Reduced Transformed Features Based Breast Mammograms Classification

Muhammad Talha[1, 2], Ghazali Solung[2]

[1]College of Science Research Center, King Saud University, Riyadh, Saudi Arabia
[2]Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia
Corresponding e-mail: mnaseem@ksu.edu.sa

Abstract: Mammography is a particular type of imaging that uses a low-dose x-ray system to examine breasts. A mammography exam, called a mammogram, is used to aid in the early detection and diagnosis of breast diseases in women. In this paper, we have proposed a method that consists of combination of different methods. First we have performed enhancement on breast mammogram to enhance the image quality. We have used Contrast Limited Adaptive Histogram Equalization (CLAHE) for enhancement. Discrete wavelet transform based features has been extracted which have been used for classification. Principle component analysis has used for features reduction and selection. Different classifiers have been used for classification into benign and malignant. It has been noted that results are very much satisfactory. We have used MIAS data set for experimentation purpose.
[Muhammad Talha, Ghazali Solung. **Reduced Transformed Features Based Breast Mammograms Classification**. *Life Sci J* 2013;10(7s):907-913] (ISSN:1097-8135). http://www.lifesciencesite.com. 145

Keywords: Breast cancer, mammogram, enhancement, Bayesian Classifier, MIAS

1. Introduction

Computer aided Diagnosis System (CAD) for breast cancer detection simulates the process of radiologist. The output of this system indicates the decision of radiologist about the case. There are many factors contributing towards growth and development of computer aided diagnostics which is a fact manifested by increased performance level and new diseases brought under this diagnostic pattern. Breast cancer is considered to be one of the leading causes of deaths among females on a global level. In Netherlands for example, approximately 10000 women are diagnosed with this disease per annum and approximately 3500 of these women die from this type of cancer. American National Cancer Institute reported that the population of the estimated new breast cancer cases for the 2006 in USA is round about 214640, while the estimation of deaths more than 41,000 [1]. Cancer statistics claim that breast cancer got the third position of appearance in diagnosed new cases following genital organs and digestive systems cancer comparing to other forms of cancer. Over the past decades it has become alarming that breast cancer incidence rates are increasing steadily. Changes in risk factors seem to contribute to the rising incidence. However, the mortality rates for breast cancer have remained relatively constant due to more effective treatment and earlier diagnosis [2]. According to American Cancer Society 2007, United States has highest figure in the world about crude and age-standardized breast cancer incidents. About 178,480 women suffer from invasive breast cancer and 62,030 from in situ breast cancer. 85% of total in situ breast were ductal carcinoma in situ (DCIS). 40,460 women in American died because of this disease in 2007.An increment is seen in breast cancer death rate between 1975 and 1990, by 0.4% annually. But due to some good treatment and mammographic diagnosis from 1990 to 2002, this is going down by an average of 2.3% per year. Infection rate is different in African black and American white woman. American white women have 20% more chances to have breast cancer then African black women. Although in early 1980's it was higher. Computerized detection of lesions is also being done for mammograms. CAD is being developed for the detection and diagnosis of breast cancer and for the assessment of breast cancer risk [3]. Computerized image analysis in screening mammography has already yielded many fruitful results. FDA (Food and Drug Administration) [4] shows that Computer Aided System are very helpful in screening method of mammogram images.

There are a number of well-known and potential risk factors for breast cancer. These can be divided into seven broad categories: age, hormonal factors, family history of breast cancer, proliferate breast disease, irradiation of the breast region at an early age, lifestyle factors and personal history of malignancy. In reality, estimates indicate that between 10 to 30% of breast cancers are missed by radiologists during routine screening. The penalty of errors in detection or classification is very high. Mammography itself cannot prove that a suspicious area is malignant or benign. To decide that, the tissue has to be removed for examination using breast biopsy techniques. A false positive detection may cause an unnecessary biopsy. Statistics show that only 20% to 30% of breast biopsy cases are proved cancerous. In a false negative detection, an actual tumor remains undetected that could lead to higher costs or even to the cost of a human life. 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 mammograms [4]. In order to develop the accuracy of interpretation, a variety of computer-aided diagnosis (CAD) systems like [3] have been proposed. CAD plays an important role in diagnosis of breast cancer and defining the extent of breast tumors. In previous twenty years, much effort has been made by computer scientists to support the radiologists in detection and diagnosis of cancerous masses by developing computer-aided tools for mammography interpretation. Image processing and intelligent systems are two important mainstreams of computer technologies that have been continuously explored in the development of computer-aided mammography systems.

