International Journal of Computational Intelligence and Informatics, Vol. 4: No. 4, March 2015

A Review on Approaches at Different Stages of Mammogram Processing

K Thangavel
Department of Computer Science
Periyar University
Salem, Tamil Nadu, India
drktvelu@yahoo.com

R Roselin
Department of Computer Science
Sri Sarada College for Women
Salem, Tamil Nadu, India
roselinjothi@gmail.com

R Subash Chandra Boss
Department of Computer Science
Periyar University
Salem, Tamil Nadu, India
rmsubash_18@yahoo.co.in

Abstract- **The aim of this paper is to review existing approaches of processing mammograms at different stages to detect breast cancer at the earliest. Moreover this paper helps to understand the different stages in mammograms and the already existing techniques in that area for further exploration. The review has been done in different stages namely mammogram preprocessing, segmentation, feature extraction, feature selection and classification in the recent years. The results obtained using different techniques are also reported.**

Keywords- Mammogram, preprocessing, segmentation, feature extraction, feature selection, classification

I. INTRODUCTION

A mammogram is a radiograph of the breast tissue. It is an effective non-invasive means of examining the breast, commonly searching for masses and/or microcalcifications. Cancer is not preventable, but early detection leads to a much higher chance of recovery and lowers the mortality rate from this disease. Mammography plays a central part in early detection of breast cancers because it can show changes in the breast up to two years before a patient or physician can feel them.

Digital Mammograms allow manipulation of fine differences in image contrast by means of image processing algorithms. Different Computer Aided Detection Systems (CAD) have been developed for the specific tasks required in breast imaging, diagnosis, and screening. The effectiveness of digital mammography in detection of breast cancer is currently under investigation. A variety of algorithms have been developed by independent investigators for use with digital mammograms. The use of computers in processing and analyzing biomedical images allows more accurate diagnosis by a radiologist. Humans are susceptible in committing errors and their analysis is usually subjective and qualitative. Objective and quantitative analysis facilitated by the application of computers to biomedical image analysis leads to a more accurate diagnostic decision by the physician [1]. This paper illustrates the review of the literature where the computer aided system is used in different stages of mammogram processing in the recent years.

An independent review has been done in the following areas: Mammogram preprocessing, mammogram segmentation, mammogram feature extraction, feature selection and classification.

A. Image Preprocessing

The pre-processing of mammogram image is essential before detection and segmentation of microcalcification. The presence of artifacts and pectoral muscle can disturb the detection of microcalcification and reduce the rate of accuracy in the CAD system. Its inclusion can affect the results of intensity-based image processing methods and needs to be identified and removed before further analysis. These processes are performed in the preprocessing stage.

B. Mammogram Segmentation

Image segmentation is one of the most critical tasks in automatic image analysis. Segmentation consists of subdividing an image into its constituent part and extracting those of interest. Many techniques for global thresholding have been developed over the years to segment images and recognize patterns but the error on the segmentation leads to misclassification.

C. Feature Extraction and Selection

Since the classification algorithm requires the classified data to be composed of feature vectors, the transformation of image into features is essential. Feature extraction is the process of creating a representation for, or a transformation from the original data. Feature selection aims to determine a minimal feature subset from a problem domain while retaining a suitably high accuracy in representing the original features.

D. Classification

Classification is assigning the objects in the dataset into a predefined set of classes. It is a type of supervised learning, because the set of classes are introduced to the system before executing classification algorithm. Classification of objects in a dataset is very useful both to understand the characteristics of existing objects and to predict the behaviors of new objects.

II. Review on Mammogram Preprocessing

Digital mammogram has become the most effective technique for early breast cancer detection. It takes an electronic image of the breast and stores it directly in a computer. Mammograms are medical images that are difficult to interpret, thus a pre-processing phase is needed in order to improve the image quality and make the segmentation result more accurate. The first step involves the removal of artifact and unwanted parts in the background of the mammogram. Artifacts reduce the quality of mammograms and may mimic or obscure abnormalities and cause interpretation errors. Recognizing artifacts improves the quality of mammographic interpretation and prevents the characterization of artifacts as breast disease. Then, an enhancement process is applied to the digital mammogram. Image enhancement operations can be used to improve the appearance of images, to eliminate noise or error, or to accentuate certain features in an image.

The overall structure of a mammogram is given in Figure 1.

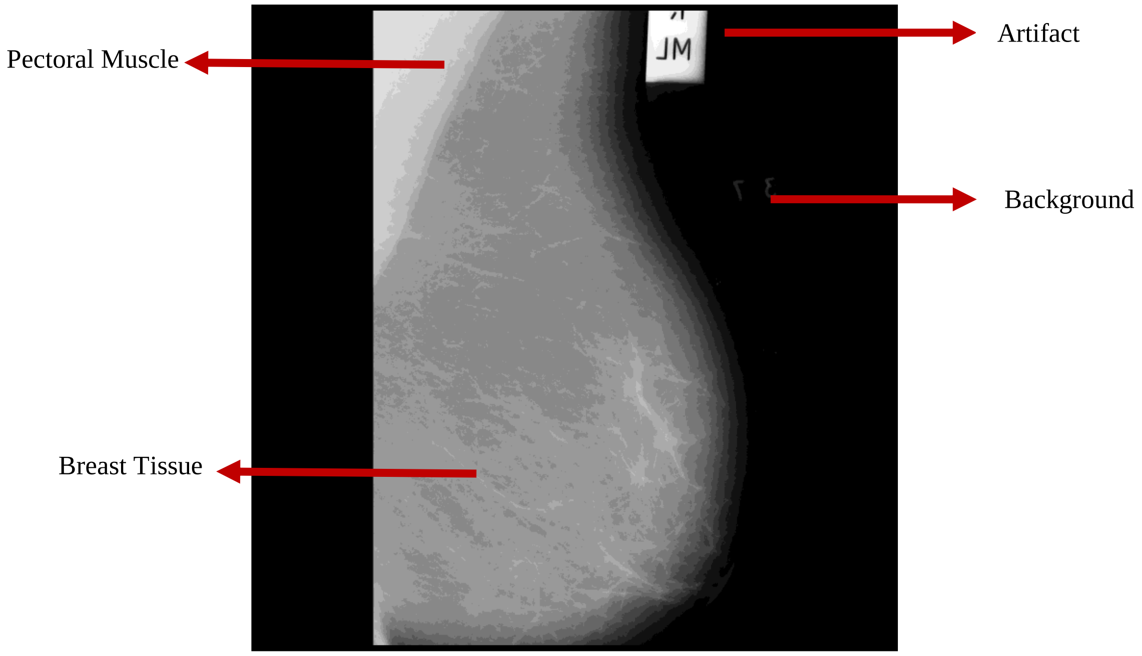

Figure 1 : Structure of a Mammogram

A. Artifact

The artifacts can be a single dead pixels, groups of dead pixels, dead or unread lines, or ghosting. Machine-related artifacts are created by components in the imaging chain that are not directly related to the detector. Patient-related artifacts may be caused by motion or by superimposition of objects or substances over the breast parenchyma or by substances on the skin.

B. Pectoral Muscle

The pectoral muscle is a high-intensity, approximately triangular region across the upper posterior margin of the image, appearing in all Medio Lateral Oblique (MLO) mammograms. Screening mammography typically involves taking two views of the breast

- Cranio-Caudal (CC) view is taken from above a horizontally-compressed breast
- Mediolateral-oblique (MLO) is taken from the side and at an angle of a diagonally-compressed breast

In 2007 Fei Ma et al. [2] proposed two image segmentation methods based on graph theory are used in conjunction with active contours to segment the pectoral muscle in screening mammograms. One method is based on adaptive pyramids (AP) and the other is based on minimum spanning trees (MST). These algorithms are tested on a public data set of mammograms and results are compared with previously reported methods. It is reported that in 80% of the images, the boundary of the segmented regions has average error less than 2 mm. In 82 of 84 images, the boundary of the pectoral muscle found by the AP algorithm has average error less than 5 mm.

