An approach for automatic lesion detection in mammograms

K.U. Sheba* and S. Gladston Raj

Abstract: Early stage breast cancer detection can reduce death rates in long term. Mammography is the current standard screening tool available for breast cancer detection, but it is found to have high false-positive and false-negative rates. This may be due to poor quality of mammograms, subtle nature of malignancies and limitations in human/brain visual system. The aim of this research work is to develop an efficient classification tool with improved breast screening accuracy to distinguish between healthy, benign and malignant breast parenchyma in digital mammograms. This paper presents a computer aided diagnosis system for automated detection and diagnosis of breast cancer in digital mammograms. The proposed system can be used as a reference reader for double reading the mammograms and thus assisting the radiologists in clinical diagnosis by indicating suspicious abnormalities. This can improve the diagnostic performance of the radiologists. In the proposed methodology, the regions of interest (ROI) are automatically detected and segmented from mammograms using global thresholding, Otsu’s method and morphological operations. Shape, texture and grey-level features are extracted from the ROIs. Optimal features are selected using Classifier and Regression Tree (CART). Classification is performed with Feed forward artificial neural networks using back propagation. Performance is evaluated using Receiver Operating Characteristic (ROC) analysis and confusion matrix. Experimental results show that the proposed method...
achieved an accuracy of 96% with 83% sensitivity and 98% specificity. The proposed methodology has been compared with several other classification models and is found to have a good performance in terms of accuracy, sensitivity and specificity.

Subjects: CAD CAE CAM - Computing & Information Technology; Computer Engineering; Medical Imaging

Keywords: grey-level co-occurrence matrix; grey level run length matrix; Otsu’s method; morphological open; feed forward artificial neural networks; global thresholding

1. Introduction

Breast cancer is the most common cancer affecting women across the world. It is the second most common cancer in the world and there is a steady increase in breast cancer cases among young women (Siegel, Naishadham, & Jeimal, 2013). Statistics reveal that India has the largest breast cancer mortality in the world (Statistics of Breast Cancer in India, 2013).

Several screening modalities for breast cancer are available like mammography, PET, MRI, Ultrasonography etc. of which mammography is considered to be the most reliable and economical (World Cancer Research Fund International, 2013–2014). Space occupying lesions are most common symptoms of breast cancer in mammograms (American Cancer Society, 2016). Lesions can be of three types-masses, architectural distortion and bilateral asymmetry of the breast (Homer, 2004). Masses can either be benign or malignant depending on their shape, margin and density properties. In terms of shape, masses that are round, oval and slightly lobular shaped are benign. An irregular, multi lobular mass may suggest malignancy. In terms of margin, masses with well-defined or circumscribed margins are considered benign. Masses with spiculated, indistinct or micro lobulated margins are highly suspicious of malignancy. In terms of density, a low density mass is probably benign where as a hard immobile mass is a sign of malignancy. Isodense masses with lobular shape and microlobulated margins are moderately suspicious (Mammographic Mass Characteristic, 2017). Second type of space occupying lesions, architectural distortion, is the tethering or indentation of breast tissue region with radiating spiculations. Architectural distortion belongs to the benign category if it is caused due to post-surgical scars or soft tissue damage. It is highly suspicious of malignancy if it is accompanied with palpable breast mass (Banik, Rangaraj, & Desautals, 2013). The third type of space occupying lesions in mammograms, bilateral asymmetry, reveals an area of high density in one of the breast as compared to the same area in the other. Breast cancer can present itself as an area of focal asymmetry or when in advanced stage can present itself as a new asymmetry (Rangayyan, Ferrari, & Frère, 2007). Hence, it requires detailed evaluation.

Mammography is found to be less effective for small lesions or tumors as it is difficult to detect them using naked eye and also for subjects less than 40 years of age with dense breasts (Wang, 2017). Some of the lesions in mammograms may go undetected or may be diagnosed incorrectly due to poor quality of mammograms, inexperience of the radiologists or due to the limitation in human/brain visual system (Sheba & Gladston Raj, 2016). To overcome the above limitations, computer aided detection and diagnosis systems are used by radiologists for breast cancer detection and diagnosis. Currently, the accuracy of CAD systems is not very high. Improving the accuracy of CAD system can lead to an improvement in detection accuracy and this can result in higher survival rate and treatment options.