From computer vision point of view, automatic segmentation and classification of mammogram images is not easy to address. Three types of anatomical variations in the tissues of mammograms image are present from person to person; i.e. fatty, fatty-glandular and dense-glandular. Inherent limitations in the imaging process due to low dose X-Rays which often result in noisy images. Radiologists are trained to differentiate between benign and malignant abnormalities but for computer it needs proper training to adopt accordingly. We have proposed fully automatic and robust technique. Strong preprocessing technique and automatic abnormality type detection method is used. No prior knowledge of the mammogram is required about its feature, type, and contents. This is a supervised method for diagnosing breast cancer. Proposed system achieved quite good accuracy for the classification of mammograms as malignant and benign.

Major Contributions:

- Proposed system is fully automatic and robust technique.
- Strong preprocessing technique and automatic abnormality type detection method is used.
- No prior knowledge of the mammogram is required about its feature, type, and contents.
- This is a supervised method for diagnosing breast cancer. Proposed system achieved quite good accuracy for the classification of mammograms as malignant and benign.
- Proposed system is very accurate system for diagnosing breast cancer.

Paper is organized as follows: Introduction is given in Section 1. Section 2 discusses the proposed method. Section 3 includes experimental results and conclusion is presented in section 4.

2. Proposed System

The proposed system is divided into four major parts as shown in Fig. 1:

- Enhancement by using CLAHE
- Features Extraction & Reduction
- Classification

Figure 1: Proposed Method

The detail of these four steps is described below one by one.

2.1 Preprocessing for Enhancement

In this step, Contrast Limited Adaptive Histogram Equalization (CLAHE) technique has been applied [11]. In CLAHE, the pixel's intensity is transformed to a value within the display range proportional to the pixel intensity's rank in the local intensity histogram. The enhancement is condensed in flat areas of the image, which prevent over enhancement of noise. It also reduces the edge shadowing effect. The CLAHE operates on small regions in the image called tiles rather than the entire image. Each tiles contrast is enhanced, so that the histogram of the output region approximately matches the uniform distribution or Rayleigh distribution or exponential distribution. Distribution is the desired histogram shape for the image tiles. The neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries.

First of all, input image is divided into equal size of number of non-overlapping regions. Then the histogram of each region has been calculated. Clip limit has to be set for clipping histograms. In our case, we have set t=0.002. Each histogram has been redistributed in such a way that its height does not exceed the clip limit. All histograms were modified by the transformation function of normal histogram. Then using bilinear interpolation, neighboring tiles has been combined. At the end, image gray scale values have been altered according to the modified histograms. Flow chart of CLAHE has been shown in Figure 2. Results have been shown in Figure 3.

