Automatic Detection and Classification of Microcalcification, Mass, Architectural Distortion and Bilateral Asymmetry in Digital Mammogram

S. Shanthi, V. Muralibhaskaran

Abstract—Mammography has been one of the most reliable methods for early detection of breast cancer. There are different lesions which are breast cancer characteristic such as microcalcifications, masses, architectural distortions and bilateral asymmetry. One of the major challenges of analysing digital mammogram is how to extract efficient features from it for accurate cancer classification. In this paper we proposed a hybrid feature extraction method to detect and classify all four signs of breast cancer. The proposed method is based on multiscale surrounding region dependence method, Gabor filters, multi fractal analysis, directional and morphological analysis. The extracted features are input to self adaptive resource allocation network (SRAN) classifier for classification. The validity of our approach is extensively demonstrated using the two benchmark data sets Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammograph (DDSM) and the results have been proved to be progressive.

Keywords—Feature extraction, fractal analysis, Gabor filters, multiscale surrounding region dependence method, SRAN.

I. Introduction

BREAST cancer is a type of cancer originating from breast tissue, most commonly from the inner lining of milk ducts, milk passages that connect the lobules to the nipple or the lobules, milk-producing glands that supply the ducts with milk. It is the second leading cause of cancer death among women mainly those who are in the 40-55 age group [1]. The breast cancer incident has increased in most countries worldwide in the last decades, with the most rapid increases occurring in many of the developing countries [2]. Mammography is an effective tool available in the recent days for the reliable detection of early and potentially curable breast cancer. Mammographic screening has been shown to be effective in reducing breast cancer mortality rates by 30-70% [1].

In the manual detection and diagnosis of breast cancer through digital mammography, abnormalities and cancers may be missed, False Negatives (FN), and non-cancerous lesions may be misclassified as cancers, False Positives (FP). Currently, Computer Aided Diagnosis (CAD) system could offer a cost effective alternative to double reading of mammograms and can suggest the radiologist about where the abnormalities present in the mammogram and what type of abnormalities present in it for reducing the errors in diagnosis [3]-[22].

The most common mammographic signs of breast cancer defined by American College of Radiology [23] are as follows:

- Calcifications (CALC): Calcifications are small calcium deposits that form in the breast because of benign or malignant processes. Mammographically, they appear as bright white spots of various sizes and shapes.
- Masses (MASS): A "Mass" is a space-occupying lesion seen in two different projections. If a potential mass is seen in, only a single projection it should be called a "Density" until its three-dimensionality is confirmed.
- Architectural distortion (ARCH): The normal architecture (of the breast) is distorted with no definite mass visible. This includes spiculations radiating from a point and focal retraction or distortion at the edge of the parenchyma.
- Bilateral asymmetry (ASYM): Asymmetry of breast parenchyma between the two sides has been one of the most useful signs for detecting primary breast cancer. Bilateral asymmetry, i.e. asymmetry of the breast parenchyma between left and right breast, may indicate breast cancer in its early stage.
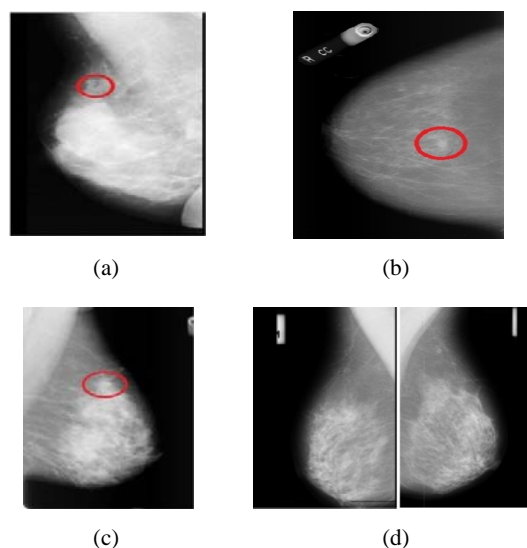

Fig. 1 Mammograms of different signs of abnormalities (a) Microcalcification (b) Masses (c) Architectural distortion (d) Bilateral asymmetry

S. Shanthi is with the Kongu Engineering College, Tamil Nadu, India (phone: 9942911551, e-mail: shanthi.kongumca@ gmail.com).
V. Muralibhaskaran is with the Dhirajlal Gandhi College of Engineering and Technology, Tamil Nadu, India (e-mail: murali66@gmail.com).