The proposed system presents a methodology with improved accuracy for automatic classification of mammograms as normal, benign or malignant. Furthermore, it aids in detecting lesions in abnormal mammograms which may indicate the presence of malignancies.

The paper is organized as follows. Section 2 presents the recent work done in this field. Section 3 explains the proposed methodology. Section 4 discusses the experimental results and Section 5 concludes the paper.
2. Related work
Recent research has laid stress on the development of computer aided detection and diagnosis system for breast cancer.

Wang, Yu, Kang, Zhao, and Qu (2014) in their paper proposed a method for breast tumor detection and classification based on ELM classifier. Here, modified wavelet transformation of the local modulus maxima algorithm is applied for segmentation. Five textural and five morphological features have been extracted from the ROIs for classification purpose. This method is found to have improved training speed and better classification accuracy than SVM. de Lima, da Silva-Filho, and Dos Santos (2016) extracted Zirinke moments from the region of interests of the images. They combined them with texture and shape features. A classification accuracy of 94.11% was obtained. The proposed method however obtained high accuracy when classifying fatty glandular and fatty breasts. A novel weight optimized multi-layer perception (MLP) based classifier with genetic algorithm is proposed for classification of lesions in mammograms by authors Valarmathi and Robinson (2016). Classification accuracy is claimed to have significantly improved by 10.53% as compared to traditional MLP neural networks. Gedik (2015) presents a CAD system which makes use of new contourlet transform algorithm (SFLCT) and least square SVM classifier to classify mammogram images. They have obtained an accuracy of 98.467%. Sensitivity is not mentioned. Hu, Gao, and Li (2011) developed a novel algorithm which utilizes a combination of adaptive global thresholding segmentation and adaptive local thresholding segmentation on multi resolution representation of original mammograms. The algorithm has a sensitivity of 91.3% with 0.71 false positives per image. Only texture features have been used. The detection results may have greatly improved if shape features and grey level intensity features had been included. Surinderan and Vadivel (2012) made use of CART classifier for classifying mammogram masses. Only shape features have been used for classification. They obtained a classification accuracy of 93.62%. The ROIs have been manually extracted. Also shape features are good at classifying one type of lesions especially circumscribed lesions. A new approach for classification of mass is proposed by Minavathi and Dinesh (2012) which make use of active contour method and measurement of the angle of curvature of each pixel at the boundary of the mass for segmentation of ROIs. This method obtains 92.7% sensitivity with 0.88 area under the ROC curve. Talha (2016) in his paper presented a fully computerized classification scheme to identify normal and abnormal mammograms using newly proposed GP filter. It obtained an accuracy of 96.97 with 98.39% sensitivity and 94.59% specificity. But, the scheme does not differentiate between benign and malignant mammograms. The authors Mehdy, Ng, Shair, Saleh, and Gomes (2017) in their paper have discussed the usage of artificial neural networks in breast cancer detection. They have discussed different variations of neural network especially the recent trend of hybrid neural networks like SOM model and have concluded that neural networks when combined with other methods can achieve better accuracy, sensitivity and positive predictive value. Wang, Nishikawa, and Yang (2017) developed a convolutional neural network (CNN) to which the input consisted of large image window for computerized detection of clustered microcalcifications. They conducted the experiment on digital and film mammograms and evaluated the detection performance using receiver characteristic analysis. CNN classifier achieved 0.971 in the area under the ROC curve. Abdel-Zaher and Eldeib (2016) in their paper developed a CAD scheme for detection of breast cancer using deep belief network unsupervised path followed by back propagation supervised path. The technique was tested on Wincosin Breast cancer data-set and was found to have high success rate in breast cancer detection. Wang, Li, and Gao (2014) developed a new classification scheme using LDA model with spatial pyramid extension incorporating spatial and marginal statistical characteristics. This classification scheme was found to be more robust compared to general feature based classification with accuracy of 92.74%
3. The proposed system

The proposed system involves the following phases- Image pre-processing, image segmentation, feature extraction, feature selection and classification. Normal mammograms do not contain lesions. But for the purpose of classification, they are also subjected to pre-processing, segmentation, feature extraction and feature selection prior to classification.