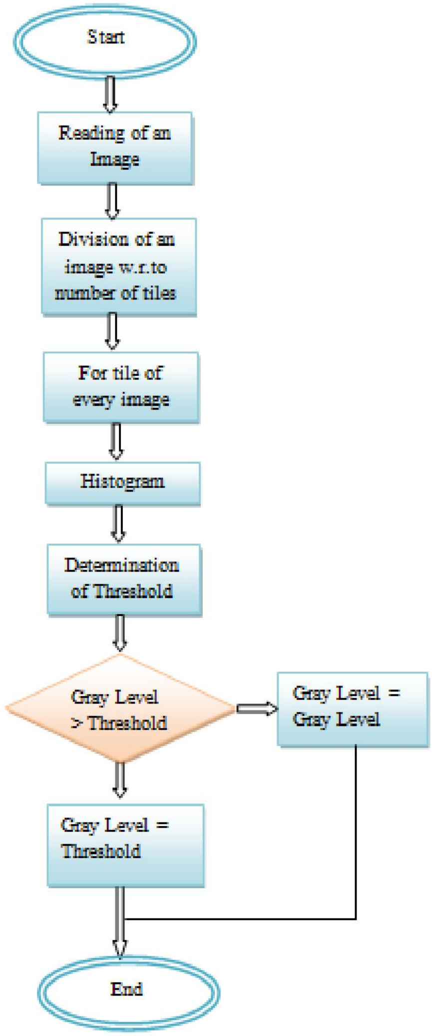


Figure 2: Flow Chart of CLAHE

2.2 Features Extraction

Features play a significant role in CAD (Computer Aided Diagnostic) environment. The transformation of an image into its set of features is known as feature extraction. Useful features of the image are extracted from the image for classification purpose. It is a challenging task to extract good feature set for classification. We have used Local windows based DCT feature for our proposed system.

2.2.1 Discrete Wavelet Transform (DWT) Features

DWT is a linear transformation in which image information is divided into detailed and approximation components. Detail components contain information of vertical horizontal and diagonal subbands of the image. These components can be obtained by applying a high pass and low pass filter on an image respectively. These components are defined by the following equations:

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} l[n-2p]a_j[n] \quad (1)$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n-2p]a_j[n] \quad (2)$$

These DWT features can be used for features reduction.

2.2.2 Principle Component Analysis for Features Reduction

Principal Components Analysis (PCA) is a multivariate procedure which rotates the data such that maximum variability's are projected onto the axes. Essentially, a set of correlated variables are transformed into a set of uncorrelated variables which are ordered by reducing variability. The uncorrelated variables are linear combinations of the original variables, and the last of these variables can be removed with minimum loss of real data [13]. The main use of PCA is to reduce the dimensionality of a data set while retaining as much information as is possible. It computes a compact and optimal description of the data set. The first principal component is the combination of variables that explains the greatest amount of variation. The second principal component defines the next largest amount of variation and is independent to the first principal component. There can be as many possible principal components as there are variables. It can be viewed as a rotation of the existing axes to new positions in the space defined by the original variables. In this new rotation, there will be no correlation between the new variables defined by the rotation. The first new variable contains the maximum amount of variation; the second new variable contains the maximum amount of variation unexplained by the first and orthogonal to the first, etc.

In other words, we wish to reduce a set of p variables to a set of m underlying super ordinate dimensions. These underlying factors are inferred from the correlations among the p variables. Each factor is estimated as a weighted sum of the p variables. The i^{th} factor is thus

$$F_i = W_{i1}X_1 + W_{i2}X_2 + \ldots + W_{ip}X_p \quad (3)$$

One may also express each of the p variables as a linear combination of the m factors,

$$X_j = A_{1j}F_1 + A_{2j}F_2 + \ldots + A_{mj}F_m + U_j \quad (4)$$

where U_j is the variance that is unique to variable j, variance that cannot be explained by any of the common factors.

PCA has been applied on these DWT coefficients to reduce data. We have selected top some features and applied different classifiers.

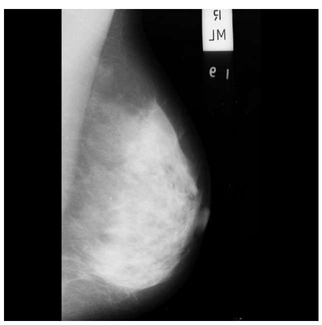
2.3 Classification

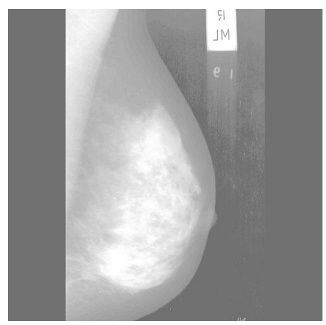
We have used Support Vector Machine (SVM), *k*-Nearest Neighbor Algorithm (KNN), Linear Discriminant Analysis (LDA), Nearest Mean Classifier (NMC), Nearest Mean Scaled Classifier (NMSC), Back Propagation Neural network (BPNN) and Bayesian (BSN) classifiers. In order to remove biasing cross validation has been used so that every sample goes into the process of training and testing at least once. We have executed many times these classifiers but averaged results of the classifiers are calculated and shown here. In the case of KNN, we have used different value of K ranging from 2 to 6. However, the best results were obtained at K=3. So we have used K=3 for presentation of results. For Back Propagation Neural Network (BPNN), we have tested the network on various numbers of hidden layers. For optimal performance, after testing, hidden layers were set at 3.

