

Breast Cancer Diagnosis by CAD

Nidhal K. El Abbadi
University of Kufa
Najaf, Iraq

Elaf J. Al Taei
University of Kufa
Najaf, Iraq

ABSTRACT

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death of female worldwide. Mammogram is one of the most excellent technologies currently being used for diagnosing breast cancer. Computer aided diagnosis helps the radiologists to detect abnormalities earlier than traditional procedures. In this paper, we suggested to use some of features selected to distinguish the benign and malignant breast cancer. Tumor segmented and denoising prior to classification. The accuracy of proposed system was 100%.

General Terms

Image processing, Pattern recognition, Medical image processing.

Keywords

Breast Cancer, Mammography, Denoising, Diagnosis, Image Features.

1. INTRODUCTION

The uncontrolled division of one cell may cause to emerge mass called tumor. Tumor can be benign or malignant. Benign tumor cannot spread or invade to other parts of the body while malignant tumor grows rapidly and invades its surrounding tissues through causing their damage. Breast cancer is a malignant tissue beginning to grow in the breast. The abnormalities like existence of a breast mass change in shape and dimension of breast, differences in the color of breast skin, breast aches, etc. are the symptoms of breast cancer [1].

Breast cancer is a disease of humans and other mammals; while the overwhelming majority of cases in humans are women, men can also develop breast cancer. Mammography is the most used screening tool for abnormality detection, because it allows an easy way to identify the cancer.

The goal of mammography is the early detection of breast cancer, typically through detection of characteristic masses and/or micro-calcifications. Radiologists interpret the mammograms and attempt to identify areas of potential abnormalities.

Therefore, the expert radiologist's is the corner stones of diagnosis the breast cancer; the screening process depends on the radiologist's ability to detect areas of subtle irregular abnormalities. It is estimated that between 10-30% of women diagnosed with breast cancer have false-negative mammograms [2].

Most of the false-negative cases can be attributed to the radiologist's failure to detect a cancer which could be due to misinterpretation, or simply that the radiologist overlooked the area.

It has been demonstrated that an independent second reading can significantly improve the detection rate and decrease the number of false positive cases.

Computerized aided can help the physicians, and radiologist's to detecting and diagnosing the breast cancer; it is a tools to processing and analyzing images as secondary reading. The first step for CAD is the ability to identify the abnormal masses in the breast, while the second step is to diagnosis the masses detected in the first step. Before these two steps implementing a really important preprocessing step has to take place which is the detection or segmentation of the breast region from the background. The analysis of mammogram image can assist the radiologist for early detection of tumor and diagnosis the breast cancer effectively [3].

Computer aided can classify into two main groups: computer-aided detection (CADE) and computer-aided diagnosis (CADx). The first group CADE works with medical images to detect and localize the lesions in the image. While CADx works directly on the lesions to classify them to benign and malignant, it is diagnosis process [4].

2. RELATED WORKS

Patel et al. suggested new method for breast image segmentation. In order to early detect of breast cancer based on detection of micro calcification and computer based decision system. The paper developed an adaptive k-means clustering algorithm [5].

Dromain et al. pointed that the mammograms image suffers from false negative mammograms which can be caused by subtle findings, radiologist distraction or complex architecture. The paper try to reduce these false negative by analyzing the image with computer aided. CAD works as a tool to help radiologist to evaluate screening mammograms [6].

R. Ramani presented new approach to classify masses in the Mammography Image. Feature extracted from mammography achieved by using Symlet, SVD, and weighted histogram. Naïve bayes, random forest and neural network algorithms used to classify extracted features [7].

3. MATERIALS AND METHODS

Breast cancer is a common cancer among women. Though, early detecting of abnormal masses can result in successful treatment. There are many steps in diagnosis of breast cancer, start from the segmentation and ended by tumor classification. The diagnosis requires precise and reliable diagnosis to ensure that doctors can distinguish between benign and malignant tumors. Mammography widely used as an effective imaging modality for tumor classification. Widespread mammography use needs aids from computer technology. Fig 1 shows block diagram for proposed computer-aided diagnosis system for diagnosing breast cancer.

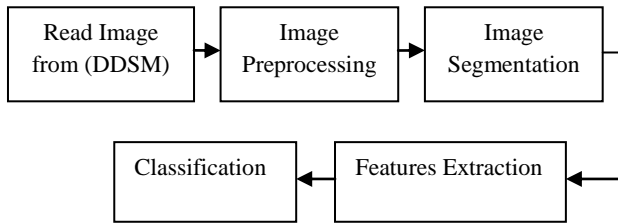


Fig 1: Block Diagram for proposed System.