3.1. Image pre-processing

The aim of pre-processing in mammograms is to enhance the breast profile from the background and to remove artifacts, labels and noise that may appear in the mammograms. Also, unrelated parts like pectoral muscles appear in the mammograms. This also needs to be eliminated to help in correct interpretation of images (Sheba & Gladston Raj, 2016). In this work, pre-processing has been carried out in two stages. The first stage consists of noise filtering, artifact and label removal and image enhancement. Median filter (Huang & Zhu, 2012) has been used for noise filtering, global thresholding (Chaubey, 2016) for artifact and label removal and adaptive fuzzy logic based bi-histogram equalization (Sheba & Gladston Raj, 2017) for controlled enhancement. Adaptive fuzzy logic based bi-histogram equalization (Sheba & Gladston Raj, 2017) is an efficient algorithm proposed in our previous paper to improve the quality of the mammograms for better perception. The algorithm combines fuzzy logic with brightness preserving bi-histogram equalization (BBHE). The merit of the proposed method is that it is fully adaptive in nature where all the parameters are computed based on the characteristics of the mammographic images and this aids in providing controlled contrast enhancement. Algorithm 1 describes the process. Figure 1 shows the results. The second stage consists of elimination of pectoral muscles. The pectoral muscles is nearly triangular in shape and appears in the upper left corner or the upper right corner of the breast contour depending on whether it is the left or right breast. In this work, the Bounding Box (Chang, 2006) has been used for the removal of pectoral muscles. Bounding Box of any image contains the coordinates of the rectangular border that fully contains the image. Bounding Box for the breast contour in the mammogram is calculated. The pectoral muscles is contained in the upper left or right corner in one third the width of the bounding box and hence an upper triangle is created with one third the width of the bounding box. All the pixels in the upper triangle are changed to binary zeros. This process has been used for the removal of pectoral muscles. Algorithm 2 describes the process and Figure 2 shows the results. All the algorithms have been implemented using MATLAB 15.0.
Algorithm 1

Aim : To enhance the mammogram image I and to remove noises, labels, artifacts.
Input : A two dimensional mammogram image I.
Output : An enhanced, noise free image with pectoral muscles removed.

Step 1 : Read the mammogram I.
Step 2 : Apply median filter to filter noises
         \[ I_2 = \text{medfilt2}(I) \]
Step 3 : Scale up every pixel by constant
         \[ I_3 = I_2 \times 10. \]
Step 4 : Apply global thresholding to the image. This finds a threshold value and generates
         a binary image with all pixel values below the threshold being converted to 0s and
         those above to binary ones.
         \[ \text{mask} = \text{im2bw}(I_3, \text{graythresh}(I_3)). \]
Step 5 : Fill the tiny holes in the binary mask.
         \[ \text{mask} = \text{imfill}('holes') \]
Step 6 : Select the largest object with binary 1’s in the binary image. The largest object is the
         breast contour.
         \[ \text{mask} = \text{bwpropfilt}('area', I_3, 1). \]
Step 7 : Apply mask to the filtered image// Image is logically ANDed with the largest object
         masked = I_2, masked (~mask)=0.
Step 8 : Apply adaptive fuzzy logic based HE to enhance the final image.
         \[ \text{preprocessed} = \text{adaptfuzzyhisteq}(	ext{masked}). \]
Step 9 : Stop.

Algorithm 2

Aim : To eliminate pectoral muscles.
Input : An enhanced noise free two dimensional mammogram image I.
Output : An image with pectoral muscles removed.