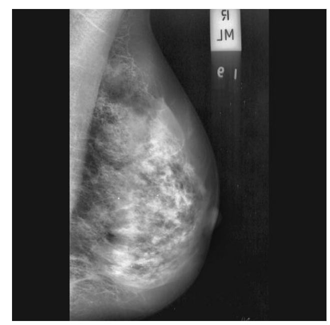
3 Results and Discussion

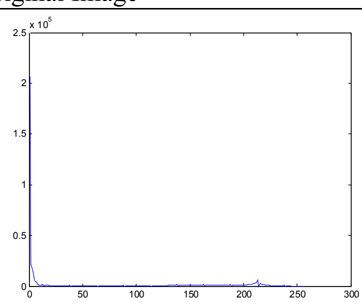
We have used publically available databases MIAS [17]. The dataset is taken from the Mammographic Institute Society Analysis (MIAS). Each mammogram is of size 1024 x 1024 pixels, and resolution of 200 micron. There are 322 mammograms of right and left breast of 161 patients in this dataset. 69 mammograms were diagnosed as being benign, 54 malignant and 207 normal.

Enhancement has been done by CLAHE. Results have been show in figure 1. Results show that CLAHE performs well as compared to histogram equalization. Visually results show that CLAHE is good. Histograms show that CLAHE is closely related to the original image.

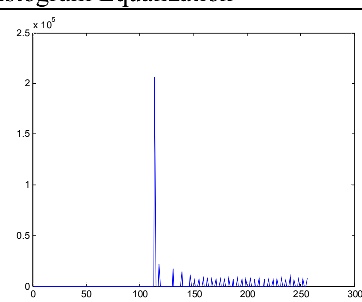

Original Image


Histogram Equalization

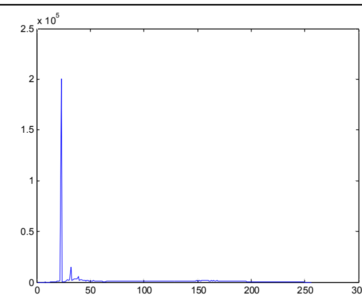

CLAHE



Original Image Histogram



Histogram Equalization Histogram



CLAHE Histogram

Figure 3: Enhancement of Breast Mammogram Images

Performance of classifiers is calculated and analyzed by the following performance measures.

1. **Accuracy**: Number of classified mass / Number of total mass

$(TP + TN)/(TP + TN + FP + FN)$ (1)

2. **Sensitivity**: Number of correct classified malignant mass / Number of total malign mass

$(TP)/(TP + FN)$ (2)

3. **Specificity**: Number of correct classified benign mass / Number of total benign mass

$(TN)/(TN + FP)$ (3)

Where TP is True positive, FP is false positive FN is false negative and TN is true negative.

Classification Results on DWT features

In this case, we have applied DWT on each image. After applying DWT, each image has been divided into four parts approximation and details components. We have selected approximation component again and applied DWT. In this way, we have applied DWT till 2nd level of each approximation component. At 2nd level, we have applied PCA and select some top features. We have applied different classifiers and check the performance. Results have been shown below. It shows that DCT features are good for classification as comparative to DWT.

Experiment I: 50-50/fixed (DWT + PCA)

In this experiment, 50% of the images of each class were used to form the training set, and the other 50% of the images were used to form the test set. Table 1 shows results by using this experiment.