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:8, No:11, 2014

Figs. 1 and 2 show the signs of breast cancer and the corresponding Region Of Interest (ROI) respectively.

The feature extraction techniques to identify microcalcifications and masses have been extensively studied, but feature extraction techniques to identify architectural distortion and asymmetry in mammograms still are challenges [9]. Even though several methods of tissue identification are available in literature, not a single method is applicable for identification of all the signs of breast cancer. Hence, we proposed novel feature extraction techniques to detect architectural distortion and asymmetry together with the detection of microcalcification and masses.
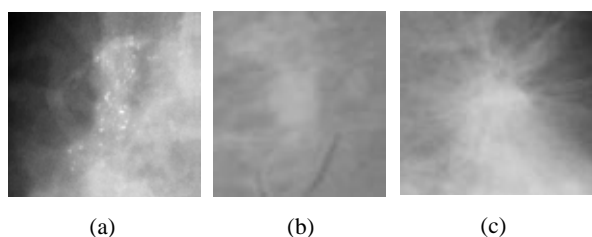

(a)   (b)   (c)

Fig. 2 ROI of different signs of abnormalities namely (a) Microcalcification, (b) Masses and (c) Architectural distortion respectively

II. RELATED WORK

Features play an important role in detecting and classifying the levels of abnormalities in any medical image. Hence feature extraction technique is very important step in CAD system because that significantly affects the success of the CAD system.

Most systems extract features to detect and classify the abnormality into benign or malignant through textures, statistical properties, spatial domain, fractal domain, and wavelet bases. Some statistical texture analysis methods have been used to detect masses or microcalcifications, such as: Gray level difference statistics (GLDS), GLCM (Gray Level co-occurrence matrix) also called SGLD (Spatial Gray Level Dependence Matrix), Gray level difference method (GLDM), Gray level run length method (GLRLM), Surrounding Region Dependence Method (SRDM), Multiscale Surrounding Region Dependence Method (MSRDM), Laws' texture energy measures [14]-[18], [24].

Similarly, multiresolution based feature extraction has proved to be useful in identifying mass and microcalcifications [5], [18].

On the contrary Gabor filters, phase portrait analysis, analysis of the angular spread of power, fractal analysis, Laws' texture energy measures, and Haralick's texture features used to detect the sites of architectural distortion [3].

Likewise mathematical morphology to detect architectural distortion around the skin line, and a concentration index to detect architectural distortion with mammary gland are used [11].

In the same way, combining directional information with morphological measures, and geometric moments related to density distributions are used to identify asymmetry in mammograms [13]. Karnan & Thangavel have devised a method to analyze the bilateral asymmetry based on Genetic Algorithm. (GA). GA is used to find the breast border and the nipple position based on which, the mammogram images are aligned and subtracted to extract the suspicious region [10].

From the existing publication, it is observed that the available feature extraction techniques identify either microcalcification, mass, architectural distortion or bilateral asymmetry. But, in this work a novel features are proposed to identify all the four signs of breast cancer.

III. PROPOSED METHODOLOGY

The diagrammatic representation of the proposed methodology is shown in Fig. 3.

A. Detection of ROI

The ROI is identified using intutionistic fuzzy c (IFCM) means clustering [4].

B. Feature Extraction Techniques

The result of IFCM clustering is the suspected ROIs from the mammogram image. Subsequently a set of features from all the ROIs are extracted. The proposed feature extraction technique is based on discrete wavelet transformation, surrounding region dependence method, Gabor filter, multifractal analysis, directional and morphological analysis.

