Mammogram Classification using Gray-Level Co-occurrence Matrix for Diagnosis of Breast Cancer

Ranjit Biswas
Department of CSE
Assam University
Silchar, India
ranjit_tb@yahoo.co.in

Abhijit Nath
Department of CSE
Assam University
Silchar, India
abhijitnath2007@gmail.com

Sudipta Roy
Department of CSE
Assam University
Silchar, India
sudipta.it@gmail.com

***Abstract*— Breast cancer is one of the most common forms of cancer in women worldwide. Most cases of breast cancer can be prevented through screening programs aimed at detecting abnormal tissue. So, early detection and diagnosis is the best way to cure breast cancer to decrease the mortality rate. Computer Aided Diagnosis (CAD) system provides an alternative tool to the radiologist for the screening and diagnosis of breast cancer. In this paper, an automated CAD system is proposed to classify the breast tissues as normal or abnormal. Artifacts are removed using ROI extraction process and noise has been removed by the 2D median filter. Contrast-Limited Adaptive Histogram Equalization (CLAHE) algorithm is used to improve the appearance of the image. The texture features are extracted using Gray Level Co-occurrence Matrix (GLCM) of the region of interest (ROI) of a mammogram. The standard Mammographic Image Analysis Society (MIAS) database images are considered for the evaluation. K-Nearest Neighbor (KNN), Support Vector Machine (SVM) and Artificial Neural Network (ANN) are used as classifiers. For each classifier, the performance factor such as sensitivity, specificity and accuracy are computed. It is observed that the proposed scheme with 3NN classifier outperforms SVM and ANN by giving 95% accuracy, 100% sensitivity and 90% specificity to classify mammogram images as normal or abnormal.**

Keywords—Mammogram; ROI; GLCM; Confusion Matrix; 3NN classifier.

I. INTRODUCTION

Breast cancer grows in the breast cells usually in the lobules (glands which produce milk) and in the milk ducts (that carry milk to the nipple). In 2015, 40290 women death reported due to breast cancer [1]. The best way to identify the presence of breast cancer at an early stage is by interpreting mammogram images. Though mammography is an effective screening tool used by the radiologist for breast cancer detection at an early stage [2], however by visual interpretation it's very difficult for the radiologist to classify the affected tissue whether it is cancerous or not. Sometimes, the mammography test can give negative results in the presence of tumor by not detecting the exact ROI where tumor is present. In such cases, further tests are required which can be expensive and time-consuming. Also, abnormal tissue region may also be missed in visual (manual) interpretation of a mammogram image. In recent years Computer Aided Detection (CAD) system are being used to solve the problem of interpretation of the mammogram image. (The mass or calcifications in the breast tissue which may get unnoticed in visual interpretation by radiologists are easily and effectively being detected using CAD systems. The development and fine tuning of the CAD system has become more crucial for early and effective detection of abnormal tissues in the given digital mammogram image [3]. Thus, the main objective of CAD system is to increase diagnosis accuracy and enhancing the mammogram interpretation. Thus, CAD system can reduce the variability in judgments among radiologists by providing an accurate diagnosis of digital mammograms.

Rest of the paper is organized as follows: section 2 discusses about computer aided breast cancer detection systems; a new automated system is proposed and discussed in section 3; section 4 depicts the experimental results; conclusion and future work is discussed in section 5.

II. RELATED WORK

A CAD system was proposed in [4] for automatic detection and classification of breast cancer where noise and pectoral region removed from the breast using morphological operations and histogram based methods respectively. Using intuitionistic FCM based clustering, ROI is extracted and that ROI is transformed into four sub bands with the help of discrete wavelet transform (DWT). Wavelet energy features and 13 gray level co-occurrence features are computed form these sub bands. Then the self-adaptive resource allocation network (SRAN) classifier is used for classification. In 2013, tumor cut algorithm was proposed by Maanasa et al. for segmentation of mammogram image for detection of breast cancer [5]. Noise and artifact is removed using Gabor filter. Then from segmented image, geometrical and textural features are extracted. The optimal features are calculated using anova test calculator. With the help of optimal features, the mammogram image is classified into normal, benign and malignant using support vector machine (SVM) classifier.