Step 1 : Find the size of the pre-processed image.
         \[ [r, c]= \text{size}(\text{preprocessed}) \] // r and c represents the number of rows and columns
Step 2 : Get the Bounding Box of the breast contour of the pre-processed mammogram. It
         contains the coordinates of the rectangular border that encloses the breast contour.
         Bbox(x, y, width, height).
         Here, Bbox(1) = x- coordinate of the origin of the bounding box
         Bbox(2) = y- coordinate of the origin of the bounding box
         Bbox(3) = width of the Bbox
         Bbox(4) = height of the Bbox
Step 3 : Create a binary mask (mask2) of the region enclosed by the bounding box with each
         pixel within the breast contour being binary ones and the remaining binary zeros.
         mask 2 = ones(r-Bbox(2), Bbox(3)) // r-Bbox(2) and Bbox(3) represents the
         // height and width of mask2
Step 4 : Find the upper triangle of the mask located within one third of the width of the
         bounding box and assign binary zeros to the region in the mask other than the upper
         triangle
         \[ k = \text{Bbox}(3)/3 \] // k represents one third the width of the mask
         mask2 = triu(mask2, k) // triu function calculates the upper triangle and
         // automatically assigns binary 0’s to the surrounding region of the triangle.
Step 5: Create two more binary masks (mask1 and mask3) which are of the same size as the regions to the left and right of the bounding box. As these regions are unwanted regions, not required for pectoral muscle elimination, all the pixels in these masks are changed to binary zeros.

\[
\text{mask 1} = \text{zeros} \left( r-Bbox(2), Bbox(1) \right) \quad // \quad r-Bbox(2) \text{ and } Bbox(1) \text{ represents the height and width of mask1.}
\]

\[
\text{mask 3} = \text{zeros} \left( r-Bbox(2), c-(Bbox(1) + Bbox(3)) \right) \quad // \quad \text{The parameters represent height and width respectively of mask3}
\]

Step 6: Concatenate all masks

\[
\text{mask} = [ \text{mask1, mask2, mask3}] 
\]

Step 7: Apply the mask to the preprocessed image such that the region in the image enclosed by the upper triangle of the bounding box becomes zero keeping the remaining portion of the image intact.

\[
\text{preprocessed (mask)} = 0
\]

Step 8: Stop.

### 3.2. Segmentation of lesions

Suspicious space occupying lesions are automatically segmented from the mammograms for further processing. Lesions tend to be brighter than the surrounding area, therefore they have higher intensity values. Multithresholding based on Otsu’s method (Chen et al., 2012) is applied on the mammograms. Multithreshold values are generated. Based on these threshold values, the image is quantized to create a label matrix. Pixels regions containing highest label values of size varying from 400 to 21,000 pixels are chosen. The contours of the pixel regions are smoothened using morphological operations open and close (Zhang, Ji, Li, & Wu, 2016). Cancerous and non-cancerous lesions are found to have a size ranging from less than 2 cm to greater than 5 cm at its widest points. From experimentation, we have found that all lesions whether cancerous or non-cancerous have an area with pixel regions ranging from 400 to 21,000 pixels in mammograms. Algorithm 3 gives the algorithm for the segmentation of lesions and Figure 3 shows the results.

**Algorithm 3**

**Aim:** To detect suspicious lesions in mammograms.

**Input:** Preprocessed mammogram image.

**Output:** Segmented lesions

Step 1: Read the preprocessed image I.

Step 2: Assign the number of grey levels (NL) to be used in multithresholding (NL) ← 5

Step 3: Multithresholding based on Otsu’s method is applied to the image

\[
\text{thresh} = \text{multithresh} \left( I, NL \right)
\]

Step 4: Breast image is quantized to create a label matrix of the same size containing pixel values ranging from 1 to NL + 1

\[
\text{seg} = \text{imquantize} \left( I, \text{thresh} \right);
\]

Step 5: Choose the pixel regions having values greater than ‘NL’

\[
\text{mask} = \text{seg} > \text{NL}
\]

Step 6: Smoothen the edges of the mask using morphological operations open and close

\[
\text{mask} = \text{imopen} \left( \text{mask}, \text{se} \right); \quad // \quad \text{se is a structuring element of disk shape. se = 5}
\]

\[
\text{mask} = \text{imclose} \left( \text{mask}, \text{se} \right);
\]

Step 7: Select objects having area between 400 to 21,000 pixels.

\[
\text{mask} = \text{bwpropfilt} \left( \text{mask}, \text{‘area’}, [400, 21,000] \right);
\]

Step 8: Apply mask to the input image

\[
I \left( \text{~mask} \right) = 0
\]

Step 9: Stop.

Figures 4–6 show the preprocessing and segmentation of the normal, malignant and benign mammograms using the algorithm 1 and algorithm 2.
Figure 3. (a) Preprocessed image mdb005 with pectoral muscles removed (b) After applying multithresholding based on Otsu’s method and (c) Segmented ROIs.