1. Multifractal Features

Multifractal analysis acts as a powerful tool in many medical applications because of its self-similarity property. The applications of multifractal analysis include segmentation, characterization of Electrocardiography (ECG) signals, characterization of brain images and characterization of mammogram [25]. Even though multifractal analysis has been widely used in the analysis of biomedical images, no method is available to study and classify abnormalities in a mammogram. Cancerous tumors exhibit a certain degree of randomness associated with their growth, and are typically irregular and complex in shape. Therefore multifractal analysis can provide a better measure of their complex patterns than the conventional Euclidean geometry.

Multifractal dimension shows different scale characteristics in different areas. It represents the global features through the local part of whole system and describes the essence of the fractal structures more accurately.

Using Renyi fractal dimension spectra, more information in the multifractal image structure can be revealed [26]. The Renyi dimension Dq is defined as:

$$D_q = \lim_{r \to 0} \frac{1}{q-1} \frac{log \sum_{j=1}^{N} p_j^q}{log(r)} \quad -\infty \leq q \leq \infty \quad (1)$$

where r is the size of the box, p_i is the frequency of the occurrence of box r and N is the number of boxes of size r that covers the underlying set.

Multifractal dimension, one of the multifractal features has better discriminating ability. In this study, the optimum

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:8, No:11, 2014

number of multifractal dimensions such as Homogeneity Dimension (HD), FD, Entropy Dimension (ED) and Correlation Dimension (CD) are used to extract different features from the ROI [26].
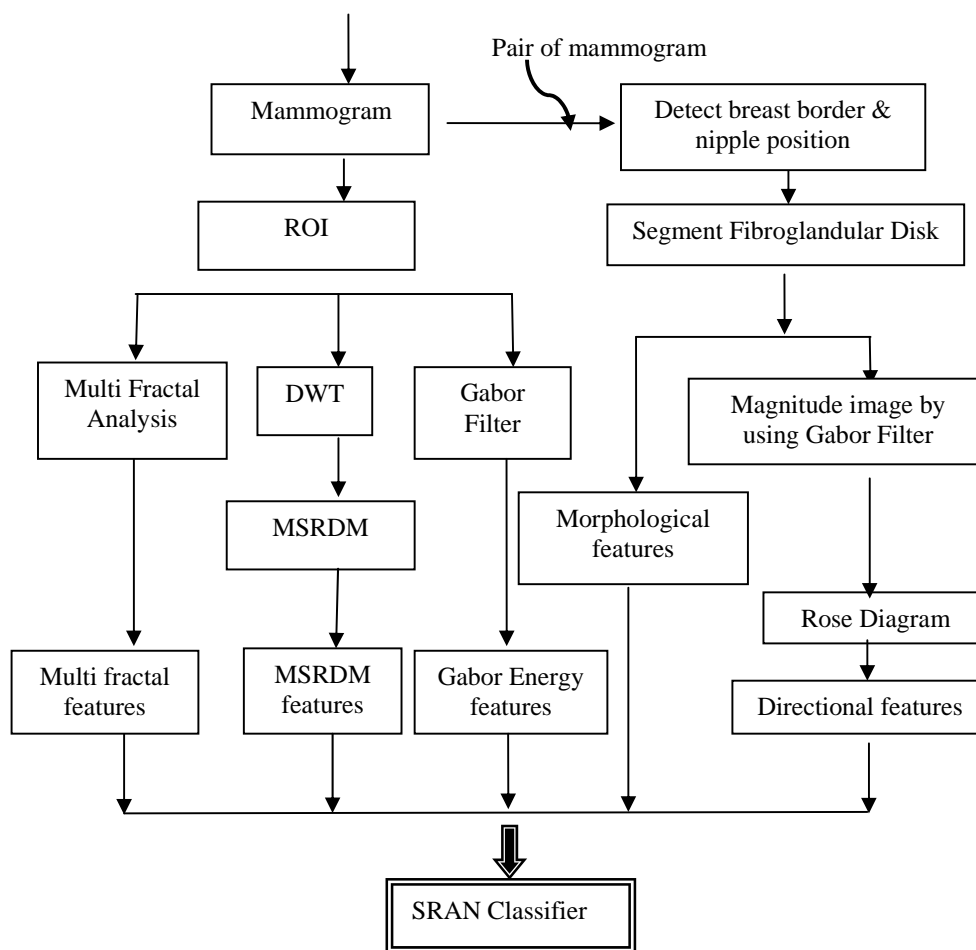