Classifier/Features	Five	Seven	Nine	Eleven	Thirteen
NMSC	0.7233	0.7323	0.8553	0.6917	0.6907
KNN	0.7263	0.7313	**0.9021**	0.7001	0.7067
NMC	0.7715	0.7725	**0.9041**	0.7273	0.7267
FISHERC	0.7313	0.7705	**0.9421**	0.6063	0.5767
BPXNC	0.8627	0.9043	**0.9542**	0.8712	0.8413
Bayesian	0.8787	0.9151	**0.9603**	0.9387	0.9131
SVM	0.9017	0.9253	**0.9721**	0.9211	0.9013

Table 1 Classification Accuracy on Different DWT+PCA Extracted Features

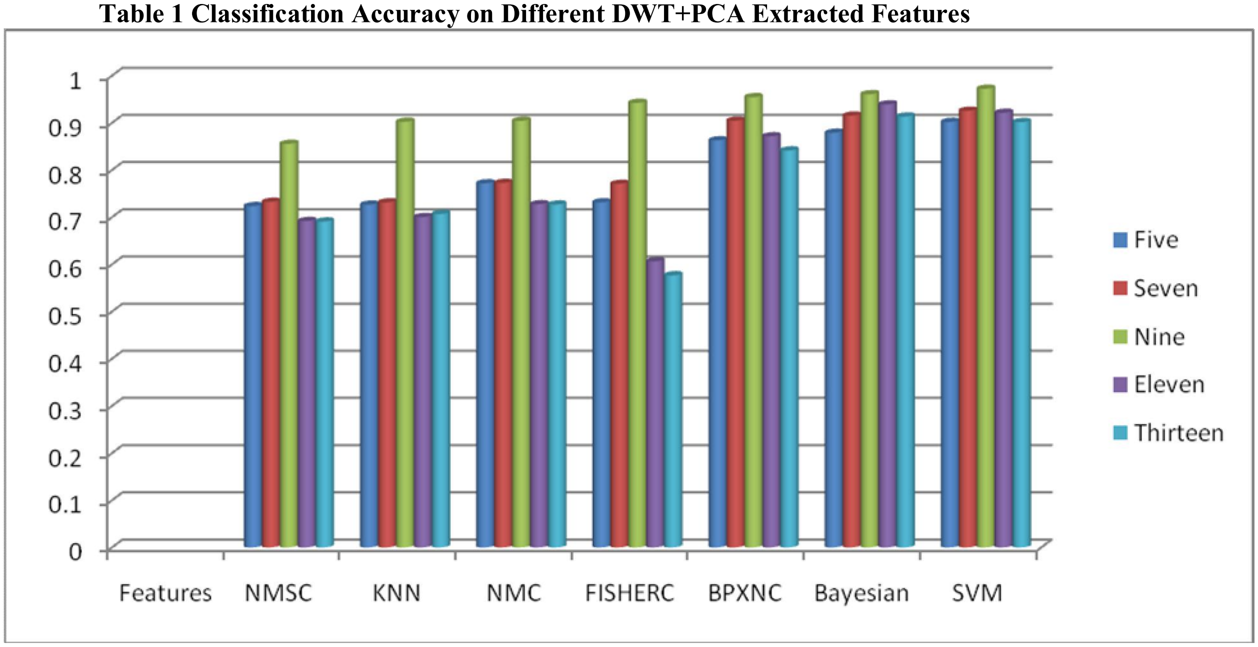


Figure 4: Comparative Classification Accuracy on Different DWT+PCA Extracted Features

Table 1 and Figure 4 present the classification accuracy using DWT+PCA with various numbers of features extracted. The range of features selected by PCA is from 5 to 20. As can be seen in table 1, the best accuracy is achieved when we use nine features. The best accuracy at nine features is demonstrated by all classifiers used in our technique. However, among different classifiers, the best performance is shown by SVM and Bayesian classifier. As the number of features exceeds eight, the accuracy decreases and then become stable at a relatively lower degree. The second best performance was achieved by Bayesian classifier with lesser number of features. However its performance increased with increase in number of features, while others showed significantly less accuracy.