Figure 4. (a) Benign image mdb013 (b) Its preprocessed image with abnormalities recorded (c) With pectoral muscles removed and (d) Segmented ROIs.

Figure 5. (a) Malignant image mdb028 (b) Its preprocessed image with abnormalities recorded (c) With pectoral muscles removed and (d) Segmented ROIs.

Figure 6. (a) Normal image mdb008 (b) Its preprocessed image (c) With pectoral muscles removed and (d) Segmented ROIs.
3.3. Feature extraction

The three different types of space occupying lesions—masses, architectural distortion of the breast and asymmetric breast tissues can be characterized as benign or malignant based on their shape, texture and grey level intensity values. Hence, it is important to extract texture, shape and grey level features from the segmented ROIs to classify them as normal, benign or malignant. In order to extract grey level intensity features, first order statistical feature analysis method (Nurhayati, Susanto, Thomas, & Maesadji, 2011) has been used. To extract texture features two methods namely, grey level co-occurrence method (Eichkitz, John, Amtmann, Marcellus, & de Paul, 2015) and grey level run length method (Bharathi & Subashini, 2013) has been used. Shape feature analysis method has been used to extract shape features.

3.3.1. First order statistical feature analysis method

First order statistical feature analysis method (Nurhayati et al., 2011) is the simplest of all feature analysis methods. This method makes use of intensity level histogram of the image to compute the grey level intensity features. Six statistical features are extracted from the ROIs—mean, variance, skewness, kurtosis, energy and entropy. These features are useful in measuring the brightness, the contrast and intensity variation of the ROIs. In mammograms, as lesions tend to be brighter and have higher contrast as compared to normal tissues, the statistical features are useful in detecting suspicious lesions.

3.3.2. Grey level co-occurrence matrix (GLCM) method

GLCM method (Eichkitz et al., 2015) is a statistical texture analysis method which characterizes the ROIs based on texture properties. GLCM method derives a large set of second order texture features from the normalized grey-level co-occurrence matrix of the image. 13 texture features are derived from the GLCM matrix each at angles 0°, 45°, 90° and 135° at distance d = 1. Hence, a total of 52 texture features values are obtained. They include—energy, contrast, correlation, variance, homogeneity, entropy, sum average, sum entropy, sum variance, difference variance, difference entropy, first correlation measure and second correlation measure. GLCM features are useful in differentiating smooth, soft texture from hard, rough textures. Benign lesions in mammograms tend to be soft & smooth whereas malignant lesions are hard and immobile. Hence, GLCM features help in distinguishing benign lesions from malignant ones.

3.3.3. Grey-level run length matrix (GLRLM) method

GLRLM method (Bharathi & Subashini, 2013) is a statistical texture analysis method which searches the image for the runs of pixels having same grey level values in a particular direction using the grey level run length matrix(GLRLM) derived from the image. 11 texture features are derived from the GLRLM of the image at angles 0°, 45°, 135° and 90°. Hence a total of 44 GLRLM texture features are obtained. The GLRLM features include short run emphasis (SRE), long run emphasis (LRE), grey level non-uniformity (GLN), run percentage (RP), run length non-uniformity (RLN), low grey level run emphasis (LGRE), high grey level run emphasis (HGRE), short run low grey level emphasis (SRLGE), short run high grey level emphasis (SRHGE), low run low grey level emphasis (LRLGE) and low run high grey level emphasis (LRHGE). GLRLM features play an important role in differentiating benign lesions which have smooth soft texture from malignant lesions which have coarse hard texture.

3.3.4. Shape feature analysis

Shape and margin of lesions are useful in distinguishing them as benign or malignant. 15 shape and margin features (Surinderan & Vadivel, 2012) have been extracted from the ROIs. These features help in measuring the circularity of the lesions, their irregularity and their margin characteristics. Shape and margin features include area, perimeter, eccentricity, equidiameter, compactness, thinness ratio, circularity1, circularity2, elongatedness, dispersion, shape index, Euler number, maximum radius, minimum radius and S.D of the edge.
3.4. Feature selection