Fig. 3 Schematic representation of the proposed methodology

2. MSRDM Features

Image analysis in multiple scales allows image resolution to change so as to process as little data as possible by selecting relevant details for a given visual task. The basic idea of multiscale analysis is to represent an image on several subimages, from coarse to fine resolution, and analyze them in the frequency domain. The multiresolution quality allows for the analysis of gray level pixels regardless of the size of the neighborhood. These properties lead to the idea that wavelets could guide researchers to better texture classification. When extracting texture features, it is necessary to measure texture features on neighborhoods of different sizes [18].

3. Gabor Energy Features

Two dimensional Gabor filter decomposes an image into components corresponding to different scales and orientations, capturing visual properties such as spatial localization, orientation and spatial frequency. It is similar to those of the human visual system, and found to be particularly appropriate for texture representation and discrimination. Sixteen Gabor filters are generated at 4 scales (2, 4, 32, 64) and 4 orientations (horizontal, vertical, 45 and 135). Therefore, 16 sub images

are obtained from Gabor filter bank and energy of each sub image is considered for future computation [18].

4. Directional Features

In order to extract the directional and morphological features, fibroglandular disk is first segmented from the breast. To segment the fibroglandular disk, the breast border and the nipple position are identified. Then directional information of fibroglandular disk is analyzed based on the detection of linear components using Gabor filter.

5. Identification of Breast Border

A histogram-based thresholding technique is used to generate a binary image to divide the breast and the non-breast region. The local optimum in the histogram is selected as the threshold value. The intensity values smaller than this threshold value are changed into black, and the intensity values greater than the threshold are changed into white in order to perform the morphological operation to remove the connected components. So the binary image contains only the breast border. The spatial coordinates of the border points are mapped with the original gray scale mammogram image.

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:8, No:11, 2014

These spatial coordinates of the border points and nipple position are used to determine the angle of rotation to align the mammogram.

6. Identification of Pectoral Region

Pectoral muscle region has to be removed to increase the detection performance. A window of size 7×7 pixels is extracted for each pixel centered at the pixel location. The median value is computed for this window. The intensity value of the center pixel is replaced with the median value. To remove the pectoral muscle region, initially the histogram is generated for the mammogram image. The global optimum in the histogram is selected as the threshold value. When the MLO view is correctly imaged, the pectoral region should always appear as a high-intensity and triangular region across the upper posterior margin of the image. In several cases, the upper part of the border line is a sharp intensity edge while the lower part is more likely to be a texture edge due to the fact that it is overlapped by a fibroglandular tissue. After finding the global optimum value, the image is scanned from top left to right in case of left breast or top right to left in case of right breast in the triangular region across the upper posterior margin of the mammogram image. The intensity values greater than this threshold are changed into black and the gray values smaller than the threshold are maintained as such so as to convert the pictorial region as black region.

7. Segmentation of the Fibroglandular Disk

In the bilateral asymmetry evaluation, only the fibroglandular disk is used as the ROI in order to compute the oriented components because most of the directional components such as connective tissue and ligaments exist in the fibroglandular region of the breast. To segment the fibroglandular disk, sIFCm clustering is used. In this study, the number of clusters is selected as two by experimentation. After alignment, fibroglandular disks are segmented from the left and right breast images.

For each left and right mammogram the rose diagram is obtained. Subsequently, the two rose diagrams are subtracted to obtain a difference rose diagram. Then the directional features, namely first angular moment, second angular moment, entropy, dominant orientation and circular variance are calculated from the difference rose diagram [13].

8. Morphological Features

Eleven morphological features including seven of Hu's moments ($\emptyset_1$ *to* $\emptyset_7$), area, average density, Eccentricity (E) and stretch (ρ) are computed from the segmented fibroglandular disks [13].

