10]. Prathibha, et al.,: mammograms analysis using SVM classifier in combined transform domain, ICTACT Journal of Image Processing, Vol. 1, No. 3, 2011.
- [11]. <http://peipa.essex.ac.uk/info/mias.html>.
- [12]. Ashraf Anwar and Hedi, "An Efficient Paroxysmal Atrial Fibrillation Prediction Method Using CWT and SVM", International Journal of Scientific & Engineering Research, Volume 4, Issue 12, 2013.

Author:

Dr. Ashraf Anwar received the PH.D degree in Electronics and Electrical Communication from Faculty of Engineering, Cairo University, Egypt in 2005. Currently, he is an Associate professor in college of computers and Information Technology, Taif University, Saudi Arabia. His research interests are mainly in the area of, computer vision, digital signal processing, control, and machine learning.