Altogether, 117 features are extracted from the ROIs of the mammograms (6 histogram features, 52 GLCM features, 44 GLRLM features and 15 shape features). It has been found that the large set of features do not necessarily lead to high classification accuracy as some of the features may be redundant, irrelevant, noisy or misleading. This can actually have a negative impact on the classification process in terms of classifier efficiency and computational time complexity. Feature selection is inevitable to select an optimal subset of significant features which not only improves classification but also leads to lower data collection. In this paper, classifier and regression tree (CART) (Hayes, Usami, Ross, & John, 2015) has been used for feature selection. CART is a decision tree induction algorithm which constructs a flow chart like structure where each internal node denotes a test on an attribute and each external node denotes a class prediction. At each internal node, the CART algorithm chooses the best feature to partition the data into individual classes using the Gini index. Hence, the features that appear on the decision tree are the relevant features. These features form the reduced subset of attributes. Figures 7–10 shows the decision trees constructed using CART algorithm for selection of optimal subset of histogram features, GLCM features, GLRLM features and shape features. Here, Class 1 = Normal class, Class 2 = Benign class and Class 3 = Malignant class.

In Figure 7, \( x_1, x_2, x_3, x_4, x_5 \) and \( x_6 \) represents features mean, variance, skewness, kurtosis, energy and entropy. Only four features namely, \( x_1, x_2, x_3 \) and \( x_6 \) appear on the decision tree and hence are the relevant features.

In Figure 8, \( x_1-x_4 \) represents contrast in four different angles, \( x_5-x_8 \) represents homogeneity, \( x_9-x_{12} \) represents energy, \( x_{13}-x_{16} \) represents correlation, \( x_{17}-x_{20} \) represents variance, \( x_{21}-x_{24} \) represents entropy, \( x_{25}-x_{28} \) represents sum average, \( x_{29}-x_{32} \) represents sum entropy, \( x_{33}-x_{36} \) represents sum variance, \( x_{37}-x_{40} \) represents difference entropy, \( x_{41}-x_{44} \) represents difference variance, \( x_{45}-x_{48} \) represents correlation1 and \( x_{49}-x_{52} \) represents correlation2 in directions 0°, 45°, 135° and 90° respectively. But, only 9 GLCM features appear on the decision tree namely, \( x_1, x_2, x_3, x_9, x_{17}, x_{25}, x_{37}, x_{39} \) and \( x_{41} \).

In Figure 9, \( x_1-x_4 \) represent SRE in four direction, \( x_5-x_8 \) represents LRE, \( x_9-x_{12} \) represents GLN, \( x_{13}-x_{16} \) represents RLN, \( x_{17}-x_{20} \) represents RP, \( x_{21}-x_{24} \) represents LGRE, \( x_{25}-x_{28} \) represent HGRE, \( x_{29}-x_{32} \) represents SGLGE, \( x_{33}-x_{36} \) represents SRHGE, \( x_{37}-x_{40} \) represents LRLGE and \( x_{41}-x_{44} \) represent...
represents LRHGE in four directions. Out of these, 13 GLRLM features appear on the decision tree and hence are relevant. They are \( x_1, x_2, x_6, x_{15}, x_{17}, x_{21}, x_{22}, x_{24}, x_{31}, x_{337}, x_{38}, x_{39}, \) and \( x_{43} \).

In Figure 10, \( x_1 \) to \( x_{15} \) represents shape features area, perimeter, max radius, min radius, Euler number, eccentricity, equidiameter, elongatedness, circularity1, circularity2, compactness, dispersion, thinness ratio, shape index and standard deviation of edge. There are only 7 relevant features appearing on the decision tree namely, \( x_1, x_2, x_4, x_9, x_{12}, x_{14} \) and \( x_{15} \).

From Figures 7–10, it is clear that 4 out of 6 histogram features, 9 out of total 52 GLCM features, 13 out of 44 GLRLM features and 7 out of 15 shape features appear as internal nodes on the decision trees which means that out of 117 features, there are only 33 features which are relevant and significant features for classification. All the remaining features are eliminated from further consideration.
3.5. Classification
Feed forward artificial neural networks with back propagation (FFANN) (Mehdy et al., 2017) has been used as the classifier for the classification phase. CART has not been used to classify mammograms after feature selection because it has been found by experiment that FFANN has better classification accuracy than CART. FFANN is one of the most popular techniques for classification as it is simple, robust and has fast training speed.