Two preprocessing algorithms, one for the breast contour extraction and the other for pectoral muscle segmentation are proposed in 2007 by Mirzaalian H. et al. [3] in which breast region extraction consist of Histogram equalization, Convolving with mask, thresholding and labeling; Modifying ends of breast border, Non-Linear Diffusion. Evaluation of the breast contour detected in the mammograms was performed by the Hausdorff Distance Measure (HDM) and also the Mean of Absolute Error Distance Measure (MAEDM) based on a distance transform and image algebra between the edges identified by radiologists and by the proposed method. The comparison of error between the proposed method and Ferrari's method (based on active contour) and Wirth's method (based on polynomial modeling of the background) shows that the proposed method outperforms two other methods.

The problem of detecting cancer masses by the application of simple thresholding followed by connected component labeling and an algorithm to remove artifacts in digital mammograms using morphological open operation followed by reconstruction are addressed by Subashini T.S. in 2010 [4]. Further the pectoral muscle was removed successfully using simple thresholding and raster scan methods.

Samir Kumar et al. in 2010 [5] proposed a method to detect the pectoral muscle. A straight line (AB) is plotted between the left background of the mammogram and starting of actual part of breast image. The second step is to determine middle point (C) at the top margin of the mammogram and plot a straight line (CD) from the middle point (C) to lower left corner point of the mammogram. The line CD crosses the line AB at point E resulting in an inverted right angle triangle (ACE) that is the region of interest (ROI) to detect the pectoral tissues from mammogram.

Jwad Nagi et al. [6] proposed algorithm that uses morphological preprocessing and seeded region growing (SRG) algorithm in 2011 in order to: (1) remove digitization noises, (2) suppress radiopaque artifacts, (3) separate background region from the breast profile region, and (4) remove the pectoral muscle, for accentuating the breast profile region. To demonstrate the capability of their proposed approach, digital mammograms from two separate sources are tested using Ground Truth (GT) images for evaluation of performance characteristics. Experimental results obtained indicate that the breast regions extracted accurately correspond to the respective GT images.

Sara et al. 2011 [7] proposed a method in which the first phase is omitting the excessive image parts which are in the two sides of the image using pixels brightness. The second phase is the distinction of the breast direction and put all images in one direction by using the threshold limit of gray level of the two halves of the image. The third phase is the breast region segmentation from the background by using the series of point operations and the growing region method and the result has been reported to 99%.

In [8] IndraKanta Maitra et al. 2011 proposed three distinct steps The initial step involves contrast enhancement by using the contrast limited adaptive histogram equalization (CLAHE) technique. Then define the rectangle to isolate the pectoral muscle from the region of interest (ROI) and finally suppress the pectoral muscle using modified seeded region growing (SRG) algorithm. The proposed algorithms were extensively applied on all the 322 mammogram images in MIAS database resulting in complete pectoral muscle suppression in most of the images. This proposed algorithm is compared with other segmentation methods showing superior results.

A new method of finding boundaries of mammogram images has been proposed by Sumit Chopra et al. 2011 [9]. This is named as extended Prewitt operator. The traditional techniques of finding edges in an image don't work well in case of mammogram images. The results are compared with the existing edge detection techniques such as Sobel, Prewitt, Laplacian of Gaussian and Canny edge detector. They evaluated that how the existing technique is better than the already present techniques. In the given technique, selecting an appropriate threshold value is the key. If the proper threshold value is not selected then it will lead to missing of the pixels which are true if the threshold value selected is high and if the threshold value is low then it will leads to the resultant image containing the false pixels.

In 2012 [10] Indra Kanta Maitra proposed three distinct steps for mammogram preprocessing. The initial step involves contrast enhancement by using the contrast limited adaptive histogram equalization (CLAHE) technique. Then define the rectangle to isolate the pectoral muscle from the region of interest (ROI) and finally suppress the pectoral muscle using the proposed modified seeded region growing (SRG) algorithm. The proposed algorithms were extensively applied on all the 322 mammogram images in MIAS database resulting in complete pectoral muscle suppression in most of the images. The proposed algorithm method is also compared with other segmentation methods showing superior results in comparison.

Frahan et al., 2013 [11] extracted the pectoral muscle by using active contours and stopping algorithm which obtains the contour which contains the boundary of the pectoral muscle. Later, it extracts the pectoral muscle binary image from the contour. The proposed algorithm was tested on the mammograms from the mini-MIAS database and it worked very efficiently. It provided very effective and accurate results for pectoral muscle segmentation. It provided up to 97.84% accuracy, computed from well segmented results.

Meenakshi Sundaram et. al., 2014 [12] made a comparative study on different types of filters to remove the noise and they concluded that adaptive median filter is best for mammogram image noise removal and gives better performance by estimating the PSNR values.

Thus in mammogram preprocessing, the methods based on graph theory, histogram equalization, convolving with mask, thresholding and connected component labeling, morphological preprocessing, seeded region growing and active contours are used in the literature to segment pectoral muscle and artifact from the mammograms

III. Review on Mammogram Segmentation

Accurate segmentation of the breast from digital mammograms is an important pre-processing step for computerized breast cancer detection. Up till now, a number of different approaches have been applied to the detection of microcalcifications. Concerning image segmentation and specification of Region of Interest (ROI), several methods have been given in the literature.

An enhancement algorithm that improves image contrast based on local statistical measures of the mammograms has been proposed by Alfonso Rojas Dom´ınguez et al. 2008[13]. After enhancement, regions are segmented via thresholding at multiple levels, and a set of features is computed from each of the segmented regions. A region-ranking system is also presented that identifies the regions most likely to represent abnormalities based on the features computed. The method was tested on 57 mammographic images of masses from the Mini-MIAS database, and achieved a sensitivity of 80% at 2.3 false-positives per image (average of 0.32 false-positives per image).

Alfonso Rojas Dom´ınguez et al. 2009 [14] explored the use of characterization features extracted based on breast-mass contours obtained by automated segmentation methods, for the classification of masses in mammograms according to their diagnosis. Two sets of mass contours were obtained via two segmentation methods a dynamic-programming-based method and a constrained region-growing method. Three popular classifiers (Bayesian classifier, Fisher's linear discriminant, and a support vector machine) were then used to predict the diagnosis of a set of 349 masses based on each of said features and some combinations of these. The systems (each system consists of a segmentation method, a feature set, and a classifier) were compared with each other in terms of their performance on the diagnosis of the set of breast masses. It was found that, although there was a percent difference of about 14% in the average segmentation quality between methods, this was translated into an average percent difference of only 4% in the classification performance. It was also observed that the spiculation feature based on edge signature information was distinctly better than the rest of the features, although it is not very robust to changes in the quality of the segmentation. All systems were more efficient in predicting the diagnosis of benign masses than that of the malignant masses, resulting in low sensitivity and high specificity values (e.g. 0.6 and 0.8, respectively).

In [15] Aminah Abdul Malek et al. 2010 proposed a method by combining seed based region growing and boundary segmentation in sequential order. The first process in region growing is to identify an initial seed point. Most of region growing techniques identify the seed point manually which engage human interaction. Thus, automated initial seed point identification for region growing algorithm is proposed by them. The boundary segmentation technique is implemented in order to improve the segmentation results. The method is tested on 50 mammogram images which contain microcalcifications confirmed by a radiologist. Experimental results show that the algorithm successfully segment the microcalcifications with accuracy of 0.94.

R.B. Dubey et al. 2010 [16] compared two methods for segmentation of masses in digital mammograms viz., level set and marker controlled watershed methods. They used 17 mammogram images for experimental purposes and concluded that the marker controlled watershed segmentation shows better results than the level set approach.