S. Deepa et al. in [6] proposed a method for classification of mammogram image. According to abnormality given in the MIAS database, the ROI of size 256x256 is extracted. From the decomposed image, contourlet co-efficient is computed using contourlet transform and from that, co-efficient, co-occurrence matrix is generated. With the help of co-occurrence matrix, optimal features are selected using sequential floating forward selection (SFFS) algorithm. Probabilistic neural network is used for classification.

978-1-5090-3411-6/16 $31.00 © 2016 IEEE
DOI 10.1109/ICMETE.2016.85

In 2014, Puneeth et al. [7] proposed a method for classification of mammogram image based on textural features, which is calculated using gray level co-occurrence matrix. K-Nearest Neighbors (KNN) classifier is used for classification.

In 2015, K. Vaidehi et al. [8] developed an intelligent system for retrieval of content-based mammogram image. In their proposed method, the features are calculated from the ROI. Using support vector machine image is classified and top 10 image is retrieved using KNN algorithm.

In 2015, Subashini et al. [9] developed a method where ROI is extracted from each mammogram image and from the ROI textural features are calculated using gray level co-occurrence matrix. Hybrid genetic algorithm-particle swarm optimization is used to select best features from the set of extracted features. With the help of optimal features, the image is classified using KNN algorithm.

In 2016, F Shirazi et al. in [10], presented a breast cancer detection system by combining mixed gravitational search algorithm (MGSA) and support vector machine (SVM). The authors used MIAS database and the features are extracted using gray level co-occurrence matrix (GLCM). In the experimental setup, the authors have use 70% of the dataset (out of 100 ROIs) as training dataset which includes normal and abnormal tissues. The rest of the 30% of the dataset is used as test objects and the results are computed upon that. By using only SVM with 24 features the performance was reported to be 86% whereas by the combination of MGSA – SVM with 12 features the performance was reported to be 93.1%.

III. PROPOSED SYSTEM

In this section, mammogram image classification system is proposed. The proposed system for classification of mammogram images is built based on GLCM by applying different classifiers. The system is divided into five stages to classify mammogram images. First step is the data set collection, second is ROI extraction process, third is the pre-processing steps which is again divided into two steps filtering and enhancement, fourth is the feature extraction from GLCM and last stage is classification. The architecture of the proposed system is shown in Fig. 1.

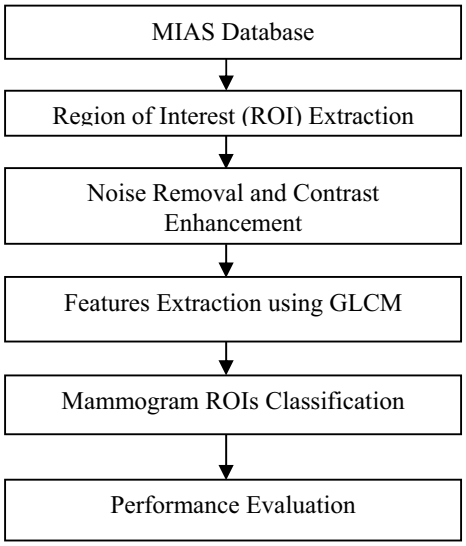


Fig. 1: Architecture of Proposed Method

A. Dataset collection

In this experiment, the mammogram images were obtained from the MIAS [11] data set. MIAS is an organization of UK research groups interested in the understanding of mammograms and has generated the digital mammograms dataset for their research. It consists of 322 images of left and right breast from 161 patients, which contains normal and abnormal images, and abnormal images are again classified into benign and malignant. The dataset provides the details, about the location and radius of the abnormalities marked by expert radiologists. In the MIAS dataset, the mammogram originally was of the size of 1024×1024 pixels.