FFANN is a network of three layers- an input layer with number of neurons equal to the number of selected features i.e., 33, a hidden layer and an output layer with three neurons each representing a target class- normal, benign and malignant. The sample data-set is divided into three sets- training set, test set and validation set. Initial weights and bias are randomly selected for FFANN usually between −1.0 to 1.0 and −0.5 to 0.5. To propagate the inputs forward, an activation function is used. The log sigmoid function is used as the activation function.

FFANN processes the data-set of training tuples comparing the network prediction of each tuple with the actual known class label. FFANN learns by using the gradient descent method in the backward direction to iteratively search for a set of weights and bias to minimize the mean square distance between networks class prediction and the known target value of the tuples. In order to avoid over fitting by the FFANN, the validation set is used during the training process. After the necessary accuracy is obtained, the weights are frozen. The test data is then fed to the FFANN and the classification accuracy is measured.

3.6. Performance evaluation
There are several evaluation measures available to evaluate the performance of the classification model. Two such measures, confusion matrix and receiver operating characteristics (ROC) analysis have been used in this paper to evaluate the performance of the classification tool.

3.6.1. Confusion matrix
Confusion matrix (Prabusankarlal, Thirumoorthy, & Manavalan, 2017) is a useful tool for analyzing how accurately the classifier recognizes tuples of different classes. Performance of a classification algorithm can be summarized in the form of a confusion matrix. Figure 11 shows a confusion matrix where the row of the matrix represents the predicted class and the columns represent the actual
class. TP and TN denote the number of positive tuples and number of negative tuples that are classified correctly as positives and negatives respectively. FP and FN represents the tuples that are wrongly misclassified as positives and negatives. Several evaluation metrics like accuracy, sensitivity and specificity can be generated from the confusion matrix.

Accuracy is the proportion of true results - both true positives and true negatives. Accuracy measures how well the classifier predicts all classes correctly. Accuracy results can be misleading if there exists a class imbalance problem where tuples in the data-set are not uniformly distributed in all classes i.e. majority of all tuples belong to one class with only very few belonging to the remaining classes. In such cases, sometimes, it can so happen that even when the accuracy rate is high it may not be acceptable because the classifier may only be correctly labeling tuples belonging to the majority class while misclassifying tuples belonging to the minority class. This is especially true in case of breast cancer. This is because breast cancer is rare and breast cancer data-set consists of tuples mostly belonging to the negative class. In such cases, evaluation metrics like sensitivity and specificity gives better evaluation than the overall accuracy as sensitivity is the proportion of true positives and specificity is the proportion of true negatives correctly identified by the classifier.

### 3.6.2. Receiver operating characteristic (ROC) analysis

ROC curves (Wang et al., 2017) are useful tools for visualizing and evaluating classifiers. ROC graphs are two dimensional graphs in which sensitivity (TP rate) is plotted on Y axis and FP rate (1 - specificity) is plotted on the X axis and it represents the trade-off between the rate at which the model can accurately identify positive tuples versus the rate at which it misclassifies negative cases as positive. For each tuple, a point is plotted in the ROC space by using its TP rate against its FP rate. A ROC curve is plotted using these ROC points. An ROC curve of a good classifier moves steeply from center (0, 0) towards the top left corner and then the curve eases off and becomes more horizontal. The area under the curve is a good measurement of accuracy. Higher the area, higher is the accuracy of the model. A model with higher accuracy will have an area closer to 1.0 and that with less accuracy will have an area closer to 0.5.