3.1 Digital Mammograms

The Digital Mammograms images used in the current research are taken from Digital Database for Screening Mammography (DDSM) established by University of South Florida. This is free available database. The suspicious area marked already from the expert, and then all the images used are tumor detecting and localized.

In the current research we select 100 images from database which belong to two types, benign and malignant (50 image as benign and 50 image as malignant). The sampling rates of mammogram images are digits between 42 to 50 microns, as shows in Fig 2.

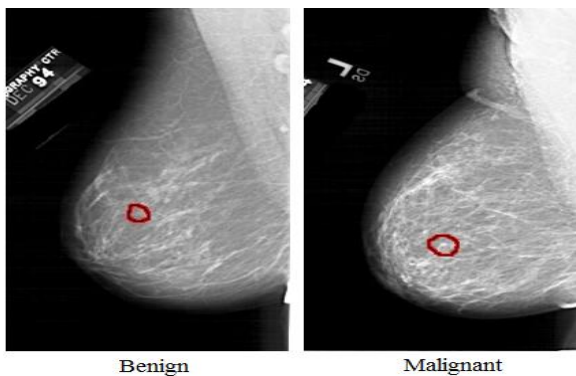


Fig 2: Breast Image

3.2 Image Preprocessing

The main objective of the preprocessing is to improve the image quality to make it ready for further processing by removing the unrelated and surplus parts in the background of the mammogram images. Mammograms are medical images that complicated to interpret. The noise removed by compute foreground markers, these are connected blobs of pixels within each of the foreground objects. Noise removing based on morphological techniques called "opening-by-reconstruction" and "closing-by-reconstruction". The opening-by-reconstruction computed by erosion followed by a morphological reconstruction and "closing-by-reconstruction" computed by dilation followed by a morphological reconstruction, the resulted images from this step shown in Fig. 3.

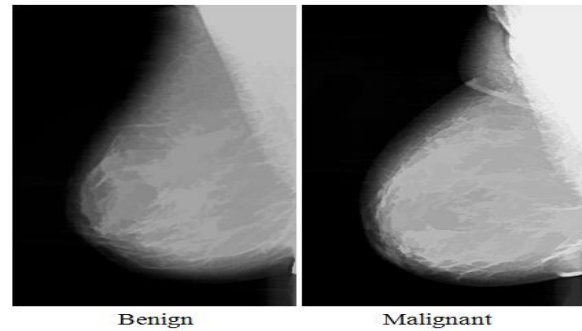


Fig 3: De-noising images

3.3 Image Segmentation

Image segmentation is the fundamental approach of digital image processing. Among all the segmentation methods, Otsu method is one of the most successful methods for image thresholding due to automatically compute the threshold. Thresholding helps to create binary images from grey-level image by converting all pixels value below a threshold to zero and the remaining pixels to one, as shown in Fig 4.

Binary image help to remove all the background and leave only the breast prior to segment the tumor from the breast (we will work on the tumor only to detect the benign or malignant breast) as shown in Fig 5. Tumor will be segmented by reconstruct the image from binary image and the result showed in Fig 6.