B. Region of Interest (ROI) Extraction process

Original mammogram images have different types of noises, artifacts in their background, pectoral muscles etc which are unwanted for feature extraction and classification. Hence a cropping operation has been applied on mammogram image to extract the ROIs which contains the abnormalities, apart from the unwanted portions of the image. MIAS database gives all the details about each mammogram image, viz., size in pixels, character of background tissue, class of abnormality, *Xc* and *Yc* coordinate value of centre of abnormality, *'r'* radius of circle enclosing the abnormality by the radiologists. ROIs extraction performed by manual cropping operation considering the centre of the abnormal area as the centre of ROI and taking the approximate radius (in pixels) of a circle enclosing the abnormal area as shown in Fig. 2. For the extraction of normal ROI, the same cropping procedure has been performed on normal mammographic images with random selection of location. In this work, all the ROIs are resized in to 128 x 128 for uniformity. In this phase, the rectangular ROIs are extracted and the ROIs are free from the background information and artifacts, which is defined by equation (1) [12].

$$I_{ROI} = I[X_c - r, (1024 - Y_c) - r, 2r, 2r] \quad (1)$$

C. Pre-Processing

A mammogram is an X-ray image of the breast. It may contain noise and image quality may be poor. So preprocessing step is mostly needed to make the image suitable for classification. Thus, filtering technique is used to

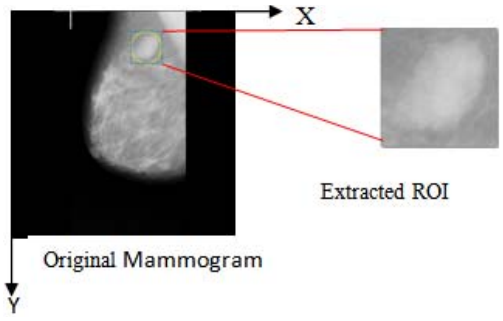


Fig. 2: Cropping of ROI

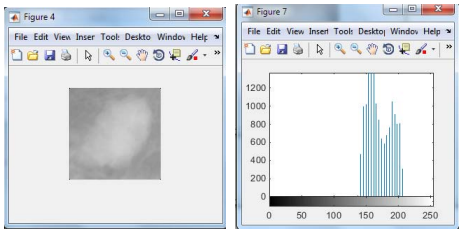
remove noise and the enhancement technique is used to improve the image quality [13].

In this work, median filter is used as it preserves the information while removing noise. PSNR and MSE [14] values of the median filter are calculated and compared with two other filter techniques. Median filter performs better by giving highest PSNR value and lowest MSE value shown in table 1. Contrast of each pixel relative to its local neighborhood is adaptively enhanced during this process which is known as Contrast Limited Adaptive Histogram Equalization and to improve the appearance of the image contrast-limited adaptive histogram equalization (CLAHE) [15] is used here. Fig. 3 shows the result of contrast enhancement process.

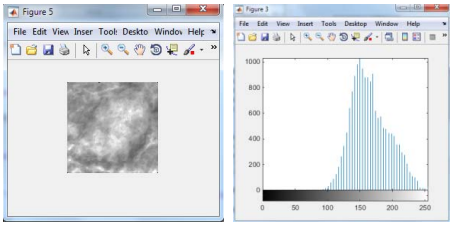
D. Feature Extraction

In image processing, processing of large data is time consuming and less efficient for classification. For reducing time, the input data is transformed into reduced set of feature vector. This transformation process is called feature extraction process. This feature vector contains relevant information and is used as input vector for classification.

Features can be classified based on color, texture and shape. In this work, we are mainly concerned about texture features and for extraction of features; Gray Level Co-occurrence Matrix (GLCM) is used since it has been proven as a powerful tool for feature extraction [16]-[17]. In this work four textural features namely contrast, correlation, Energy, homogeneity are extracted with $d = 1$ and $\theta = 0^0, 45^0, 95^0, 135^0$ and then take the average of these four direction.



(a) (b)



(c) (d)

Fig. 3: (a) De-noised image; (b) Histogram of (a); (c) Contrast enhanced image (d) Histogram of (c)

TABLE 1: PSNR AND MSE VALUE OF MAMMOGRAM SAMPLES

Image	*Averaging filter*		*Wiener filter*		*Median filter*	
	PSNR	***MSE***	***PSNR***	***MSE***	***PSNR***	***MSE***
Mdb015	3.52	28929.74	26.63	141.23	**37.28**	**12.17**
Mdb145	3.28	30543.91	26.64	140.84	**38.86**	**8.46**

E. Gray Level Co-occurrence Matrix (GLCM)

Grey level co-occurrence matrices (GLCM) are introduced in [18]-[19]. It explains the occurrence of certain grey levels in relation to other grey levels using statistical sampling. The process statement is reproduced as it is from [18]-[19] in the following paragraph.