### 4. Experimental results

The proposed system was tested using mammograms obtained from mini-MIAS database (Suckling et al., 1994). The database includes 322 breast images of 161 patients which have been carefully selected, expertly diagnosed and positions of abnormalities have been recorded in case of malignant and benign mammograms. Only normal mammograms and those cases from mini-MIAS database containing space occupying lesions are considered in this work. They are selected based on the appropriate description provided for each image. 251 mammograms are selected of which 178 are normal, 44 are with benign lesions and 29 are with malignant space occupying lesions.
All the mammograms are preprocessed and ROIs are detected. Histogram, GLCM, GLRLM and shape features are extracted from the ROIs. 33 optimal features are selected using the decision trees. Normal mammograms do not contain lesions. But for the purpose of classification, they also undergo all the steps mentioned above. For training purpose, a total of 171 mammograms containing 124 normal mammograms, 24 mammograms with benign lesions and 23 mammograms with malignant lesions are used. Hence, a matrix of size $33 \times 171$ is given as input to the input layer which is made up of 33 neurons. Each row of the matrix contains the value of a particular feature for 171 mammograms. Using this feature matrix, FFANN is trained and once the necessary accuracy is obtained, it is saved as an object file. Finally, FFANN containing 33 neurons in the input layer, 50 neurons in the hidden layer and 3 neurons in the output layer is obtained as the most accurate FFANN in the training stage. During the testing phase, for each test mammogram, matrix of size $33 \times 1$ is created where each row represents one of the 33 features for the respective mammogram. The class of the mammogram is then decided by the FFANN based on the training results. Figure 12(a)–(c) provides the FFANN design and demonstrates the classification performance using the ROC curve and all confusion matrix. From Figure 12(b), the all confusion matrix can be interpreted as follows.

1. From 178 normal cases, 177 have been classified as normal and 1 has been misclassified as malignant.
2. From 44 benign cases, 40 have been truly classified as benign and 4 has been misclassified as malignant.
3. Out of 29 malignant cases, 24 have been classified as malignant whereas 1 has been misclassified as normal and 4 as benign.

The overall classification accuracy is 96.0 with 83% sensitivity and specificity of 98%. From Figure 12(c), which is the ROC curve analysis of the classifier, the area under the curve is large which means the performance of the classifier is good. A simple comparison in accuracy of the proposed method with other methods is given in Table 1. The comparison has not been easy as different databases and different number and types of cases have been used by the authors in their work. From the comparison results, it is possible to conclude that the proposed method has achieved high accuracy rate in automated detection of space occupying lesions and classification of mammograms as normal, benign and malignant.
The types of features extracted, the selection of optimal subset of features and the segmentation process greatly contribute to classifier performance. A better classification accuracy rate can be obtained only if the region of interest is segmented in such a manner that it is restricted to contain only tumour and delineates the enormous unwanted region surrounding it. The ground truth data in mini-MIAS database provides the location of the abnormalities. By comparing with the ground truth data, it has been found that the segmentation process which consisted of multithresholding using Otsu’s method, morphological operations open and close and a constraint value for choosing pixel regions yielded good segmentation results. The types of features extracted from the ROIs are also important in enhancing the classifier performance. Besides the histogram features, texture and shape features have also been extracted as texture and shape are two important characteristics that distinguish benign lesions from malignant lesions. Also, the optimal subset of features selected using CART classifier has also contributed to the efficiency of the classifier.

5. Conclusion and future work
The use of computer aided detection and diagnosis system for breast cancer has received widespread acceptance among the radiologists in the recent years. They contribute as second readers in the early detection and diagnosis of breast cancer. The proposed methodology in this paper can be used in the development of CAD system for breast cancer as it achieves high accuracy rate. In our future work, we would like to make the following enhancements.

1. Though accuracy and specificity are high, sensitivity needs to be improved. The experiment needs to be conducted on larger number of malignant mammograms as this can avoid class imbalance problem and lead to higher sensitivity results.
2. Optimal features play an important role in improving the classifier performance. In future, we would like to compare different feature selection methods and choose the best method
3. The technique for pectoral muscles removal has to be improved, as in some cases, lesions occurring very close to the pectoral muscles have been partially removed leading to misclassifications.
4. Segmentation process does not yield good results with masses having obscure margins. This needs to be addressed.

Funding
The authors received no direct funding for this research.

Author details
K.U. Sheba1
E-mail: shebaku72@gmail.com
ORCID ID: http://orcid.org/0000-0002-3232-7118
S. Gladston Raj2
E-mail: gladston@rediffmail.com

1 Department of Computer Applications, BPC College, Piravom, India.
2 Department of Computer Science, Government College, Nedumangad, India.

Citation information
Cite this article as: An approach for automatic lesion detection in mammograms, K.U. Sheba & S. Gladston Raj, Cogent Engineering (2018), 5: 1444320.
References


