British Journal of Applied Science & Technology
9(1): 53-64, 2015, Article no.BJAST.2015.246
ISSN: 2231-0843

SCIENCEDOMAIN *international*
www.sciencedomain.org

Automated Detection of Breast Cancer's Indicators in Mammogram via Image Processing Techniques

J. O. Olaleke[1*], A. O. Adetunmbi[2], O. O. Obe[2] and O. G. Iroju[1]

[1]Computer Science Department, Adeyemi College of Education, Ondo, Ondo State, Nigeria.
[2]Computer Science Department, Federal University of Technology, Akure, Ondo State, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Author JOO initiated the study and wrote the first draft of the manuscript. Authors AOA and OOO supervised the study while, author OGI assisted in the design of the system. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJAST/2015/13675
Editor(s):
(1) Patrice Wira, Faculty of Science and Technology, University of Haute Alsace, France.
Reviewers:
(1) Sadaf Alipour, Department of Surgery, Arash Women's Hospital, Tehran University of Medical Sciences, Tehran, Iran.
(2) Anonymous, USA.
(3) Gustavo Machado Badan, Breast Radiology, Santa Casa de Misericórdia de São Paulo, Brazil.
(4) Anonymous, USA.
(5) Saylee Gharge, V. E. S. Institute of Technology, University of Mumbai, India.
(6) Marcelo Duarte, Industrial Automation, SENAI Rio Faculty, Brazil.
Complete Peer review History: http://www.sciencedomain.org/review-history.php?iid=1136&id=5&aid=9161

Original Research Article

Received 27th August 2014
Accepted 15th April 2015
Published 8th May 2015

ABSTRACT

Aims: The detection of abnormalities in mammographic images is an important step in the diagnosis of breast cancer. The indicators of cancer in mammograms can be in form of calcification, mass and stellate lesion. This paper proposed a two-stage procedure for the detection of these cancer's indicators.
Methodology: Twenty images were used for the study. The images were obtained from Mammographic Image Analysis Society (miniMIAS) database. The images were pre-processed and enhanced using top hat filtering method and the enhanced images were segmented using Otsu's method. Four features were extracted and selected from the mammographic images using Gray Level Concurrence Matrix (GLCM). The features extracted and selected include energy, homogeneity, contrast, and correlation. Subtractive clustering and fuzzy logic techniques were employed for the classification of the cancer's indicators in the mammograms. The implementation of the image processing techniques was done with matrix laboratory.
Results: The result showed that seven of the images were affected by stellate lesion, nine of the

**Corresponding author: E-mail: shollyjane@yahoo.com;*

images were affected by microcalcification while four of the images were affected by mass.
Conclusion: The method presented in this paper would enhance the detection of cancerous cells in the breasts.

Keywords: Mammogram; breast cancer; image processing; fuzzy logic; subtractive clustering.

1. INTRODUCTION

Breast cancer, a malignant tumour in the glandular tissue of the breast is one of the major causes of deaths among women [1,2]. These tumours are referred to as carcinomas which occur when the processes that control the normal cell growth break down, thus enabling a single abnormal cell to multiply at a rapid rate, thereby destroying a proportion of the normal breast tissue over time [3]. Different image modalities are used in the detection and evaluation of breast abnormalities. The most common of these modalities is called the mammography. The result obtained from the mammography is called the mammogram. Mammogram is capable of showing indicators that help to determine the likelihood of cancerous cells in the breasts. These indicators include the mass, the calcifications and the stellate lesion amongst others. A breast mass can be defined as a localized swelling, protuberance, or lump in the breast, which usually is described by its location, size, shape, margin characteristics, and any other associated findings such as distortion and X-ray attenuation. Calcifications are very small bits of calcium that appears within the soft tissue of the breast [3]. They usually appear as white dots on the mammogram. Calcifications are divided into two kinds. These are macro calcification and micro calcification. Macro calcification is a coarse calcium deposit in the breast. Macro calcification is usually caused by natural aging of the breast or it can be as a result of previous injuries or inflammation in the breast. Macro calcification is usually harmless and it is not linked to breast cancer [3]. Macro calcification is not benign and it does not require any form of monitoring or treatment. Micro calcification on the other hand can be used as the primary indicator of breast cancer detection [4]. It may appear alone or in clusters. Its layout and shape can assist radiologists to determine the occurrence of cancer in the mammogram. The detection of stellate lesions is very important in the characterization of breast cancer [3]. Stellate lesion often appears as a central mass surrounded by spicules radiating outwards. In general, it is difficult to detect because it has an irregular centre with ill-defined borders radiating spicules that may extend from several millimeters to centimeters in size [3]. As a cancer cell proliferates, it shows up as a star-shaped with spiky lines radiating in all directions from a central region. A white star shape is a characteristic of a malignant stellate lesion whereas the black star indicates a radial scar and post-traumatic fat necrosis.