Assume that an image to be analyzed is rectangular and has N_X rows and N_y columns. The gray level appearing at each pixel is quantized to N_g levels. Let, $L_x = \{1,2,....,N_x\}$ be the rows, $L_y=\{1,2,....,N_y\}$ be the columns and $G=\{0,1,2,....N_g\text{-}1\}$ is the total number of gray levels quantized up to N_g levels. The set L_xxL_y is the set of pixels of the image ordered by their row-column designations. Then, the image I can be represented as a function of co-occurrence matrix that assigns some gray level in L_xxL_y as I: L_xx$L_y \rightarrow G$.

The texture-context information is specified by the matrix of relative frequencies $p_{i,j}$ with two neighbouring pixels separated by distance$d$, one with gray level i and the other with gray level j. Such matrices of gray-level co-occurrence frequencies are a function of the angular relationship θ and distance d between the neighbouring pixels. By using a distance of one pixel and angles quantized to 45^0intervals, four matrices of horizontal, first diagonal, vertical, and second diagonal (0, 45, 90 and 135 degrees) are used. Then, the un-normalized frequency in those four directions is defined by equation (2).

$$p(i,j,d,\theta) = \#\left\{\begin{array}{c}((k,l),(m,n)\epsilon\left|(L_x\,X\,L_y)\,X\,(L_xXL_y)\right| \\ (k-m=0,|l-n|=d)\;or\;(k-m=d,l-n=-d) \\ or\;(k-m=-d,l-n=d)\;or\;(|k-m|=d,l-n=0), \\ or\;(k-m=d,l-n=d)\;or\;(k-m=-d,l-n=-d), \\ I(k,l)=i,\quad I(m,n)=j\end{array}\right. \quad (2)$$

Where # is the number of elements in the set, *(k, l)* the coordinates with gray level *i, (m, n)* the coordinates with gray level *j*.

Consider $p(i,j)$ be the $(i,j)^{th}$ entry in a normalized GLCM. G is the number of gray levels range from 0 to N_g-1. μ is the mean value of *p*. $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the means and standard deviations of p_x and p_y and presented in Equations (3), (4), (5) and (6) respectively.

$$\mu_x = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} i.p(i,j) \quad (3)$$

$$\mu_y = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} j.p(i,j) \quad (4)$$

$$\sigma_x = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}(i-\mu_x)^2.p(i,j) \quad (5)$$

$$\sigma_y = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}\left(j-\mu_y\right)^2.p(i,j) \quad (6)$$

In this paper, four textural features namely contrast, correlation, Energy, homogeneity are extracted from mammogram ROIs using GLCM as formulated in [18]-[19]

Contrast:

$$F1 = \sum_{n=0}^{N_g-1} n^2\{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\}, |i-j| = n \quad (7)$$

Contrast is a relative measure of the intensity between a pixel and its neighbours over the whole image and is presented in equation (7). It is the quantity of local variation present in an image.

Correlation:

$$F2 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{(ij)p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \quad (8)$$

How a pixel is correlated to its neighbor over the whole image is known as correlation and is presented in equation (8). Feature values range from -1 to 1, these extremes indicating perfect negative and positive correlation respectively.

Energy:

$$F3 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{p(i,j)\}^2 \quad (9)$$

Energy also known as uniformity or angular second moment (ASM). Energy measure the sum of squared elements in the GLCM as presented in equation (9). Basically the property of energy is provide how uniform the texture image. The range of energy is [0 1], where Energy is 1 for a constant image.

Homogeneity:

$$F4 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i,j)}{1+(i-j)^2} \quad (10)$$

It measures the closeness of the distribution of the elements in the GLCM to the GLCM diagonal and is defined in equation (10). The range of homogeneity is [0 1], where homogeneity is 1 for a diagonal GLCM. It is high when local gray level is uniform and inverse GLCM is high.

F. Classification

Last step of the proposed method is the classification of the mammogram image into normal or abnormal. Here, the classification is done mainly with the help KNN algorithm and compared with SVM and ANN classifiers.