The segmentation methodology presented by Wenda He et al. 2011 [17] consists of five distinct steps: (1) feature extraction using mammographic patches, (2) deriving local image properties, (3) feature transformation, (4) mammographic building block based model generation by clustering, and (5) model driven segmentation. The Mammographic Image Analysis Society database was used to ease the quantitative and qualitative evaluation, with respect to mammographic risk assessment, based on both Tabár and Breast Imaging Reporting and Data System schemes. Classification accuracies of 71% and 79% were achieved in the corresponding low and high risk categories for Tabár and Breast Imaging Reporting and Data System schemes, respectively. They reported that the proposed segmentation approach can produce consistent and realistic segmentation results, with respect to breast anatomy and Tabár tissue modelling. For screening mammography and computer aided diagnosis, this approach is very useful in aiding radiologists' estimation of breast cancer risk and treatment planning prior to biopsies.

Quintanilla-Dominguez et al. 2011[18] used K-Means, Fuzzy C-Means and Possibilistic Fuzzy C-Means to segment mammogram images. A comparison of the advantages and drawbacks offered by these algorithms in mammograms is given and which should help to improve the detection of microcalcification clusters in digitized mammograms.

In [19] Sundaram et. al. 2011, proposed the Histogram Modified Local Contrast Enhancement (HM-LCE) to adjust the level of contrast enhancement, which in turn gives the resultant image a strong contrast and also brings the local details present in the original image for more relevant interpretation. It incorporates a two stage processing both histogram modifications as an optimization technique and a local contrast enhancement technique. This method is tested for Mias mammogram images. The performance of this method is determined using three parameters like Enhancement Measure (EME), Absolute Mean Brightness Error (AMBE) and Discrete Entropy (H) for all 22 numbers of Mias mammogram images with microcalcification. Its enhancement potential is also tested by sobel and otsu methods for the detection of microcalcification in the mammogram image. From the subjective and quantitative measures it is interesting that this proposed technique provides optimum results by giving better contrast enhancement and preserving the local information of the original mammogram images in the Mias data base and the method has increased the delectability of micro calcifications present in the given mammogram image.

A fully automated segmentation method was proposed by Pelin Kus et. al. 2012 [20]. In their experiment, noise on the acquired mammogram is reduced by median filtering; multidirectional scanning is then applied to the resultant image using a moving window 15X1 in size. The border pixels are detected using the intensity value and maximum gradient value of the window. The breast boundary is identified from the detected pixels filtered using an averaging filter. The segmentation accuracy on a dataset of 84 mammograms from the MIAS database is 99%.

Wei-Yen Hsu et al. 2012 [21] used improved watershed transform using prior information, and also done region segmentation and region compression. He showed that the proposed method gives promising results in the compression applications.

Thangavel et al. 2013 [22] proposed mammogram image segmentation using Rough K-Means clustering algorithm. In which the median filter is used for pre-processing of image and it is normally used to reduce noise in an image. The 14 Haralick features are extracted from mammogram image using Gray Level Cooccurrence Matrix (GLCM) for different angles. The features are clustered by K-Means, Fuzzy C-Means (FCM) and Rough K-Means algorithms to segment the region of interests for classification. The result of the segmentation algorithms compared and analyzed using Mean Square Error (MSE) and Root Means Square Error (RMSE). It is observed that the proposed method produces better results that the existing methods.

Pereira DC et al. [23] proposed a method for detection and segmentation of masses using multiple thresholding, wavelet transform and genetic algorithm in mammograms which were randomly selected from the Digital Database for Screening Mammography (DDSM). The developed computer method was quantitatively evaluated using the area overlap metric (AOM). The mean ± standard deviation value of AOM for the proposed method was 79.2 ± 8%. The experiments demonstrate that the proposed method has a strong potential to be used as the basis for mammogram mass segmentation in CC and MLO views. Another important aspect is that the method overcomes the limitation of analyzing only CC and MLO views.

Kanimozhi Suguna et. al. [24] proposed Monkey Search Optimization (MSO) which is based on Metaheuristic Algorithm. It is used for selecting region of interest in mammogram image. Monkey Search Optimization (MSO) algorithm is considered as a new algorithm for searching optimum solution based on the foraging behavior of monkeys. Pectoral region removed image is given as input for feature extraction. The proposed algorithm can be implemented for various applications as the time consumption for the process is reduced greatly. The proposed algorithm is compared with few other metaheuristics algorithms such as Ant Colony Optimization (ACO), Artificial Bee Colony Optimization (ABC) and Particle Swarm Optimization (PSO); from the results it shows that the proposed approach can be considered to be an appropriate algorithm for image segmentation. Results are presented based on simulation made with the implementation in MATLAB which is tested on the images of MIAS database.

The overall results arrived in the literature are listed in Table 1.

TABLE I. REVIEW ON MAMMOGRAM IMAGE SEGMENTATION

Year	Author(s)	Technique and Result
2008	Alfonso Rojas Dom´ınguez and Asoke K. Nandi	Enhancement algorithm that improves image contrast based on local statistical measures. A region-ranking system 57 mammographic images sensitivity of 80% at 2.3 false-positives per image
2009	Alfonso RojasDomínguez and Asoke K.Nandi	Dynamic-programming based boundary tracing method and a constrained region-growing method 349 Masses Resulting in low sensitivity and high specificity values (e.g. 0.6 and 0.8, respectively)
2010	Aminah Abdul Malek, Wan Eny Zarina Wan Abdul Rahmana, Arsmah Ibrahima, Rozi Mahmudb, Siti Salmah Yasirana and Abdul Kadir Jumaata	Combining seed based region growing and Boundary segmentation in sequential order 50 mammogram images Accuracy of 0.94

Year	Author(s)	Technique and Result
2010	R.B. Dubeya, M. Hanmandlub and S.K. Guptac	Level set and Marker controlled watershed methods 17 mammogram images The marker controlled watershed segmentation shows better results than the level set approach
2011	Wenda He, Erika R.E. Denton, Kirsten Stafford and Reyer Zwiggelaar	Based on geometric moments, and prior information of the mammographic building blocks as described by Tabár tissue modeling schemes. Classification accuracies of 71% to 79%
2011	J. Quintanilla-Dominguez, B. Ojeda-Magañaa, M.G. Cortina-Januchs, R. Ruelas, A. Vega-Coronab and D. Andinaa	K-Means, Fuzzy C-Means and Possibilistic Fuzzy C-Means A comparison of the advantages and drawbacks offered by these algorithms in mammograms is given.
2011	M. Sundarama, K. Ramarb, N. Arumugama and G. Prabina	The Histogram Modified Local Contrast Enhancement. The performance of this method is determined using three parameters like Enhancement Measure (EME), Absolute Mean Brightness Error (AMBE) and Discrete Entropy (H). 22 numbers of Mias mammogram images with microcalcification in the Mammogram Image.
2012	Pelin Kus and IrfanKaragoz	Noise on the acquired mammogram is reduced by median filtering multidirectional scanning and it is then applied to the resultant image using a moving window15X1 in size. The border pixels are detected using the intensity value and maximum gradient value of the window. The breast boundary is identified from the detected pixels filtered using an averaging filters. 84 mammograms from the MIAS. Segmentation accuracy 99%.
2013	Subash Chandra Boss R., Thangavel K., Arul Pon Daniel K	Median filter is used for preprocessing 14 Haralick features are extracted Segmented using K-Means, Fuzzy C Means and Rough K-Means Results are compared and analysed using Root Square and root means square method. Rough K-Means outperforms others
2014	Pereira DC, Ramos RP, do Nascimento MZ	Multiple thresholding, wavelet transform and genetic algorithm The developed computer method was quantitatively evaluated using the area overlap metric (AOM). The mean ± standard deviation value of AOM for the proposed method was 79.2 ± 8%.
2014	Kanimozhi Suguna S. and Uma Maheswari S	Monkey Search Optimization 95.6 % the region segmented by the Monkey Search Optimization (MSO), contains the actual mass

Thus in the literature, Enhancement algorithm, Dynamic-programming based boundary tracing method, a constrained region-growing method, combining seed based region growing and boundary segmentation in sequential order, Level set and Marker controlled watershed methods, Tabár tissue modeling schemes, K-Means, Fuzzy C-Means, Possibilistic Fuzzy C-Means, The histogram modified local contrast enhancement, rough K-Means, multiple thresholding, wavelet transform, genetic algorithm and Monkey Search Optimization are used for mammogram segmentation.