In spite of the benefits of the mammogram in the detection of abnormal cells in the breast, its accuracy is not guaranteed [2]. For instance, scientists estimate that mammograms miss about 25 percent of breast cancer in women of 40 to 49 years of age and about 10 percent of breast cancer in older women is missed in mammograms [3]. In addition, mammograms contain low signal to noise ratio (low contrast) and a complicated structured background [3]. Owning to these limitations, there is a need for a better method that will facilitate the detection of this terminal disease. Hence, this paper proposes an image-based, efficient and automated method that will enhance the detection of cancerous cells in the breasts.

2. METHODOLOGY

Twenty mammographic images were used for this study. These images were obtained from the miniMIAS (Mammographic Image Analysis Society) database. The miniMIAS database is provided by the Mammographic Image Analysis Society (MIAS) in the United Kingdom (UK). It has 200 micron resolution [5].

Five of the images were affected by masses, seven images were affected by micro calcifications and eight of the images were affected by satellite lesions.

2.1 Image Selection Criteria

The original images were selected purposively from the miniMIAS database. This is to ensure that the selected images were affected by the three cancer indicators. Images a, c, f, h, k, p, r and t were affected by stellate lesion, images b, d, e, g, i, q and s were affected by calcification while j, l, m, n and p were affected by masses.

2.2 Description of the Tool Employed

The image processing tool box in matrix laboratory (Matlab) was employed. This is because the Image Processing Toolbox provides a comprehensive set of functions for image manipulation, analysis, digital imaging, computer vision, and digital image processing. Its capabilities include colour space transformations, linear filtering, mathematical morphology, texture analysis, pattern recognition and image statistics.

2.3 Proposed Model

The proposed model is a two-stage detection procedure as shown in Fig. 1. The purpose of the first stage is to determine the location and border of the suspected areas in the mammogram and to obtain an enhanced image [5]. The image enhancement was achieved with the use of top hat filtering method. In addition, the top hat filtering method was used to segment objects that differ in brightness from the surrounding background in images with uneven background intensity. The Otsu method was used to segment the enhanced borders from the background image. Morphological closing was applied to the image result of the Otsu method to eliminate noise [6]. The second stage involved the extraction of four features from the images using Gray Level Concurrence Matrix (GLCM) [7]. The features extracted and selected included energy, homogeneity, contrast, and correlation. The classification of the cancer type was done using subtractive clustering and fuzzy logic techniques.

2.3.1 Image Conversion

The original images were converted to gray scale images. This is because medical images are better worked upon in gray scale and they also reduce the memory location of each image. The syntax used for the image conversion in Matlab is as shown in Equation (1).

$$I = \text{rgb2gray (RGB)} \quad (1)$$

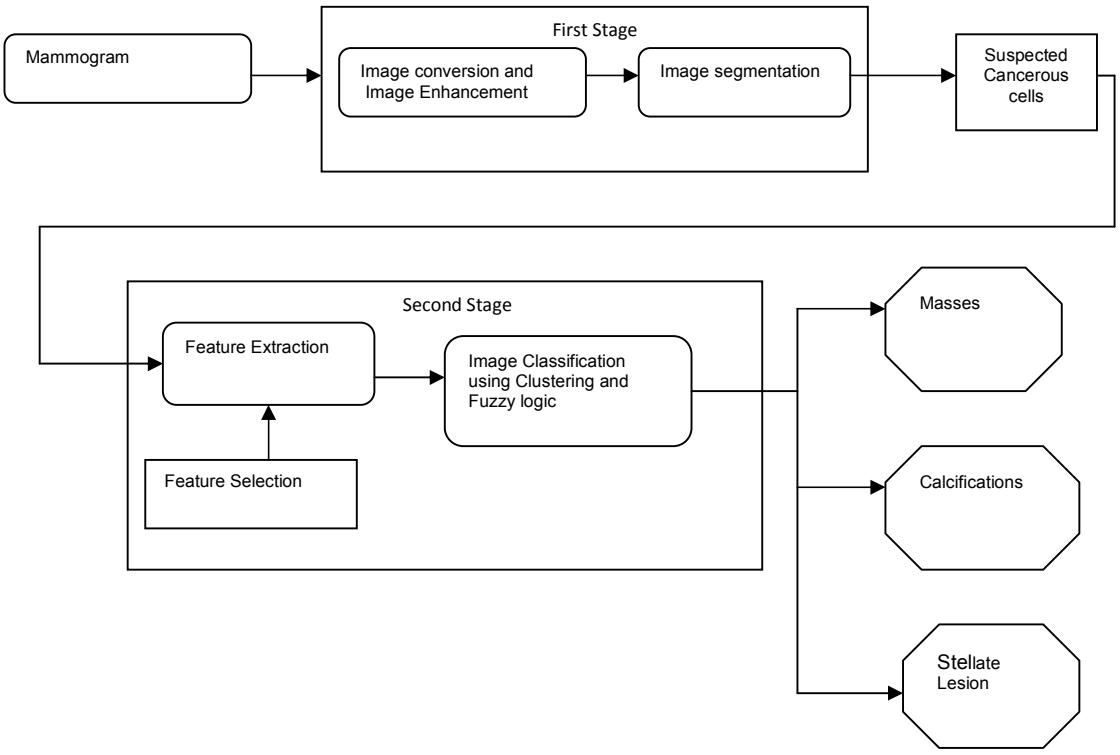

Fig. 1. Flow chart of proposed model

Rgb2gray converts the true colour image rgb (red green blue) to the grayscale intensity image I. Rgb2gray also converts RGB images to grayscale by eliminating the hue and saturation information while retaining the luminance [8]. For the conversion from RGB to grayscale image, the mathematical function is based on 30%red + 59%green +11%blue.