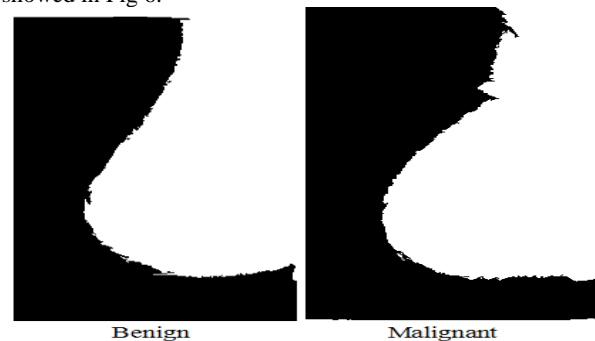


Fig 4: Breast binary image.\

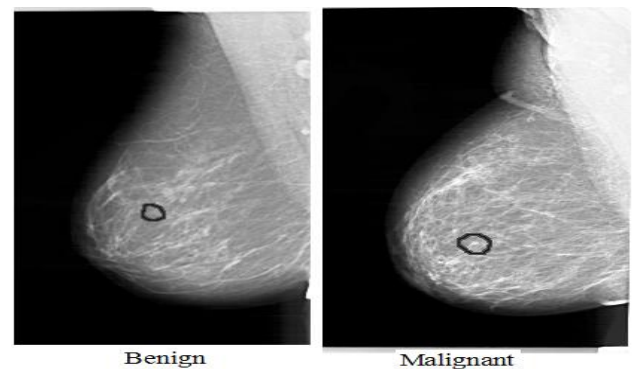


Fig 5: Detection tumor in breast images.\

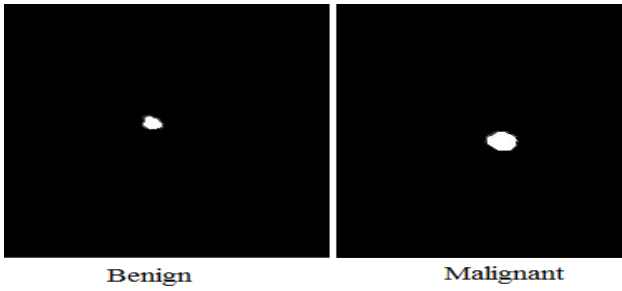


Fig 6: Segmented tumor from breast images

3.4 Features Extraction

Statistical features extraction is important steps in breast cancer detection and classification. Among the statistical features the following first-order statistics is used as texture features in representing images. First-order texture measures are calculated from the original image values. They do not consider the relationships with neighborhood pixel.

Histogram-based approach to texture analysis is based on the intensity value concentrations on all or part of an image represented as a histogram. Features derived from this approach include moments such as mean, variance, standard deviation and entropy. The histogram of intensity levels is a simple summary of the statistical information of the image and individual pixels are used to calculate the gray-level histogram. Therefore, the histogram contains the first-order statistical information about the image (or sub image). These statistics are defined by the following relations, and are applied to images in Fig 6.

$$\text{Mean } (\mu_i) = \frac{\sum_{x=1}^M \sum_{y=1}^N I_i(x, y)}{M \times N}$$

$$\text{Variance } (\sigma^2) = \frac{\sum_{x=1}^M \sum_{y=1}^N (I_i(x, y) - \mu)^2}{M \times N}$$

$$\text{Standard Deviations } (\sigma_i) = \sqrt{\frac{\sum_{x=1}^M \sum_{y=1}^N (I_i(x, y) - \mu)^2}{M \times N}}$$

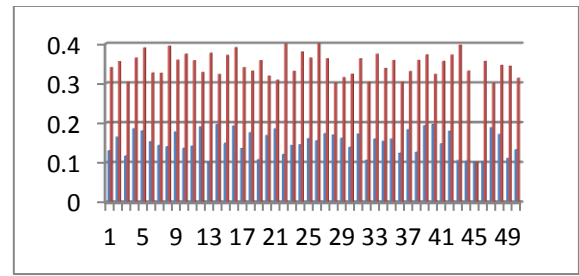
$$\text{Entropy}_i = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N I_i(x, y) (-\ln I_i(x, y))$$

3.5 Classification

After the features extracted, mammogram images classified into benign or malignant images depending on values of features extracted. The features values highly varied between benign and malignant.

4. RESULTS AND DISCUSSIONS

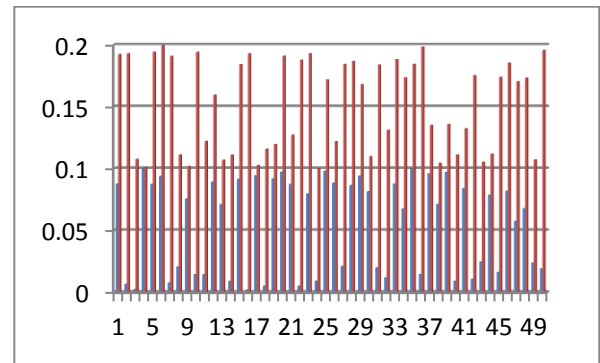
Images from the DDSM dataset mammograms used to evaluate the proposed method. 100 images selected from the database which is 50 images benign and 50 images malignant. The features extracted from the tumor were varied from the benign to malignant as shown in figures (7, 8, 9, and 10).



Mean of benign Mean of Malignant

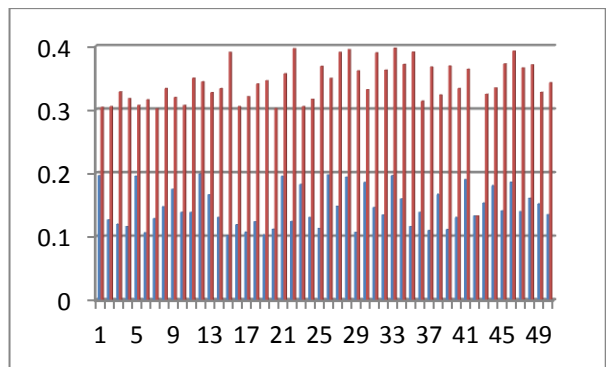


Fig 7: Mean to Benign and Malignant Breast Cancer



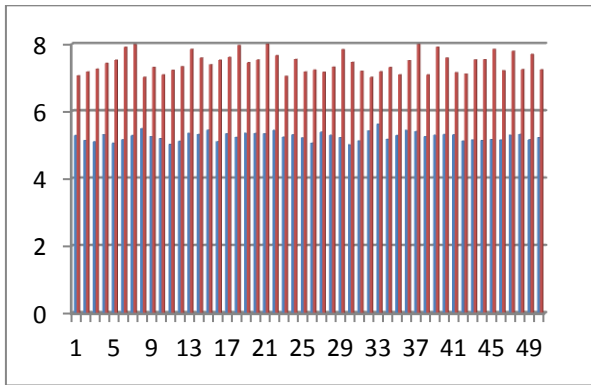
Variance to Benign Variance to Malignant