C. Classification

Self adaptive resource allocation network classifier is used for classification. The SRAN classifier is a sequential learning algorithm with self-regulated control parameters. Since, the SRAN classifier uses explicit classification error in growing/learning criterion and discarding similar samples, it prevents overtraining and provides better generalization performance [18].

IV. EXPERIMENTAL RESULTS AND DISCUSSION

For the current study, the images are taken from two public and widely known databases: the Mammographic Image Analysis Society (MIAS) database [27] and Digital database for screening mammography (DDSM) database [28]. The performance of the proposed approach can be estimated based on accuracy, sensitivity and specificity.

$$Accuracy = \frac{TP+TN}{TP+TN+TP+FN} \tag{2}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{3}$$

$$Specivicity = \frac{TN}{TN+FP} \tag{4}$$

TP => True Positive: a patient predicted with cancer when the patient actually has cancer.

TN => True Negative: a patient predicted healthy when the patient is actually healthy.

FN => False Negative: a patient predicted healthy when the patient actually has cancer

FP => False Positive: a patient predicted with cancer when the patient is actually healthy.

The classification performance of the method in terms of accuracy, sensitivity and specificity is shown in Table I.

TABLE I
CLASSIFICATION RESULTS OF THE PROPOSED METHOD

Dataset	Accuracy	Sensitivity	Specificity
MIAS	94.72	0.92	0.96
DDSM	93.75	0.93	0.95

From Table I, it is observed that the proposed feature set produces an accuracy of 94.72% and 93.75% for MIAS and DDSM with the sensitivity of 0.92 and 0.93, respectively.

As there is no work in the literature to detect and classify all the signs of the breast cancer together, the comparison is done in a twofold manner. First, the proposed feature set is used to classify each sign separately. Second, the same feature set is used to classify all the signs together. Then both the type of results is compared with the existing techniques and summarized in Table II.

It is found from Table II that the classification performance of the proposed technique is comparable and better than that achieved by a number of existing techniques.

V. CONCLUSION

This chapter has proposed a feature extraction approach that uses MSRDM, Gabor filters, multifractal analysis, directional and morphological analysis for analyzing and classifying microcalcification, mass, architectural distortion and bilateral asymmetry in digital mammograms. The approach is evaluated on a subset of the MIAS and DDSM benchmark databases.The advantage of this approach lies in the fact that it concentrates and extracts features to detect all the signs of the breast cancer such as microcalcification, masses, architectural distortion and bilateral asymmetry in mammogram images.

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:8, No:11, 2014

TABLE II
SUMMARY OF CLASSIFICATION PERFORMANCE OF EXISTING TECHNIQUES AND THE PROPOSED METHOD

Method	Signs for Breast cancer	Dataset / No. of images used	Accuracy
[14]	CALC	DDSM/85	AUC of 0.99
[12]	CALC	MIAS/(25 MC+297 Normal) Digital(90 MC+190 Normal)	ROC of 0.90 0.92
[17]	CALC	Real dataset/66 59 MCCs, 683 MCs	89%
Proposed	CALC	MIAS DDSM	95.68% 94.33%
[19]	Mass	MIAS & real dataset/111	ROC of 0.95
[20]	Mass	DDSM/433	Sensitivity of 86%
[16]	Mass	DDSM/16	AUC of 0.88
Proposed	Mass	MIAS DDSM	96.59% 94.67%
[21]	ARCH	Real dataset/ 25	Sensitivity of 80%
[22]	ARCH	4224 ROIs	sensitivity of 80%
[3]	ARCH	Real Prior mammogram and 106 case with 4224 ROI	ROC of 0.75
Proposed	ARCH	MIAS DDSM	95.15% 94.4%
[10]	ASYM	MIAS/114	90.6%
[13]	ASYM	MIAS/88	84.4%
Proposed	ASYM	MIAS DDSM	95.06% 94.39%
Proposed	All Signs	MIAS/322 DDSM/464	94.72% 93.75%

REFERENCES

[1] American Cancer Society, "Glopbal Cancer Facts & Figures", American Cancer Society, Atlanda, 2011.