2.3.2 Image Enhancement

The grayscale mammographic images were enhanced using top hat filtering [6]. The top hat filtering permits the extraction of light objects from an uneven background. The proposed model employed the top hat filtering formula in [9] which is given as:

$$B = I - [(I \Theta SE) \oplus SE]$$

Where

I is the gray scale image
B is the enhanced image
SE is the structuring element
Θ is the morphological erosion operation which adds pixels to the boundaries of objects in the image [10]
$\oplus$ is the morphological dilation operation which removes pixels on the image boundaries
$-$ is the image subtraction [10]

The syntax used for the image enhancement is shown in Equations (2) and (3) respectively.

se = strel('disk',12) (2)

J = imtophat (I,se) (3)

where

strel is a disk-shaped structuring element which removes the uneven background illumination from the images.
I is the grayscale image
J is the enhanced image

2.3.3 Image segmentation

The enhanced images were segmented using Otsu's technique [6]. This was used to segment the enhanced borders from the background image. The Otsu's method was employed because it is the most accurate technique in detecting cancer in digital mammograms as it has shown a more satisfactory performance in medical image segmentation [10]. It has also been found to perform well compared to other thresholding methods in segmenting masses in digital mammogram [11]. The equations for the Otsu method are shown in equations (4), (5) and (6) [12].

$$d_W^2(T) = \omega_0 \delta_0^2 + \omega_1 \delta_1^2 \quad (4)$$

Where δ_0 and δ_1 are the variances of the pixels below and above the threshold t, respectively. ω_0 and ω_1 are defined as,

$$\omega_0 = \sum_{i=0}^{T-1} p(i) \quad (5)$$

$$\omega_1 = \sum_{i=T}^{L} p(i) \quad (6)$$

With p(i) indicating the probabilities of the occurrence of gray level i in the image.

Sobel and Canny edge detection was then used to segment the enhanced borders from the background image. The Sobel edge detector applies Sobel approximation to the derivative of the image and detected edges whenever the gradient of reconstructed input image was at its maximum [5]. The Canny edge detector found edges by finding local maximum of the gradient of unprocessed input image. In each edge detection algorithm, the gradient was calculated using the derivative of a Gaussian filter, and the output is a binary image, where 1 represents edges and 0 represents background [5]. The Sobel and Canny binary output images were post processed by the flood-filling operation to fill all objects with closed borders. The flood fill operation fills holes in the binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image [5]. They were then logically ORed together to produce a new image that includes false positives but discourages false negatives. This new image undergoes morphological closing to eliminate noise.

2.3.4 Feature extraction and selection

Four features were extracted from the images using Gray Level Concurrence Matrix (GLCM). These features extracted include contrast, homogeneity, correlation and energy. Contrast is a measure of intensity between a pixel and its neighbour over the entire image. Homogeneity measures the spatial closeness of elements' distribution in the image. The range of values is between [0, 1], with the maximum being achieved when the image is a diagonal matrix. The homogeneity descriptor refers to the closeness of the distribution of elements in GLCM. Correlation is a measure of how correlated a pixel is to its neighbour over the entire image while energy or uniformity is the sum of squared elements in GLCM [9]. Energy may also be referred to as the angular second moment.

The goal of feature selection is to choose the optimal feature vector, which consists of the features that minimize the classification error [9]. The equation for the features extraction is as shown in Equation (7) [12]

$$C(i,j) = |\{m,n\} : P(m,n) = i \; and \; P(m + d.\cos\theta, n + d.\sin\theta = j)| \quad (7)$$

In Equation (7), the co-occurrence matrix C, computed on a gray-level image P, is defined by a distance d and an angle θ. C (i, j) is the number of times that the gray value i co-occurs with the gray value j in a particular spatial relationship defined by d and θ.

A Matlab file containing the features extracted was loaded into the Matlab workspace as the input variables. The file was named datain. The cancer indicators identified from each of the mammograms were written into an output file. The output file tagged dataout was loaded into the Matlab workspace as the output variables. The two variables had five columns. Four of the columns represented the four input variables while one column represented the output variable. The twenty rows in datain and dataout represented the number of observations or samples or data points available. A row in datain constitutes a set of observed values of the four input variables (contrast, homogeneity, correlation, and energy) and the corresponding row in dataout represents the value of cancer indicator given the observations made for the input variables. The relationship between the input variables and the output variable was modeled by clustering the data.

2.3.5 Image classification using subtractive clustering and fuzzy logic

Subtractive clustering was employed to identify natural groupings in data from the data set. This allows the concise representation of relationships embedded in the data [13]. Subtractive clustering is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a dataset [13]. The syntax used for generating this is:

```
[C,S] = subclust([datain dataout],0.5)
```

where C is the cluster center and S is the sigma value that specifies the range of influence of a cluster center in each of the data dimensions. All cluster centers share the same set of sigma values and 0.5 is the radius which marks the cluster's radius of influence in the input space. S has 5 columns representing the influence of the cluster centers on each of the 5 dimensions. This clustering technique allowed the groupings of the input variables into broad categories of the output variable, that is, cancer indicators hence allowing for easy classification.