ROC Curve for Classification Performance on Fused features

From previous experiments, we have found out that SVM and Bayesian are good classifiers for classification. Therefore, here we have only selected SVM for ROC curve to show performance of classification.

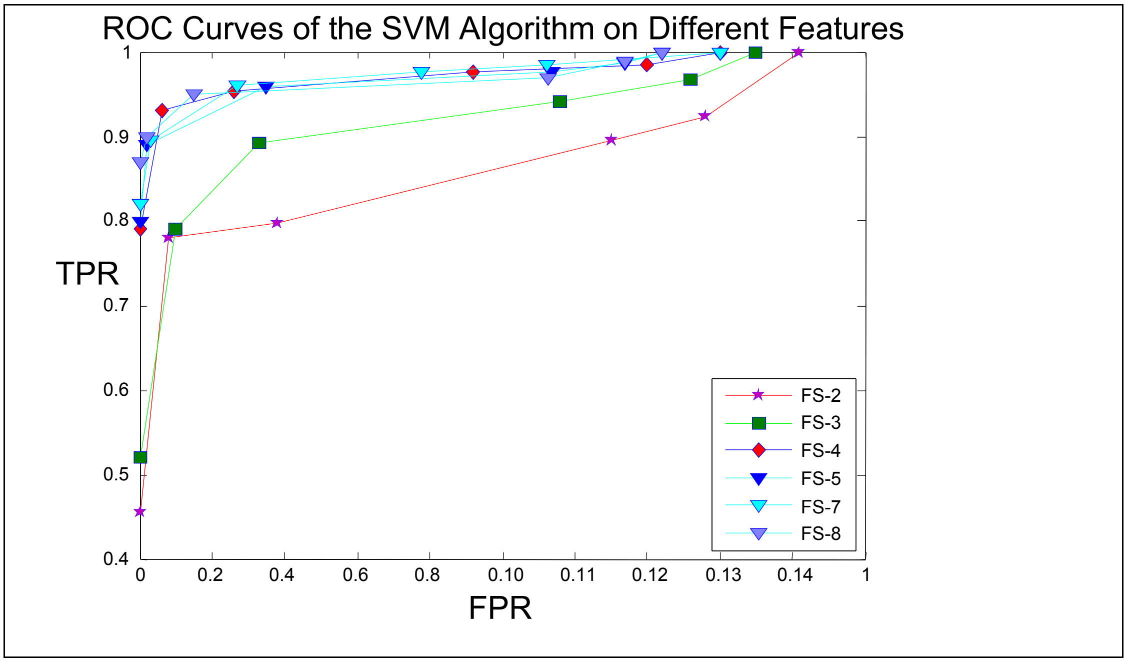


Figure 5. ROC curve of SVM for fusion feature set [2-14] for classification

In the figure 5, we show ROC curve for various numbers of features on the DWT+PCA features. The figure shows that best results are achieved using fourteen features. The performance of SVM classifiers using these features from DWT are shown in above ROC. So we can safely say that best classification can be achieved using fourteen DWT features.

5. Conclusion & Future Work

Proposed system is developed for diagnosing the breast cancer from mammogram images. This system performs this diagnosis in multiple phases. In first phase preprocessing on mammogram image is done to enhance image quality using CLAHE. Then features extraction and reduction has been performed.

Different classifier has been used for classification. All experiments show that the proposed system gives exceptionally good results. In the future, we will perform to classify the malignancy of breast images.

Acknowledgement

The authors extend their appreciation to the Deanship of Scientific Research, College of Science Research at King Saud University for funding this research work.