IV. Review on Feature Extraction and Feature Selection on Mammogram Processing

A fundamental problem of automating the detection and recognition of abnormalities in digital mammograms utilizing computational statistics is one of the methods for extracting the appropriate features for use in a classification system. Several feature extraction techniques have been proposed although none have been shown to be sufficient for the problem. Many of these features tend to be local in nature. Feature selection in the design of pattern classifiers has three goals:

i. To reduce the cost of extracting features,

ii. To improve the classification accuracy, and

iii. To improve the reliability of the estimate of performance.

The important highlights of the review of the feature extraction and feature selection in the literature are given:

Roman W. Swiniarski et. al. 2003 [25] used Histogram based Feature Extraction, Principal Component Analysis (PCA) for dimensionality Reduction, Rough set for Feature Selection and Back propagation network for classification and obtained 75.0% accuracy for the test set.

In [26] Ping Zhang et. al. 2005 extracted 14 Statistical Features; Neural-genetic algorithm is used for feature selection.

J.C. Fu et. al. 2005 [27] extracted features for each suspected microcalcification, representing texture, the spatial domain and the spectral domain. A sequential forward search (SFS) algorithm selects the classification input vector, which consists of features sensitive only to microcalcifications. Classification performance is compared using Az value of the ROC curve. The reported accuracy before feature selection is SVM - 97.01% GRNN - 96 and after feature selection SVM - 98% GRNN - 97.80%

In [28] Dar-Ren Chen et. al. 2005 used the fractal dimension to quantify the texture information. 110 malignant and 140 benign tumors are tested K-Means classification method is used to classify benign tumors from malignant ones. The ROC area index Az is 0.9218

Gobert N. et. al. 2006 [29] extracted image energy features, Intensity gradient features and Co-occurrence matrix features. Empirical distributions are used to estimate statistical significance of classification scores. In an example study, eleven high classification scores are obtained but only three are found to be significant at $p = 0.05$.

In [30] H.S. Sheshadri et. al. 2007 extracted six textural features for mammogram images. They are Mean, Standard deviation, Smoothness, Third moment, Uniformity, Entropy. They used 320 Mammograms. The classification is based on the standard parameters of image histograms as prescribed by ACRBIRADS.

Harris Georgiou et. al. 2007 [31] used Fractal features. Uniresolution curve feature values extracted from the radial distance signal, the DFT spectrum signal and the DWT multi-scale decomposed signals are used. MANOVA significance analysis was carried out and reduced subsets of sizes up to 20 features. LDA model, LSMD classifier were used. The result shows that utilizing the shape characterization alone as discrimination measure between benignancy and malignancy can establish a success rate over 93%.

In [32] Brijesh Verma et. al. 2007 extracted 14 Features ,10 Existing Features in the Literature, which are histogram, average grey level, energy, entropy, number of pixels, standard deviation, skew, average boundary grey level, difference and contrast. Four of them are modified, which are modified energy, modified entropy, modified standard deviation and modified skew. Neural-genetic algorithm is developed for feature selection. The highest classification rate achieved for testing set by Neural Network is 85.0%. Five features are considered to be the most significant features of a digital mammogram for microcalcification classification. They are modified skew, boundary average grey level, standard deviation, skew and modified standard deviation.

In [33] Alfonso Rojas-Domínguez et. al. 2009 presented four new features for the analysis of breast masses. Two of the features measure the degree of speculation of a mass and its likelihood of being speculated. The other two features measure the local fuzziness of the mass margins based on points defined automatically. 319 masses are experimented. The first classifier is a Bayesian classifier, the second classifier is the Fisher's linear discriminant, and the third classifier is a support vector machine. Classification Accuracy reported is 89%.

In [34] Wagner Coelho et. al. 2010, extracted Morphometric features from Convex Polygon and the Normalized Radial Length techniques. From the seven investigated features, both MIFS-U and LDA revealed the normalized residual mean square value and the circularity as the most relevant. An important addition of MIFS-U was the identification of the Mshape as the third in importance, while LDA set this feature at the fifth position. This may be attributed to the MIFS-U capacity, incontrast to LDA, to perceive non-linear behaviours.

Sung-Nien Yu et. al. 2010 [35] extracted Textural features based on Markov random field (MRF) and fractal models together with statistical textural features based on the surrounding region-dependence method (SRDM) from the neighbourhood of the suspicious MCs and Classified by a three-layer BPNN. Twenty mammograms containing 25 areas of MCs are used in this experiment. The free-response operating characteristic (FROC) curve was used to evaluate the performance of classification. A true positive rate of about 94% is achieved at the rate of 1.0 false positive per image.

In [36] Ioan Buciu et. al. 2011 Gabor wavelets and directional features are extracted at different orientation and frequencies. Principal Component Analysis is employed to reduce the dimension of filtered and unfiltered high-dimensional data. Results outperform the radiologist sensitivity reported as being 75%. For the normal versus tumor case, though the specificity is relatively low, a promising value for the sensitivity is achieved.

In [37] Mohamed Meselhy Eltoukhy et. al. 2012, constructed a matrix by putting wavelet or curvelet coefficients of each image in row vector. A feature extraction method is developed based on the statistical t-test method. The method is ranking the features according to its capability to differentiate the classes. Then a dynamic threshold is applied to optimize the number of features, which can achieve the maximum classification accuracy rate. SVM is used for classification and the accuracy is 94.79%.

In [38] Rodrigo Pereira Ramos et. al. 2012, used Features derived from co-occurrence matrices, wavelet and ridgelet transforms of mammographic images are used. The data set consisted of 120 cranio-caudal mammograms, half containing a mass, rated as abnormal images, and half with no lesions. To select the best set of features, genetic algorithm (GA) is used. Experimental results showed that the best classification rates were obtained with the wavelet based feature extraction using GA for selection of the most relevant features, giving an Az = 0.90.

Belal K. Elfarra et. al. [39] used both human features, which are obtained by Digital Database for Screening Mammography (DDSM), and computational features, which are extracted using new feature extraction method called Square Centroid Lines Gray Level Distribution Method (SCLGM). The experimental results are obtained from a data set of 410 images taken from DDSM for different types. 31 features are selected from 145 extracted features; 18 of the selected features are from the proposed feature extraction method (SCLGM). Receiver Operating Characteristics (ROC) and confusing matrix are used to measure the performance. In training stage, the proposed method achieved an overall classification accuracy of 96.3%, with 92.9% sensitivity and 94.3% specificity. In testing stage, the proposed method achieved an overall classification accuracy of 89%, with 88.6% sensitivity and 83.3% specificity

In [40] Yi-Jhe Huang et. al. proposed fully automated algorithm that is able to select a discriminative feature set from a training database via sequential forward selection (SFS), sequential backward selection (SBS), and F-score methods. This feature sets are used to microcalcifications cluster (MCC) detection in digital mammograms for early breast cancer detection. The system was able to select features fully automatically, regardless of the input training mammograms used. The proposed scheme was tested using a database of 111 clinical mammograms containing 1,050 microcalcifications (MCs). The accuracy of the system was examined via a free response receiver operating characteristic (fROC) curve of the test dataset. The system performance for MC identifications was Az = 0.9897, the sensitivity was 92%, and 0.65 false positives (FPs) were generated per image for MCC detection.