Fuzzy logic was then employed to capture the broad categories identified during clustering into a Fuzzy Inference System (FIS). Genfis2 is the function that creates a FIS using subtractive clustering. Genfis2 employs subclust behind the scenes to cluster the data and it uses the cluster centers and their range of influences to build a FIS. The FIS acts as a model that reflects the relationship between the input variables and output variables. The FIS is composed of inputs, outputs and rules. Each input in the FIS represents an input variable in the input dataset and each output in the FIS represents an output variable in the output dataset. Therefore each input and output was characterized by eight membership functions. The membership function type is a Gaussian type membership function and the parameters of the membership function are [0.0051 0.2196], where 0.0051 represents the spread coefficient of the Gaussian curve and 0.2196 represents the center of the Gaussian curve.

3. RESULTS

Fig. 2 shows the original images from miniMIAS database

3.1 Result of Image Conversion

The result of the image conversion is shown in Fig. 3.

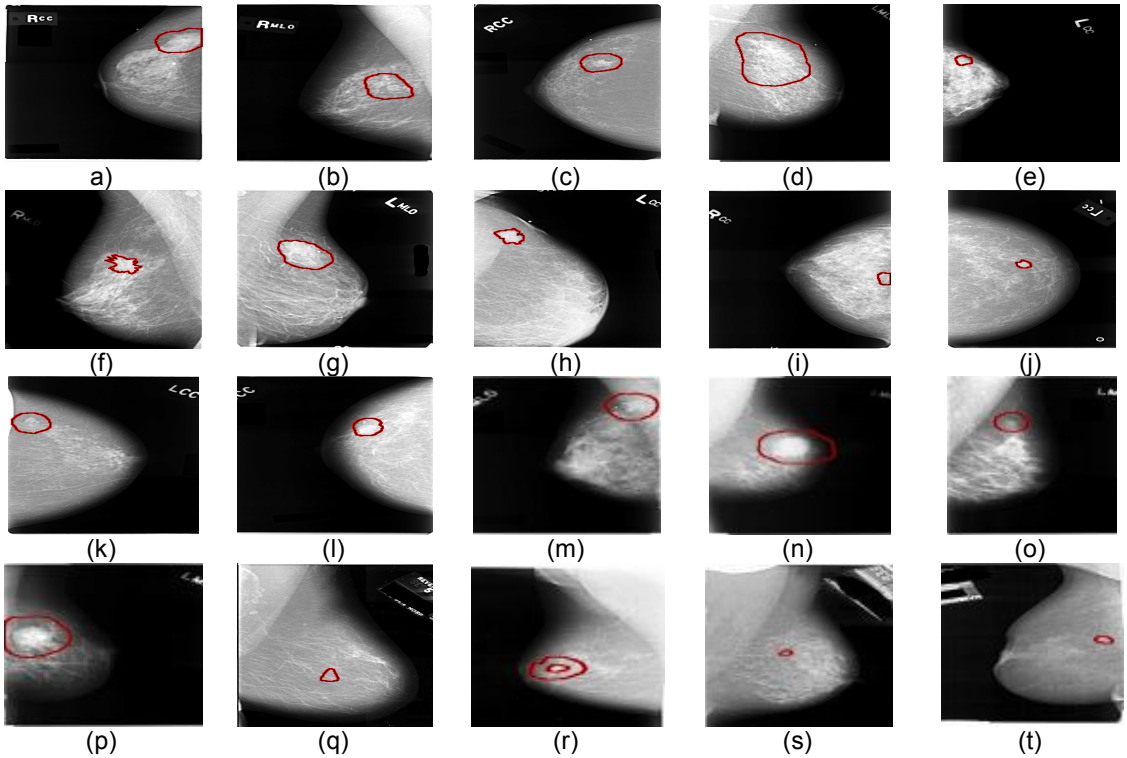

Fig. 2. The original images from miniMIAS database

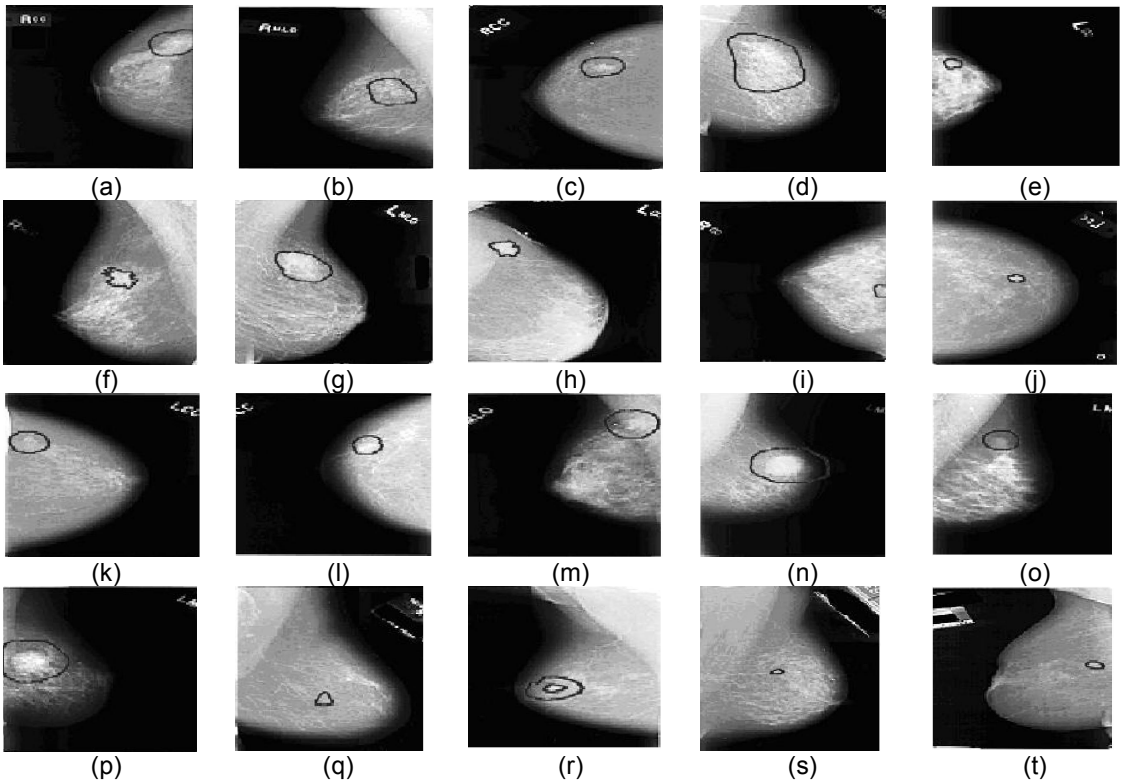

Fig. 3. The grayscale images

3.2 Result of the Image Enhancement

The result of the image conversion is shown in Fig. 4.

3.3 Result of the Image Segmentation

The result of the image segmentation is as shown in Fig. 5.