[2] International Agency for Research on Cancer IARC, World Cancer Report 2008: IARC.

[3] S.Banik, R.M. Rangayyan, and J.E.L. Desautels, "Measures of angular spread and entropy for the detection of architectural distortion in prior mammograms", International Journal of Computer Assisted Radiology and Surgery, vol. 8, pp. 121–134, 2013.

[4] S. Shanthi and V. Murali Bhaskaran, "Intuitionistic fuzzy c-means and decision tree approach for breast cancer detection and classification", European Journal of Scientific Research, vol. 66, no.3, pp. 345-351, 2011.

[5] S. Shanthi and V. Murali Bhaskaran, "A Novel Approach for Detecting and Classifying Breast Cancer in Mammogram Images", International Journal of Intelligent Information Technologies, vol. 9, no. 1, pp. 21-39, 2013.

[6] S. Shanthi and V. Murali Bhaskaran, "Computer aided system for detection and classification of breast cancer", International Journal of Information Technology, Control and Automation, vol. 2, no.4, pp. 87-98, 2012.

[7] V. Nguyen, "An Automated Method to Segment and Classify Masses in Mammograms", World Academy of Science, Engineering and Technology, vol. 3, no. 4, pp. 761 – 766, 2009.

[8] A. Sindhuja and V. Sadasivam, "Automatic Detection of Breast Tumors in Sonoelastographic Images Using DWT", World Academy of Science, Engineering and Technology, vol. 7, no. 9, pp. 596 – 602, 2013.

[9] S. Banik, R.M. Rangayyan and J.E.L. Desautels, "Detection of architectural distortion in prior mammograms", IEEE Transactions on Medical Imaging, vol. 30, no. 2, pp. 279-294, 2011.

[10] M. Karnan and K. Thangavel, "Automatic detection of the breast border and nipple positionon digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications", Computer Methods and Programs in Biomedicine, vol. 87, pp. 12–20, 2007.

[11] T. Matsubara, T. Ichikawa, T. Hara, H. Fujita, S. Kasai, T. Endo and T. Iwase, "Automated detection methods for architectural distortions around skin line and within mammary gland on mammograms", International Congress Series, vol. 1256, pp. 950–955, 2003.

[12] A. Oliver, A. Torrent, X. Llado, M. Tortajada, L. Tortajada, M. Sentis, J. Freixenet, and R. Zwiggelaar, "Automatic microcalcification and cluster detection for digital and digitised mammograms", Knowledge-Based Systems, vol. 28, pp. 68–75, 2012.

[13] R.M. Rangayyan, R.J. Ferrari, and A.F. Frere, "Analysis of bilateral asymmetry in mammograms using directional, morphological, and density features", Journal of Electronic Imaging, vol. 16, no. 1, pp. 013003-1-013003-12, 2007.

[14] A.N. Karahaliou, I.S. Boniatis, S.G. Skiadopoulos, F.N. Sakellaropoulos, N.S. Arikidis, E.A. Likaki, G.S. Panayiotakis, and L.I. Costaridou, "Breast cancer diagnosis: analyzing texture of tissue surrounding microcalcifications", IEEE Transactions on Information Technology in Biomedicine, vol. 12, no. 6, pp. 731– 738,2008.

[15] J.K. Kim, and H.W. Park, "Statistical textural features for detection of microcalcifications in digitized mammograms", IEEE Transactions on Medical Imaging, vol. 18, no. 3, pp. 231-238, 1999.

[16] A. Mencattini, M. Salmeri, G. Rabottino, and S. Salicone, "Metrological characterization of a CADx system for the classification of breast masses in mammograms", IEEE Transactions on Instrumentation and Measurement, vol. 59, no. 11, pp. 2792–2799, 2010.