Shobha Jose et. al. [41] proposed texture feature extraction of mammogram images based on Biorthogonal wavelet filter via lifting scheme. Maximum likelihood estimator (MLE) is used for texture feature estimation. DDSM is used as the database. Here biorthogonal wavelets are used in the lifting scheme to get texture feature vectors of mammogram images. By using lifting scheme in all biorthogonal wavelets, predict and update filter coefficients are also got. These coefficients are adapted later and thus found the optimal wavelet filter bank for increasing the retrieval performance of the retrieval system. By using lifting scheme methodology decomposition of images are done and thus got approximation and detail coefficients of image.

The Results are summarized in Table 2.

TABLE II. REVIEW ON FEATURE EXTRACTION AND FEATURE SELECTION

Year	Author(s)	Technique and Result
2003	Roman W., Swiniarski and Andrzej Skowron	Histogram based Feature Extraction Principal Component Analysis (PCA) for dimensionality Reduction Rough set for Feature Selection Backpropagation network has provided 75.0% accuracy for the test set
2005	Ping Zhang, Brijesh Verma and Kuldeep Kumar	14 Statistical Features are extracted Neural-genetic algorithm is used for feature selection
2005	J.C. Fu, S.K. Lee, S.T.C. Wong, J.Y. Yeh, A.H. Wang and H.K. Wu	61 features are extracted for each suspected microcalcification, representing texture, the spatial domain and the spectral domain. A sequential forward search (SFS) algorithm selects the classification input vector, which consists of features sensitive only to microcalcifications. Classification performance is compared using Az value of the Receiver Operating Characteristic curve. Before Feature Selection SVM - 97.01% GRNN - 96 After Feature Selection SVM - 98% GRNN - 97.80%
2005	Dar-Ren Chen, Ruey-Feng Chang, Chii-Jen Chen, Ming-Feng Ho, Shou-Jen Kuo, Shou-Tung Chen, Shin-Jer Hung and Woo Kyung Moon	The fractal dimension is used to quantify the texture information. 110 malignant and 140 benign tumors are tested K-Means classification method is used to classify benign tumors from malignant ones. The ROC area index Az is 0.9218
2006	Gobert N. Lee and Murk J. Bottema	Image energy features Intensity gradient features Co-occurrence matrix features Empirical distributions are used to estimate statistical significance of classification scores. In an example study, eleven high classification scores are obtained but only three are found to be significant at p = 0.05.
2007	H.S. Sheshadri and A. Kandaswamy	Six textural features for mammogram images are defined. They are Mean, Standard deviation, Smoothness, Third moment, Uniformity, Entropy.

Year	Author(s)	Technique and Result
		320 Mammograms The classification is based on the standard parameters of image histograms as prescribed by ACRBIRADS.
2007	Harris Georgiou , Michael Mavroforakis, Nikos Dimitropoulos, Dionisis Cavouras and Sergios Theodoridis	Fractal features. Uniresolution curve feature values extracted from (a) the radial distance signal, (b) the DFT spectrum signal and (c) the DWT multi-scale decomposed signals. MANOVA significance analysis was carried out and reduced subsets of sizes up to 20 features. LDA model, LSMD classifier The result shows that utilizing the shape characterization alone as discrimination measure between benignancy and malignancy can establish a success rate over 93%.
2007	Brijesh Verma and Ping Zhang	14 Features are extracted 10 Existing Features in the Literature Four of them are modified, which are modified energy, modified entropy, modified standard deviation and modified skew. Neural-genetic algorithm is developed for feature selection. The highest classification rate achieved for testing set by Neural Network is 85.0%. Five features are considered to be the most significant features of a digital mammogram for microcalcification classification. They are modified skew, boundary average grey level, standard deviation, skew and modified standard deviation.
2009	Alfonso Rojas-Domínguez and Asoke K.Nandi	Four new features for the analysis of breast masses are presented. Two of the features measure the degree of speculation of a mass and its likelihood of being speculated. The other two features measure the local fuzziness of the mass margins based on points defined automatically. 319 masses The first classifier is a Bayesian classifier, the second classifier is the Fisher's linear discriminant, and the third classifier is a support vector machine. Classification Accuracy 89%.
2010	Wagner Coelho A. Pereira , Andre´ V. Alvarenga , Antonio Fernando C.Infantosi, Leonardo Macrini and Carlos E. Pedreira	Morphometric features were extracted from Convex Polygon and the Normalized Radial Length techniques. From the seven investigated features, both MIFS-U and LDA revealed the normalized residual mean square value and the circularity as the most relevant. An important addition of MIFS-U was the identification of the Mshape as the third in importance, while LDA set this feature at the fifth position. This may be attributed to the MIFS-U capacity, incontrast to LDA, to perceive non-linear behaviours.
2010	Sung-Nien Yu and Yu-Kun Huang	Textural features based on Markov random field (MRF) and fractal models together with statistical textural features based on the surrounding region-dependence method (SRDM) were extracted from the neighbourhood of the suspicious MCs. Classified by a three-layer BPNN. Twenty mammograms containing 25 areas of MCs are used in this experiment. The free-response operating characteristic (FROC) curve was used to evaluate the performance of classification. A true positive rate of about 94% is achieved at the rate of 1.0 false positive per image.
2011	Ioan Buciu and Alexandru Gacsadi	Gabor wavelets and directional features are extracted at different orientation and frequencies. Principal Component Analysis is employed to reduce the dimension of filtered and unfiltered high-dimensional data. Results outperform the radiologist sensitivity reported as being 75%. For the normal versus tumor case, though the specificity is relatively low, a promising value for the sensitivity is achieved.
2012	Mohamed Meselhy Eltoukhy, Ibrahima Faye and Brahim Belhaouari Samir	A matrix is constructed by putting wavelet or curvelet coefficients of each image in row vector. A feature extraction method is developed based on the statistical t-test method. The method is ranking the features according to its capability to differentiate the classes. Then a dynamic threshold is applied to optimize the number of features, which can achieve the maximum classification accuracy rate. SVM is used for Classification Accuracy is 94.79%.
2012	Rodrigo Pereira Ramos, Marcelo Zanchetta, do Nascimento and Danilo Cesar Pereira	Features derived from co-occurrence matrices, wavelet and ridgelet transforms of mammographic images are used. The data set consisted of 120 cranio-caudal mammograms, half containing a mass, rated as abnormal images, and half with no lesions. To select the best set of features, genetic algorithm (GA) is used. Experimental results showed that the best classification rates were obtained with the wavelet based feature extraction using GA for selection of the most relevant

Year	Author(s)	Technique and Result
		features, giving an Az = 0.90.
2013	Belal K. Elfarra and Ibrahim S. I. Abuhaiba	Human features obtained from DDSM Computational features, which are extracted using Square Centroid Lines Gray Level Distribution Method (SCLGM). 410 images taken from DDSM for different types. 31 features are selected from 145 extracted features; 18 of the selected features are from the proposed feature extraction method Receiver Operating Characteristics (ROC) and confusing matrix are used to measure the performance. In training stage, the proposed method achieved an overall classification accuracy of 96.3%, with 92.9% sensitivity and 94.3% specificity. In testing stage, the proposed method achieved an overall classification accuracy of 89%, with 88.6% sensitivity and 83.3% specificity.
2013	Yi-Jhe Huang, Ding-Yuan Chan, Da-Chuan Cheng,	Sequential forward selection (SFS), Sequential backward selection (SBS), and F-score methods. A database of 111 clinical mammograms containing 1,050 microcalcifications (MCs). The accuracy of the system was examined via fROC curve of the test dataset. The system performance for MC identifications was Az = 0.9897, the sensitivity was 92%, and 0.65 false positives (FPs) were generated per image for MCC detection.
2014	Shobha Jose	Biorthogonal wavelet filters via lifting scheme. Maximum likelihood estimator (MLE) is used for texture feature estimation. DDSM is used as the database.