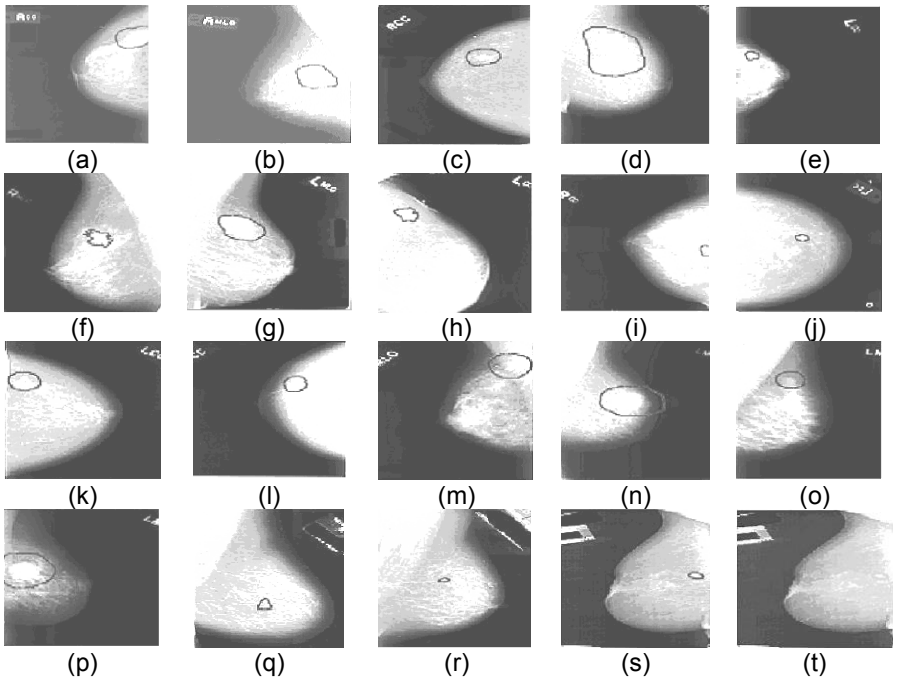

Fig. 4. The enhanced images

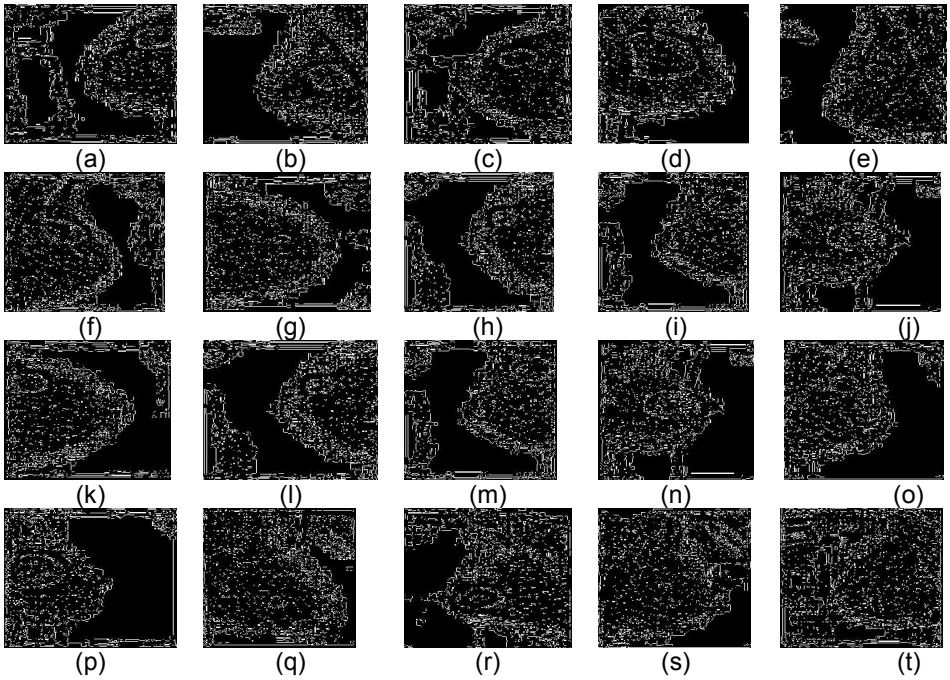

Fig. 5. The segmented images

3.4 Result of Feature Extraction and Selection

Table 1 shows the results of the features extracted and selected.

3.5 Result of Image Classification using Subtractive Clustering and Fuzzy Logic

Tables 2 and 3 show the results of the cluster center and sigma values respectively. Table 4 shows the rules Generated from the FIS. Table 5 shows the output of the rules used to generate the output of the FIS.

In Table 4, In1cluster1, in1cluster2, in1cluster3, in1cluster4, in1cluster5, in1cluster6, in1cluster7 and in1cluster8 respectively capture the position and influence of the clusters for the input variable contrast. Similarly, the other seven inputs follow the exact pattern mimicking the position and influence of the other seven clusters along their respective dimensions in the dataset. The rules dictate the behaviour of the fuzzy system based on the inputs variables and output variable [13]. Therefore, the number of rules equals the number of clusters and since eight clusters were created, eight rules were generated. The first rule indicates that if the inputs to the FIS, that is, contrast, homogeneity, correlation, and energy strongly belong to their respective cluster1 membership function then the output, that is cancer indicator, must strongly belong to its cluster1 membership function. The (1) at the end of the rule is to indicate that the rule has a weight or an importance of 1. Weights take its values between 0 and 1. Rules with lesser weights count for less in the final output. The significance of the rule is that it concisely maps cluster 1 in the input space to cluster 1 in the output space. Similarly, the other seven rules map cluster2, cluster3, cluster4, cluster5, cluster6, cluster7 and cluster8 respectively in the input space to cluster2, cluster3, cluster4, cluster5, cluster6, cluster7 and cluster8 respectively in the output space. If a data point closer to the first cluster has a strong membership to the first cluster, it is fed as an input to the FIS, then rule1 will fire with more firing strength than the other seven rules. Similarly, an input with strong membership to the second cluster will fire the second rule with more firing strength than the other seven rules and so on. The output of the rules (firing strengths) is then used to generate the output of the FIS. The result generated from FIS was also supported by the mathematical parameters proposed by [7].

