# Enhanced Breast Mammograms Classification using Bayesian Classifier

Muhammad Talha [1, 2,] Ghazali Bin Sulong[2], Abdulrahman Alarifi[3]

[1]College of Science Research Centre, King Saud University, Riyadh, Saudi Arabia.
[2]Department of Computer Graphics and Multimedia, Faculty of Computer Science and
Information System, Universiti Teknologi Malaysia, Malaysia.
[3]Computer Research Institute, King Abdulaziz City for Science and Technology, Riyadh, Saudi
Arabia.

mnaseem@ksu.edu.sa, ghazalis@utm.my, aarifi@kacst.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. After that local window based discrete cosine transform has been applied for features extraction. Principle component analysis has used for features reduction and selection. Bayesian Classifier has 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.

**Keywords:** Classification, breast cancer, mammogram, enhancement, Bayesian

## 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.

## 2 Proposed Method

The proposed system is divided into four major parts as shown in Fig. 1:
- Enhancement by using CLAHE
- Features Extraction & Reduction
- Classification

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



**Fig. 1**: Proposed Method

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.

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. Discrete cosine transform (DCT) is used for converting the signal into its frequency components. In image processing DCT attempts to de-correlate the image data [12]. DCT has the ability to pack the image data into as few DCT coefficients as possible without any distortion. DCT has the property of separability and symmetry.

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].

For estimating the class of the new data item a probabilistic model is defined which is known as Bayesian classification. Bayesian classifier is a statistical classifier. Bayesian classification is based on the Bayes theorem [14]. Bayesian classification is used for classifying objects into associated classes based on the attributes of these objects. Attributes of the data/object are considered as independent of each other in Naive Bayes classification [15][16].

# 3    Conclusions

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. Bayesian 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

## 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.