Thus in the literature, histogram based features, statistical features, texture features, features from the spatial domain, spectral domain and fractal domain, Morphometric features, Textural features based on Markov random field (MRF), Gabor wavelets and directional features and features derived from co-occurrence matrices are used. Principal Component Analysis (PCA) for dimensionality Reduction, Rough set for Feature Selection and sequential forward search are used for feature selection.

V. Review on Mammogram Image Classification

Classifiers play an important role in the implementation of computer-aided diagnosis of mammography. The features or a subset of these features are employed by classifiers to classify microcalcifications into benign and malignant.

In 2007 J. Jiang et. al. [42] used Genetic algorithm for classification. They used 188 mammograms from DDSM. Extensive experiments show that the proposed GA design is able to achieve high performances in microcalcification classification and detection, which are measured by ROC curves, sensitivity against specificity, areas under ROC curves and benchmarked by existing representative techniques.

Nikhil R. Pal et. al. 2008 [43] used neural networks for classification. The system is tested on a set of 17 mammograms comprising 10 abnormal and seven normal images which are not used in training and the system is found to perform very well. Moreover for each abnormal image, the system is able to locate the calcified regions quite accurately.

In 2009 Liyang Wei et. al. [44] used Support Vector Machine for classification. They used 200 mammograms from the Department of Radiology at the University of Chicago. Their experimental results reported 0.78 to 0.82 in terms of the area under the ROC curve. Sumeet Dua et. al. [45] used Weighted Association Rule based Classifier. He tested 322 mammograms from MIAS database they attained accuracy of 89%. M. Muthu Rama Krishnan et. al. used Support Vector Machine for classification. They have experimented with two data sets Data Set – I : 699 instances and Data Set – II : 569 instances. Database was created from the University of Wisconsin Hospitals and the classification accuracy attained is : 99.385% for dataset-I and 93.726% for dataset-II.

In 2011 N.C. Tsai et. al. [46], used Multi-layer Perceptron for classification and they reported the classification accuracy as 97.12 at 0.08 false positive per images. Wener BorgesSampaio et. al. [47] used Cellular Neural Networks for classification. They attained the Sensitivity of 80% and rates of 0.84 false positives per image and 0.2 false negatives per image, and an area under the ROC curve of 0.87. Amir Tahmasbi et. al. [50] used Multi-layer Perceptron for 322 Mammograms from MIAS data base. The designed systems yield Az = 0.976, representing fair sensitivity, and Az = 0.975 demonstrating fair specificity. Stylianos D. Tzikopoulos et. al. [48] used Support Vector Machines for classification of 322 Mammograms from MIAS data base. They achieved Classification Accuracy as 84.47%.

In 2012 Jinchang Ren [49] used Support Vector Machine (SVM) and Artificial Neural Network for classification. 748 suspicious MCCs are collected from DDSM. A new strategy namely balanced learning with optimized decision making is proposed to enable effective learning from imbalanced samples, which is further employed to evaluate the performance of ANN and SVM in this context. When the proposed learning strategy is applied to individual classifiers, the performance from both ANN and SVM has been significantly improved. Although ANN outperforms SVM when balanced learning is absent, the performance from the two classifiers becomes very comparable when both balanced learning and optimized decision making are employed. Consequently, an average improvement of more than 10% in the measurements of F1 score and Az measurement are achieved for the two classifiers.

Iuan F. Ramirez Villegas et. al. [50] used Support Vector Machines. They used 23 mammograms from Mias Database and attained 93.75 % accuracy. Again in 2012 Arnau Oliver et. al. [51] used Neural Network is for classification. 23 mammograms from Mias Database are used and that classification accuracy reported is 93.75. Arnau Oliver et. al. [52] used Neural Network for classification. 322 mammograms of the MIAS database and a set of 280 mammograms extracted from a non-public full-field digital database are used for experimental purpose. The result of classification is reported as 80% sensitivity at 1 false positive cluster per image.

Loris Nanni et. al. [53] used support vector machine for classification. 584 Mammograms from DDSM are used for experimental analysis and they are able to attain the area under the ROC curve as Az of 0.97. Discriminant Fusing analysis based Classifier is used by Jun-Bao Li et. al. [54] in which 42 mammogram from MIAS Database were taken and the classification accuracy reported is 95.88%.

Nasseer M. et. al [54] proposed a Computer Aided Diagnosis (CADx) system for classifying abnormal masses in digital mammograms using Support Vector Machines (SVM). The proposed system successfully achieved 93% classification accuracy, which is considered as a good result when compared with similar works in the same research field.

Meenakshi Sundaram K. et. al. [55] applied image mining technique on mammogram to classify the cancer diseases. It can be classified into normal, benign and malignant. They proposed Fuzzy Association Rule Mining. Experiments have been taken dataset with 300 images taken from MIAS of various types to improve accuracy using minimum number of rules to patterns. The experiments and results of the FARM gives better performance compared with existing method.

Jog N. V. et. al [56] used GLDM feature extraction method and SVM classifier. Experiments were conducted on MIAS database. The results show that combination of GLDM feature extractor with SVM classifier is found to give appropriate results.

The results are summarized in Table 3.

TABLE III. REVIEW ON MAMMOGRAM CLASSIFICATION

Year	Author	Classifier	No. of Mammograms used and Results
2007	J. Jiang, B. Yao and A.M. Wason	Genetic Algorithm	188 mammograms from DDSM 300 MC present ROI and 300 Non-MC present ROI Mc100nm300 0.9919 Mc300nm100 0.9868 Mc300nm300 0.987
2008	Nikhil R. Pal, Brojeshwar Bhowmick, Sanjaya K. Patel and Srimanta Pal and J. Das	Neural Networks	17 Mammograms
2009	Liyang Wei, Yongyi Yang and Robert M.Nishikaw	Support Vector Machine	200 mammograms from the Department of Radiology at the University of Chicago. 0.78 to 0.82 in terms of the area under the ROC curve.
2009	Sumeet Dua, Harpreet Singh and H.W.Thompson	Weighted Association Rule based Classifier	322 mammograms from MIAS database. Accuracy 89%.
2011	Stylianos D. Tzikopoulos, Michael E. Mavroforakis, Harris V. Georgiou, Nikos Dimitropoulos and Sergios Theodoridis	Support Vector Machines	322 Mammograms from MIAS data base. Classification Accuracy - 84.47%
2012	Jinchang Ren	Support Vector Machine (SVM) and Artificial Neural Network	748 suspicious MCCs are collected from DDSM No Balanced learning (%) ANN SVM Φ(.,Accu) 4.3 17.6 Φ(.,F1) 4.4 22.2 Φ(.,Az) 5.1 25.7 Balanced Learning (%) ANN SVM Φ(.,Accu) 1.7 1.5 Φ(.,F1) 1.6 1.4 Φ(.,Az) 2.0 1.7
2012	Juan F. Ramirez Villegas and David F.Ramirez-Moreno.	Support Vector and Neural-based Classifier	23 mammograms from Mias Database. Classification Accuracy 93.75 %
2012	Arnau Oliver, Albert Torrent , Xavier Llado, Meritxell Tortajada,	Neural Network	322 mammograms of the MIAS database and a set of 280 mammograms extracted from a non-public