The experimental result was generated based on the following mathematical parameters:

$$\text{Mass}(c, h, c_0, e) = \begin{cases} \text{if } 0.20 <= c <= 0.25 \\ 0.60 <= h <= 0.65 \\ 0.70 <= c_0 <= 0.85 \\ 0.50 <= e <= 0.60 \end{cases}$$

$$\text{calcification}(c, h, c_0, e) = \begin{cases} \text{if } 0.25 <= c <= 0.30 \\ 0.60 <= h <= 0.65 \\ 0.85 <= c_0 <= 0.90 \\ 0.65 <= e <= 0.70 \end{cases}$$

$$\text{stellate}(c, h, c_0, e) = \begin{cases} \text{if } 0.30 <= c <= 0.35 \\ 0.60 <= h <= 0.65 \\ 0.90 <= c_0 <= 1.00 \\ 0.50 <= e <= 0.60 \end{cases}$$

where c = contrast
h = homogeneity
c_0 = correlation
e = energy

4. DISCUSSION

A breast cancer detection system has been introduced in this paper. The breast cancer detection system employed image processing techniques such as top hat filtering, Otsu method, Canny and Sobel method, Gray Level Concurrence Matrix, subtractive clustering and fuzzy logic techniques. The proposed breast cancer detection system is a two stage detection procedure which was based on [9]. The major distinction between the existing system and the proposed system is that the existing system was used for microcalcification detection while the proposed system dealt with the detection of three cancer indicators which are microcalcification, stellate lesion and masses. The proposed system was tested with twenty mammographic dataset obtained from the miniMIAS database. These mammographic images were affected by the three cancer indicators. Seven of the images were affected by stallete lesion, eight of the images were affected by microcalcifications while five of the images were affected by mass. However, the result of the proposed system showed that seven of the images were affected by stellate lesion; nine of them were affected by microcalcifications while four of the images were affected by mass. This showed a discrepancy between the result of original mammographic images and those of the proposed system. For instance, the result from the miniMIAS database showed that images c and j were affected by stellate lesion and mass respectively while the proposed system showed that images c and j were affected by microcalcification.