Fig 8: Variance to Benign and Malignant Breast Cancer



Std to Benign Std to Malignant

Fig 9: Standard Deviation to Benign and Malignant Breast Cancer



■ Entropy to Benign ■ Entropy to Malignant

Fig 10: Entropy to Benign and Malignant Breast Cancer

Precision and **recall** are the basic measures used in evaluating search strategies and they computed as follows:

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{total No. of images that are relevant}} \times 100$$

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{total No. of images retrieved}} \times 100$$

The precision and recall value for images benign and malignant breast cancer are 100%.

In addition, the suggested algorithm compared with other algorithms and the suggested algorithm was better than the others were as shown in table 1.

Table 1: Comparing the performance of proposed algorithm with other algorithms.

Papers	Classification methods	Classification Accuracy (%)
Breast cancer diagnosis using least square support vector machine	LS-SVM	98.53
Diagnose Breast Cancer through Mammograms Using EABCO Algorithm	EABCO and Bilateral Subtraction	96.40
Evaluation of Breast Cancer with a Computer-Aided Detection System by Mammographic Appearance and Histopathology	Histopathology	95
Detection of Breast Cancer with Full-Field Digital Mammography and Computer-Aided Detection	FFDM	94
A Computer Aided Diagnosis System for Breast Cancer	K-nearest neighbors, probabilistic neural networks	98.80 96.33
Computer Aided Detection of Tumours in Mammograms	naïve bayes, random forest, neural network	81.21 86.67 89.7
The proposed system	Results of mean, variance, standard deviation and entropy	100

5. CONCLUSION

This paper suggested to use the Computer aided diagnosis systems to help the doctors in detection/diagnosis of abnormalities faster than the traditional procedures. The automated mammogram classification method suggested based on first-order statistics used for features extraction post of tumor segmented. Number of features selected to distinguish between benign and malignant breast cancer, such as mean, entropy, standard deviation, and variance. The selected features worked very well and give very good results.

The proposed algorithm tested with 100 images and gives very promised results. The precision and recall was 100% for both benign and malignant. Also the algorithm performance was better than the other algorithms as shown in the table 1.

6. REFERENCES

- [1] Polat K., and Gne S. 2006 Breast Cancer Diagnosis Using Least Square Support Vector Machine, Elsevier, DOI: 10.1016/j.dsp.2006.10.008.
- [2] Cheng, H.D., Cai, X., Chen, X.W., Hu, L., and Lou, X., 2003 Computer Aided Detection and Classification of Micro Calcifications in Mammograms: A Survey, Pattern Recognition, vol. 36, pp: 2967–2991.
- [3] Sivakumar R., and Karnan M. 2012 Diagnose Breast Cancer through Mammograms Using EABCO Algorithm, International Journal of Engineering and Technology (IJET), Vol 4, No 5.

- [4] Suzuki K., 2012 A review of computer-aided diagnosis in thoracic and colonic imaging, *AME*, **DOI:** 10.3978/j.issn.2223-4292.2012.09.02.
- [5] Bhagwati Charan Patel, and G. R. Sinha, 2010 An Adaptive K-means Clustering Algorithm for Breast Image Segmentation, *International Journal of Computer Applications*, Volume 10, No 4. **DOI:** 10.5120/1467-1982
- [6] Dromain, C., Boyer, B., Ferré, R., Canale, S., Delalogue, S., Balleyguier, C., 2012 Computer Aided Diagnosis (CAD) in the Detection of Breast Cancer, *European Journal of Radiology* 82, pp 417– 423. Epub 2012, **DOI:** 10.1016/j.ejrad.2012.03.005.
- [7] Ramani, R. Suthanthira, N. Vanitha, 2014 Computer Aided Detection of Tumours in Mammograms, *International Journal of Image Graphics and Signal Processing*, 4, 54-59, **DOI:** 10.5815/ijigsp.2014.04.07
- [8] Amjath Ali J. and J. Janet, 2013 Mass Classification in Digital Mammograms Based on Discrete Shearlet Transform, *Journal of Computer Science* 9 (6): 726-732.