Year	Author	Classifier	No. of Mammograms used and Results
	Lidia Tortajada, Melcior Sentis, Jordi Freixenet and Reyer Zwiggelaar		full-field digital database. 80% sensitivity at 1 false positive cluster per image.
2012	Loris Nanni , Sheryl Brahnam and Alessandra Lumini	Support Vector Mission	584 Mammograms from DDSM. It obtains an area under the ROC curve Az of 0.97.
2012	Jun-Bao Li, Yun-Heng Wanga and Lin-Lin Tang	Discriminant Fusing analysis based Classifier	42 Mammograms from MIAS Database. Classification Accuracy 95.88%
2013	Nasseer M. Basheer and Mustafa H. Mohammed	Support Vector Mission	105 mammography images are acquired from two databases (61 images from the online available MIAS database and 44 images from the Al-joumhory hospital (in Mosul/Iraq) 93% classification accuracy
2014	Meenakshi Sundaram K, Aarthi Rani P, Sasikala D.	Fuzzy Association Rule Mining	300 images taken from MIAS of various types to improve accuracy using minimum number of rules to patterns. The average accuracy of 95% by using precision and recall measures to evaluation method for mammogram classification.
2014	Jog N. V. and Mahadik S.R.,	Support Vector Mission	79 Images from MIAS database Appropriate results were derived

Thus in the literature for mammogram image classification Genetic Algorithm, Neural Networks, Support Vector Machine, Weighted Association Rule based Classifier, Discriminant Fusing analysis based Classifier, Fuzzy Association Rule Mining are used for mammogram classification

VI. CONCLUSION

Review on several mammogram pre-processing techniques, feature extraction, feature selection and classification available in the literature to classify mammograms are highlighted. To remove the noise, artifact and the pectoral muscles methods such as Adaptive pyramids and Minimum Spanning Trees, Non-Linear Diffusion, Raster scanning, Relative Dependency Measure Using Rough Set Theory and Otsu thresholding and multiple regression analysis are used in the literature. A review is also made in mammogram image segmentation and some of the techniques used are Extended Prewitt Operator Kirsch Edge Detector, A region-ranking system, Dynamic-programming based boundary tracing method and a constrained region-growing method, Level set and Marker controlled watershed methods, Tabár tissue modeling, K-Means, Fuzzy C-Means and Possibilistic Fuzzy C-Means and the Histogram Modified Local Contrast Enhancement. In features extraction features such as Fractal features, Morphometric features, Textural features, Gabor wavelets and directional features are studied from the literature. Some of the feature selection techniques available are a sequential forward search, Empirical distributions, MANOVA significance analysis, Neural-Genetic algorithm and Rough Fuzzy based feature selection. Regarding mammogram classification mainly Neural Networks and Support Vector Machines are used as classifier.

REFERENCES

[1] Rangayyan R.M. , “Biomedical Image Analysis”, CRC Press LLC, Boca Raton, 2005.

[2] Fei Ma, Mariusz Bajger, John P. Slavotinek and Murk J. Bottema, “Two graph Theory based Methods for Identifying the Pectoral Muscle in Mammograms”, Pattern Recognition, Vol. 40, pp. 2592 – 2602, 2007.

[3] Mirzaalian H., Ahmadzadeh M.R., Sadri S. and Jafari M., “Pre-processing Algorithms on Digital Mammograms”, In Proceeding of MVA2007 on Machine Vision Applications, Tokyo, Japan May 16-18, 2007.

[4] Subashini T.S., Ramalingam V. and Palanivel S., “Pectoral Muscle removal and Detection of masses in Digital Mammogram using CCL”, International Journal of Computer Applications ,Vol. 1, No. 6, pp 66-70, 2010.

[5] Samir Kumar and Bandyopadhyay, “Pre-processing of Mammogram Images”, International Journal of Engineering Science and Technology, Vol. 2, No. 11, pp 6753-6758, 2010.

[6] Jwad Nagi, Sameem Abdul Kareem, Farrukh Nagi and Syed Khaleel Ahmed, “Automated Breast Profile Segmentation for ROI Detection Using Digital Mammograms”, IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES 2010), Kuala Lumpur, 2011.

[7] Sara Dehghani and Mashallah Abbasi Dezfooli, “A Method for Improve Preprocessing Images Mammography”, Vol. 1, No. 1, pp 90-93, 2011.

[8] Indra Kanta Maitra, Sanjay Nag and Samir K. Bandyopadhyay, “Accurate Breast Contour Detection Algorithms in Digital Mammogram”, International Journal of Computer Applications, Vol. 25, No.5, pp 1-13, 2011.

[9] Sumit Chopra, Pankaj Bhambri, Baljit Singh, “Segmentation of the Mammogram Images to find Breast Boundaries”, IJCST Vol. 2, Issue 2, pp. 164-167, 2011.

[10] Indra Kanta Maitra, Sanjay Nag, Samir Kumar and Bandyopadhyay, Technique for preprocessing of digital mammogram, Computer Methods and Programs in Biomedicine, Vol 107, Issue 2, pp. 175-188, 2012.

[11] Farhan AKRAM, Jeong Heon KIM and Kwang Nam CHOI, A Preprocessing Algorithm for The CAD System of Mammograms using The Active Contour Method, Vol 32, Issue 2, pp 1-13, 2013

[12] Meenakshi Sundaram K. Sasikala D. and Aarthi Rani P. , A Study on Preprocessing a Mammogram Image Using Adaptive Median Filter, Vol. 3, Issue 3, pp. 10333-10337, 2014.

[13] Alfonso Rojas Dom´ınguez and Asoke K. Nandi, "Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection", Computerized Medical Imaging and Graphics, Vol. 32, pp. 304–315, 2008.

[14] Alfonso RojasDomínguez and Asoke K. Nandi, "Toward breast cancer diagnosis based on automated segmentation of masses in mammograms", Pattern Recognition, Vol. 42, pp. 1138 – 1148, 2009.

[15] Aminah Abdul Malek, Wan Eny Zarina Wan Abdul Rahmana, Arsmah Ibrahima, Rozi Mahmudb, Siti Salmah Yasirana and Abdul Kadir Jumaata, "Region and Boundary Segmentation of Microcalcifications using Seed-Based Region Growing and Mathematical Morphology", Procedia Social and Behavioral Sciences, Vol. 8, pp. 634–639, 2010.

[16] Dubeya R. B., Hanmandlub M. and Guptac S.K., "A comparison of two methods for the segmentation of masses in the digital mammograms", Computerized Medical Imaging and Graphics, Vol. 34, pp. 185–191, 2010.

[17] Wenda He, Erika R.E. Denton, Kirsten Stafford and Reyer Zwiggelaar, "Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments", Biomedical Signal Processing and Control, Vol. 6, pp. 321–329, 2011.

[18] Quintanilla-Dominguez J., Ojeda-Magañaa B., Cortina-Januchs M.G., Ruelas R., Vega-Coronab A. and Andinaa D., "Image segmentation by fuzzy and possibilistic clustering algorithms for the identification of microcalcifications", Scientia Iranica, Vol. 18, No. 3, pp. 580–589, 2011.

[19] Sundarama M., Ramarb K., Arumugama N. and Prabina G., "Histogram Modified Local Contrast Enhancement for mammogram images", Applied Soft Computing, Vol. 11, pp. 5809–5816, 2011.

[20] Pelin Kus and IrfanKaragoz, "Fully automated gradient based breast boundary detection for digitized X-ray mammograms", Computers in Biology and Medicine, Vol. 42, pp. 75–82, 2012.

[21] Wei-Yen Hsu, "Improved watershed transform for tumor segmentation, Application to mammogram image compression", Expert Systems with Applications, Vol. 39, pp. 3950–3955, 2012.

[22] Subash Chandra Boss R., Thangavel K., Arul Pon Daniel K., "Mammogram Image Segmentation Using Rough Clustering", International Journal of Research in Engineering and Technology, Vol 2, pp. 2321-7308 , 2013.

[23] Pereira DC, Ramos RP, do Nascimento MZ, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm", Computer Methods and Programs in Biomedicine, Vol. 114, pp. 88-104.