Corresponding Author:
Muhammad Talha
Denaship of Scientific Research.
King Saud University, Riyadh,
Saudi Arabia
E-mail: mnaseem@ksu.edu.sa

References

[1] M J M Broeders and A L M Verbeek. Breast cancer epidemiology and risk factors. *Q-J-Nucl-Med*, 41(3):179–88, 1997

[2] Nawazish Naveed, Tae Sun Choi, M. Arfan Jaffar, Malignancy and abnormality detection of mammograms using DWT features and ensembling of classifiers, International Journal of the Physical Sciences Vol. 6(8), pp. 2107–2116,18 April, 2011

[3] M. Wallis, M. Walsh, and J. Lee, "A review of false negative mammography in a ymptomatic population," Clin. Radiology, vol. 44, pp. 13-15, 1991

[4] J. Tang, R. Rangayyan, J. Xu, I. El Naqa, Y. Yang, Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances, Information Technology in Biomedicine, IEEE Transactions on 13 (2) (2009) 236–251

[5] G. Kom, A. Tiedeu, M. Kom, Automated detection of masses in mammograms by local adaptive thresholding, Comput. Biol. Med. 37 (1) (2007) 37–48

[6] N. Eltonsy, G. Tourassi, A. Elmaghraby, A concentric morphology model for the detection of masses in mammography, Medical Imaging, IEEE Transactions on 26 (6) (2007) 880–889

[7] Nawazish Naveed, M. Arfan Jaffar, Tae Sun Choi, Segmentation and texture-based classification of breast mammogram images, Microscopy Research and Techniques, DOI: 10.1002/jemt.21070
[8] X. Wang, B. Zheng, W.F. Good, J.L. King, and Y. Chang. Computer assisted diagnosis of breast cancer using a data-driven bayesian belief network. International Journal of Medical Informatics, 54(2):115–126, May 1999.
[9] S. Timp. Analysis of Temporal Mammogram Pairs to Detect and Characterise Mass Lesions. PhD in medical sciences, Radboud University Nijmegen, 2006
[10] M. Mavroforakis, H. Georgiou, N. Dimitropoulos, D. Cavouras, and S. Theodoridis, Significance analysis of qualitative mammographic features, using linear classifiers, neural networks and support vector machines, European Journal of Radiology, 54(1):80–89, 2004.
[11] Antonis Daskalakis, et al, "An Efficient CLAHE-Based, Spot Adaptive, Image Segmentation Technique for Improving Microarray Genes' Quantification", 2nd International Conference on Experiments/Process/System Modelling/Simulation & Optimization, Athens, 4-7 July, 2007.
[12] G. Strang, "The Discrete Cosine Transform," SIAM Review, Volume 41, Number 1, 1999, pp.135-147
[13] Jolliffe, I. T. (1986). Principal Component Analysis. Springer-Verlag. pp. 487
[14] Stephen M. Stigler, The history of statistics, Inverse Probability, Harvard University press, 1986, pp. 94-126
[15] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern classification, 2nd edition, John Wiley and Sons, New York, 2001
[16] Qurat-ul Ain, Syed M. Naqi and M. Arfan Jaffar, Bayesian Classification Using DCT Features for Brain Tumor Detection, Lecture Notes in Computer Science, 2010, Volume 6276, Knowledge-Based and Intelligent Information and Engineering Systems, Pages 340-349
[17] J Suckling et al (1994): The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series 1069 pp375-378
[18] Campanini R, Dongiovanni D, Iampieri E, Lanconelli N, Masotti M, Palermo G, Riccardi A, Roffilli M (2004). A novel featureless approach to mass detection in digital mammograms based on support vector machines. Phys. Med. Biol., 49: 961-975
[19] Guo Q, Shao J, Ruiz V (2005). "Investigation of support vector machine for the detection of architectural distortion in mammographic images." J. Phys. Conf. Ser., 15: 88–94
[20] Miller P, Astley S (1994). "Automated detection of breast asymmetry using anatomical features," in State of the Art in Digital Mammographic Image Analysis. Ser. Machine Perception Artif. Intell., 9: 247-261
[21] C Daniels, G H Schaller, N L McNeil, P Salvador, T Matcalfe, J S Caines, and N A Wadden. Fundamentals of breast imaging, a multimedia teaching file for residents, technologists, radiologists and physicists. Radiology dept. Dalhousie University, Halifax, Nova Scotia; Radiology dept. University of Toronto, 1997. ISBN: 0-7703-2806-7.
[22] N Karssemeijer. Automated classification of parenchymal patterns in mammograms. Phys Med Biol, 43:365–378, 1998.

5/22/2013