Table 1. Features extracted and selected

Images	Contrast	Homogeneity	Correlation	Energy
a	0.3418	0.9684	0.9253	0.5822
b	0.2520	0.6102	0.8632	0.6675
c	0.2741	0.6314	0.8710	0.6645
d	0.2513	0.6500	0.8512	0.6551
e	0.2704	0.6301	0.8711	0.6345
f	0.3015	0.9854	0.9668	0.5517
g	0.2815	0.6048	0.8519	0.6748
h	0.3398	0.9934	0.9550	0.5238
i	0.2679	0.6456	0.8602	0.6528
j	0.2553	0.6080	0.8718	0.6900
k	0.3458	0.9545	0.9420	0.5904
l	0.2171	0.6312	0.7015	0.5022
m	0.2302	0.6059	0.7204	0.5210
n	0.2050	0.6218	0.7311	0.5008
o	0.3015	0.9917	0.9358	0.5818
p	0.2213	0.6034	0.7032	0.5153
q	0.2917	0.6114	0.8500	0.6512
r	0.3405	0.9535	0.9732	0.5712
s	0.2708	0.6708	0.8614	0.6740
t	0.3005	0.9559	0.9645	0.5825

Table 2. Result of the cluster centers

Contrast	Homogeneity	Correlation	Energy	Output
0.2741	0.6314	0.8710	0.6645	0.6103
0.2213	0.6034	0.7032	0.5153	0.5108
0.3418	0.9684	0.9253	0.5822	0.7044
0.3015	0.9854	0.9668	0.5517	0.7014

Table 3. Result of the sigma value

Contrast	Homogeneity	Correlation	Energy	Output
0.0249	0.0689	0.0480	0.0334	0.0351

Table 4. The rules generated from the FIS

Rules	Weights
If (Contrast is in1cluster1) and (Homogeneity is in2cluster1) and (Correlation is in3cluster1) and (Energy is in4cluster1) then (Cancer Type is out1cluster1)	1
If (Contrast is in1cluster2) and (Homogeneity is in2cluster2) and (Correlation is in3cluster2) and (Energy is in4cluster2) then (Cancer Type is out1cluster2)	1
If (Contrast is in1cluster3) and (Homogeneity is in2cluster3) and (Correlation is in3cluster3) and (Energy is in4cluster3) then (Cancer Type is out1cluster3)	1
If (Contrast is in1cluster4) and (Homogeneity is in2cluster4) and (Correlation is in3cluster4) and (Energy is in4cluster4) then (Cancer Type is out1cluster4)	1
If (Contrast is in1cluster5) and (Homogeneity is in2cluster5) and (Correlation is in3cluster5) and (Energy is in4cluster5) then (Cancer Type is out1cluster5)	1
If (Contrast is in1cluster6) and (Homogeneity is in2cluster6) and (Correlation is in3cluster6) and (Energy is in4cluster6) then (Cancer Type is out1cluster6)	1
If (Contrast is in1cluster7) and (Homogeneity is in2cluster7) and (Correlation is in3cluster7) and (Energy is in4cluster7) then (Cancer Type is out1cluster7)	1
If (Contrast is in1cluster8) and (Homogeneity is in2cluster8) and (Correlation is in3cluster8) and (Energy is in4cluster8) then (Cancer Type is out1cluster8)	1

Table 5. The output of the rules used to generate the output of the FIS

Images	Contrast	Homogeneity	Correlation	Energy	Output
a	0.3418	0.9684	0.9253	0.5822	0.7044
b	0.2520	0.6102	0.8632	0.6675	0.6002
c	0.2741	0.6314	0.8710	0.6645	0.6103
d	0.2513	0.6500	0.8512	0.6551	0.6019
e	0.2704	0.6301	0.8711	0.6345	0.6015
f	0.3015	0.9854	0.9668	0.5517	0.7014
g	0.2815	0.6048	0.8519	0.6748	0.6040
h	0.3398	0.9550	0.9550	0.5238	0.7030
i	0.2679	0.6456	0.8602	0.6528	0.6066
j	0.2553	0.6080	0.8718	0.6900	0.6062
k	0.3458	0.9545	0.9420	0.5904	0.7080
l	0.2171	0.6312	0.7015	0.5022	0.5130
m	0.2302	0.6059	0.7204	0.5210	0.5268
n	0.2050	0.6218	0.7311	0.5008	0.5146
o	0.3015	0.9917	0.9358	0.5818	0.7027
p	0.2213	0.6034	0.7032	0.5153	0.5108
q	0.2917	0.6114	0.8500	0.6512	0.6010
r	0.3405	0.9535	0.9732	0.5712	0.7096
s	0.2708	0.6708	0.8614	0.6740	0.6192
t	0.3005	0.9559	0.9645	0.5825	0.6192

In addition, the parameter used for the performance evaluation was detection rate. The performance of the systems detection rate was evaluated using two parameters, True Positive (TP), False Positive (FP). A TP is obtained when a mammogram is correctly detected as mass, stellate lesion and micro calcification as obtained in the miniMIAS database. When a normal mammogram is incorrectly classified, then it is defined as a FP.