[24] Kanimozhi Suguna S. and Uma Maheswari S., Performance Analysis of Feature Extraction and Selection of Region of Interest by Segmentation in Mammogram Images between the Existing Meta-heuristic Algorithms and Monkey Search Optimization (MSO), Vol 11, pp. 72-82, 2014.

[25] Roman W. Swiniarski and Andrzej Skowron, "Rough set methods in feature selection and recognition", Pattern Recognition Letters,Vol. 24, pp. 833–849, 2003.

[26] Ping Zhang, Brijesh Verma and Kuldeep Kumar, Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection, Pattern Recognition Letters Vol. 26, pp. 909–919, 2005.

[27] Fu J.C., Lee S.K., Wong S.T.C., Yeh J.Y., Wang A.H. and Wu H.K., "Image segmentation feature selection and pattern classification for mammographic microcalcifications", Computerized Medical Imaging and Graphics, Vol. 29, pp. 419–429, 2005.

[28] Dar-Ren Chen, Ruey-Feng Chang, Chii-Jen Chen, Ming-Feng Ho, Shou-Jen Kuo, Shou-Tung Chen, Shin-Jer Hung and Woo Kyung Moon, "Classification of breast ultrasound images using fractal feature", Journal of Clinical Imaging, Vol. 29, pp. 235–245, 2005.

[29] Gobert N. Lee, Murk J. Bottema, "Significance of classification scores subsequent to feature selection", Pattern Recognition Letters, Vol. 27, pp.1702–1709, 2006.

[30] Sheshadri H.S. and Kandaswamy A., "Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms", Computerized Medical Imaging and Graphics, Vol. 31, pp. 46–48, 2007.

[31] Harris Georgiou , Michael Mavroforakis , Nikos Dimitropoulos, Dionisis Cavouras and Sergios Theodoridis , "Multi-scaled morphological features for the Characterization of mammographic masses using statistical classification schemes", Artificial Intelligence in Medicine, Vol. 41, pp. 39-55, 2007.

[32] Brijesh Verma and Ping Zhang, "A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms", Applied Soft Computing, Vol. 7, pp. 612–625, 2007.

[33] Alfonso Rojas-Domínguez and Asoke K.Nandi, "Development of tolerant features for characterization of masses in mammograms", Computers in Biology and Medicine, Vol. 39, pp. 678 – 688, 2009.

[34] Wagner Coelho A. Pereira , Andre´ V. Alvarenga , Antonio Fernando C.Infantosi, Leonardo Macrini and Carlos E.Pedreira, "A non-linear morphometric feature selection approach for breast tumor contour from ultrasonic images", Computers in Biology and Medicine, Vol. 40, pp. 912–918, 2010.

[35] Sung-Nien Yu and Yu-Kun Huang , "Detection of microcalcifications in digital mammograms using combined model-based and statistical textural features", Expert Systems with Applications, Vol. 37, pp. 5461–5469, 2010.

[36] Ioan Buciu and Alexandru Gacsadi, "Directional features for automatic tumor classification of mammogram images", Biomedical Signal Processing and Control, Vol. 6, pp. 370– 378, 2011.

[37] Mohamed Meselhy Eltoukhy , Ibrahima Faye and Brahim Belhaouari Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation", Computers in Biology and Medicine, Vol. 42, pp. 123–128, 2012.

[38] Rodrigo Pereira Ramos , Marcelo Zanchetta, do Nascimento , Danilo Cesar Pereira, "Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms", Expert Systems with Applications, Vol. 39 , pp. 11036–11047, 2012.

[39] Belal K. Elfarra and Ibrahim S. I. Abuhaiba, "New Feature Extraction Method for Mammogram Computer Aided Diagnosis", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 6, pp. 13-36, 2013.

[40] Yung-Jen Ho, Po-Pang Tsai, Wu-Chung Shen, and Rui-Fen Chen, "Automated Feature Set Selection and Its Application to MCC Identification in Digital Mammograms for Breast Cancer Detection", Sensors, Vol. 13, pp. 4855-4875, 2013.

[41] Shobha Jose, "Texture Feature Extraction for Mammogram Images Using Biorthogonal Wavelet Filter via Lifting Scheme", International Journal of Science and Research, Vol. 3, pp. 1252-1256, 2014.

[42] Jiang J., Yao B. and Wason A.M., "A genetic algorithm design for microcalcification detection and classification in digital mammograms", Computerized Medical Imaging and Graphics, Vol. 31, pp. 49–61, 2007.

[43] Nikhil R. Pal, Brojeshwar Bhowmick, Sanjaya K. Patel , Srimanta Pal and J. Das, "A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms", Neuro computing, Vol. 71 pp. 2625– 2634, 2008.

[44] Liyang Wei, Yongyi Yang and Robert M.Nishikaw, "Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis", Pattern Recognition, Vol. 42, pp. 1126 – 1132, 2009.

[45] Sumeet Dua, Harpreet Singh and Thompson H.W., "Associative classification of mammograms using weighted rules", Expert Systems with Applications, Vol. 36, pp. 9250–9259, 2009.

[46] Muthu Rama Krishnan M., Shuvo Banerjee, Chinmay Chakraborty, Chandan Chakraborty and Ajoy K. Ray, "Statistical analysis of mammographic features and its classification using support vector machine", Expert Systems with Applications, Vol. 37, pp, 470–478, 2010.

[47] Tsai N.C., Chen H.W. and Hsu S.L., "Computer-aided diagnosis for early-stage breast cancer by using Wavelet Transform", Computerized Medical Imaging and Graphics, Vol. 35, pp. 1–8, 2011.

[48] Wener BorgesSampaio, EdgarMoraesDiniz, Aristo fanes Correa Silva , Anselmo CardosodePaiva and MarceloGattass , "Detection of masses in mammogram images using CNN, geostatistic functions and SVM", Computers in Biology and Medicine, Vol. 41, pp. 653–664, 2011.

[49] Amir Tahmasbi, FatemehSaki and Shahriar B.Shokouhi, "Classification of benign and malignant masses based on Zernike moments", Computers in Biology and Medicine, Vol. 41, pp. 726–735, 2011.

[50] Stylianos D. Tzikopoulos, Michael E. Mavroforakis, Harris V. Georgiou, Nikos Dimitropoulos and Sergios Theodoridis, "A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry", Computer Methods and Programs in Biomedicine, Vol. 102, pp. 47–63, 2011.

[51] Jinchang Ren, "ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging", Knowledge-Based Systems, Vol. 26, pp. 144–153, 2012.

[52] Juan F. Ramirez Villegas and David F. Ramirez Moreno, "Wavelet packet energy, Tsallis entropy and statistical parameterization for support vector-based and neural-based classification of mammographic regions", Neurocomputing, Vol. 77, pp. 82–100, 2012.

[53] Arnau Oliver, Albert Torrent , Xavier Llado, Meritxell Tortajada, Lidia Tortajada, Melcior Sentis, Jordi Freixenet and Reyer Zwiggelaar, Automatic microcalcification and cluster detection for digital and digitised mammograms, Knowledge Based Systems, Vol. 28, pp. 68–75, 2012.

[54] Loris Nanni , Sheryl Brahnam and Alessandra Lumini, "A very high performing system to discriminate tissues in mammograms as benign and malignant", Expert Systems with Applications, Vol. 39, pp. 1968–1971, 2012.

[55] Jun-Bao Li, Yun-Heng Wanga and Lin-Lin Tang, "Mammogram-based discriminant fusion analysis for breast cancer diagnosis", Clinical Imaging, Vol 36, Issue 6, pp. 710-716, 2012.

[56] Nasseer M. Basheer and Mustafa H. Mohammed, "Classification of Breast Masses in Digital Mammograms Using Support Vector Machines", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, pp. 57-63.