[17] A. Tiedeu, C. Daul, A. Kentsop, P. Graebling, and D. Wolf, "Texture-based analysis of clustered microcalcifications detected on mammograms", Digital Signal Processing, vol. 22, no. 1, pp. 124–132, 2012.

[18] S. Shanthi and V. Murali Bhaskaran, "A Novel Approach for Classification of Abnormalities in Digitized Mammograms", Sadhana - Academy Proceedings in Engineering Science journal, vol. 39, no. 5, pp. 1141–1150, 2014.

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:8, No:11, 2014

[19] T. Mu, A.K. Nandi, and R.M. Rangayyan, "Classification of breast masses using selected shape, edge-sharpness, and texture features with linear and kernel-based classifiers", Journal of Digital Imaging, vol. 21, no. 2, pp. 153-169, 2008.

[20] L de O. Martins, G.B. Junior, A.C. Silva, A.C.D.E. Paiva, and M. Gattass, "Detection of Masses in Digital Mammograms using K-means and Support Vector Machine", Electronic Letters on Computer Vision and Image Analysis, vol. 8, no. 2, pp. 39-50, 2009.

[21] M. Nemoto, S. Honmura, A. Shimizu, D. Furukawa, H. Kobatake, and S. Nawano, "A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows", International Journal of Computer Assisted Radiology and Surgery, vol. 4, no. 1, pp. 27–36, 2009.

[22] R.M. Rangayyan, S.Banik, and J.E.L. Desautels, "Computer-aided detection of architectural distortion in prior mammograms of interval cancer", Journal of Digital Imaging, vol. 23, no. 5, pp. 611-631, 2010.

[23] American College of Radiology, 1998, "Illustrated Breast Imaging Reporting and Data System BI-RADS", American College of Radiology, Reston, PA, 3rd edition, 1998.

[24] R.M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification", IEEE Transactions of System, Man, Cybernetics, SMC-3, pp. 610–621, 1973.

[25] R. Lopes and N. Betrouni, "Fractal and multifractal analysis: a review", Medical Image Analysis, vol. 13, pp. 634–649, 1999.

[26] W. Kinsner, "A unified approach to fractal dimensions", International Journal of Cognitive Informatics and Natural Intelligence, vol.1, pp. 26-46, 2007.

[27] J. Suckling, J. Parker, D.R. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, SL. Kok, P. Taylor, D. Betal, and J. Savage, "The Mammographic Image Analysis Society digital mammogram database", Proceeding of International Workshop on Digital Mammography, pp. 211–221,1994.

[28] M. Heath, K. Bowyer, D. Kopans, R. Moor, and W.P. Kegelmeyer, Proceedings of the Fifth International Workshop on Digital Mammography, M.J. Yaffe, ed. pp. 212-218, Medical Physics Publishing, ISBN 1-930524-00-5, 2001.

S. Shanthi received her BSc Degree in Computer Science, MCA degree in Computer Applications, Master of Philosophy in Computer Science from Bharathiar University, Coimbatore, India in the year 1993, 1996 and 2004 respectively and ME degree in Computer Science and Engineering from Anna University, Chennai, India in the year 2006. Currently, she is pursuing the PhD degree in Computer Science and Engineering at Anna University, Chennai, India.

She is presently working as an Assistant Professor (SLG) in the Department of Computer Applications, Kongu Engineering College, Tamil Nadu, India. Her area of interest includes, Data Mining, Image Processing, Pattern Recognition and Soft Computing.

V. Murali Bhaskaran is a Principal of Dhirajlal Gandhi College of Engineering and Technology, Tami Nadu, India. He obtained his B.E. Degree in Computer Science and Engineering from Bharathidasan University in the year 1989, M.E. degree in Computer Science and Engineering from Bharathiar University in the year 2000 and secured University III rank and Ph.D in Computer Science and Engineering from Bharathiar University in the year 2008.

He has published 40 papers in Journals and 40 papers in Conferences both at National and International level. He is a life member of ISTE, CSI and ACS. Also he is a fellow member in IE (FIE). His areas of interest include Computer Architecture, Computer Networks, Network Security, Information Security, etc.