The measures are based on the formula:

$$Detection\ Rate = \frac{TP}{TP + FP} \times 100\%$$

The percentage of detection rate of the FIS system is as computed below:

TP=18
FP=2

Hence, the detection rate of the FIS is 90%. However, the high detection rate of cancer indicators in the proposed system could be because the proposed system was tested with a small dataset.

4.1 Clinical Applications and Limitations of the Proposed System

This section highlights the clinical benefits and limitations of the proposed system.

4.1.1 Clinical Applications of the Proposed System

The clinical applications of the proposed system are as highlighted below:

4.1.1.1 Facilitates the extraction of clinically useful information

During the process of diagnosis and treatment, patients routinely undergo some procedures which require diverse techniques. Clinical information is usually generated during these procedures. However, to make sense of this information, it is important to provide reliable computational tools that can automatically analyze and extract information from images. Hence, this system will aid clinicians extract useful information for clinical decision making.

4.1.1.2 Enhancement of Mammographic Images

Mammograms usually contain low signal to noise ratio (low contrast) and a complicated structured

background. Hence, there is a need to enhance mammographic imaging. Thus, this system will aid in the enhancement of mammographic images by making it clearer. There by, increasing the quality of the images.

4.1.1.3 Interpretation of mammographic images

This system will help humans and computers to easily interpret mammographic images once the images have been enhanced.

4.2 Limitations of the Proposed System

The proposed system extracted only four features namely contrast, homogeneity, correlation and energy from the mammographic images. However, images are composed of diverse features such as entropy, auto-correlation, sum of squares and variance, standard deviation, skewness, average and dissimilarity. The purpose of extracting these features is to minimize the classification error. Hence, the more features extracted, the less the classification error. This study is also limited to the detection of three cancer indicators which are mass, microcalcification and stellate lesion.

5. CONCLUSION

Breast cancer is one of the major causes of death among women since the last decades and it has become an emergency for the healthcare systems of both developing and industrialized countries. Mammography is the main test used for screening and early diagnosis of breast cancer. However, the accuracy of mammography cannot be guaranteed. Consequently, this paper focused on the detection of breast cancer's indicators in mammography images using digital image processing techniques. The breast cancer detection system employed image processing techniques such as top hat filtering, Otsu method, Canny and Sobel detection, subtractive clustering and fuzzy logic. The detection system is a two stage detection procedure. The first stage determines the location and border of the suspected areas in the mammogram while the second stage involved the extraction of features from the images and the classification of the images into the three cancer indicators. Hence, this method can be used to assist radiologists by helping them to improve the accuracy of breast cancer diagnosis.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. California Department of Public Health. Breast cancer. California Department of Public Health, Cancer Surveillance Section; 2010.
2. Radiological Society of North America. Mammography. Accessed June 4 2013. Accessed 29 March 2012. Available: http://www. radiologyInfo.org
3. Saslow D, Boetes C, Burke W. American cancer society guidelines for breast screening with MRI as an adjunct to mammography. A Cancer Journal for Clinicians. 2007;57:79-89
4. Mohiy H, Mohamed A, Walid D. Detection of breast cancer tumor algorithm using mathematical morphology and wavelet analysis. GVIP Conference, Cairo, Egypt; 2005.
5. Fei G. An efficient approach to automated segmentation in medical image analysis. Master of Science in electrical engineering, Texas Technical University, USA; 2005.
6. Sheshadri HS, Kandaswamy A. Detection of Breast Cancer by Mammogram Image Segmentation. Department of Electronics & Communication Engineering, PSG College of Technology, Coimbatore-641004 Tamilnadu, India; 2005.
7. Olaleke JO. Development of a Breast Cancer Detection System. MTech Thesis, School of Science, Federal University of Technology, Akure, Ondo State, Nigera; 2012.
8. Seamann T. Digital image processing using local segmentation. Ph.D Thesis, School of Computer Science and Software Engineering, Faculty of Information Technology Monash University, Australia; 2003
9. Fu JC, Lee SK, Wong STC, Yeh JY, Wang AH, Wu HK. Image segmentation feature selection and pattern classification for mammographic micro calcifications. Computerized Medical Imaging and Graphics. 2005;29:419–429.
10. Duarte MA, Alvarenga AV, Azevedo CM, Calas MJG, Infantosi AFC, Pereira WCA. Segmenting mammographic microcalcifications using a semi-automatic

procedure based on Otsu's method and morphological filters. Brazilian Journal of Biomedical Engineering. 2013;29(4):377-388.

11. Mohammed WA, Kadah YM. Computer aided diagnosis of digital images. International Conference on Computer Engineering & Systems. 2009;299-303.
12. Otsu N. A Threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979;9(1):62-66.
13. Jain Anil K. Data Clustering: 50 years beyond K-Means. Department of Computer Science and Engineering, Michigan State University, USA; 2008.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=1136&id=5&aid=9161