KNN is a supervised learning that has been used in many applications in the field of data mining and pattern recognition. This classifier computes the distance from the unlabeled data to every training data point and selects the best k-neighbor with the shortest distance. This classifier implementation is very simple since there is no need of explicit training step. The input to the classifier is the k-closest training samples and the output is the class name or a class membership. The output is decided by majority vote of its neighbours, where the input is being assigned to the class most frequent among its k-nearest neighbours. When k is considered as 1, then the input is just allocated to that class.

In this present study 1NN and 3NN are considered and Euclidean distance is used for calculating the distance between the new samples and training samples. Different k values give different results.

IV. EXPERIMENTAL RESULT AND DISCUSSION

To carry out the research, 208 mammogram images are considered out of which 188 images are used as a training set and 20 (10 normal and 10 abnormal) images are taken as testing set. From the de-noised and contrast enhanced ROI based mammogram images, the four texture features are calculated with the help of gray level co-occurrence matrix. These features are used as an input to the KNN classifier for training and testing.

TABLE 2: GLCM FEATURES VALUE FOR TEST DATASET

Image	*Contrast*	*Correlation*	*Energy*	*Homogeneity*
mdb190	0.1606	0.9073	0.2173	0.9225
mdb193	0.1313	0.8719	0.3492	0.9366
mdb199	0.1493	0.8560	0.2947	0.9282
mdb204	0.1768	0.9293	0.1929	0.9134
mdb208	0.1354	0.9616	0.2595	0.9347
mdb209	0.1510	0.9518	0.1814	0.9267
mdb212	0.1527	0.9180	0.2150	0.9255
mdb214	0.1868	0.9414	0.1504	0.9083
mdb219	0.1919	0.9470	0.1537	0.9060
mdb226	0.1868	0.9507	0.1382	0.9101
mdb234	0.2052	0.9331	0.1580	0.9000
mdb245	0.1767	0.8382	0.2810	0.9150
mdb246	0.0784	0.8818	0.5188	0.9639
mdb252	0.1616	0.9620	0.1402	0.9206
mdb257	0.1307	0.8268	0.3763	0.9362
mdb264	0.1496	0.9113	0.2591	0.9281
mdb272	0.1812	0.8195	0.2968	0.9122
mdb282	0.1501	0.7991	0.3470	0.9270
mdb302	0.1571	0.8367	0.3081	0.9236
mdb304	0.1476	0.9099	0.3128	0.9288

To evaluate the performance of the proposed system the following parameters are used which is defined by the equations (11),(12) and (13)[20].

$$\text{Accuracy} = \frac{TP+TN}{N} \times 100\% \quad (11)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100\% \quad (12)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100\% \quad (13)$$

Where, True positive (TP) = the mammogram image predicted with abnormal when actually the mammogram image is abnormal.

True negative (TN) = the mammogram image predicted with normal when the mammogram image actually normal.

False positive (FP) = the mammogram image predicted with abnormal when the mammogram image is normal.

False negative (FN) = the mammogram image predicted with normal when the mammogram image is abnormal.

Accuracy defines the overall correctness of the classifier, specificity defines true negative rate and sensitivity defines true positive rate. The higher values of both sensitivity and specificity show better performance of the system. Tables 3, 4, 5 and 6 show the classification result using 1NN, 3NN, SVM and ANN respectively with the help of confusion matrix. The overall performance of the proposed method is shown in table 7. The performance of the four classifiers are compared and represented graphically in fig. 4.

From the graphical representation of fig. 4, it is revealed that 3NN classifier outperforms other three in terms of accuracy and specificity whereas in terms of sensitivity 3NN outperforms SVM and ANN clearly and equals with 1NN. Thus, 3NN can be considered as a more efficient classifier for the classification of the mammogram ROIs images than 1NN, SVM and ANN.

TABLE 3: CONFUSION MATRIX FOR 1NN

Target class	**Predicted class**	
	Normal	***Abnormal***
Normal	7(TN)	3(FP)
Abnormal	0(FN)	10(TP)

TABLE 4: CONFUSION MATRIX FOR 3NN

Target class	**Predicted class**	
	Normal	***Abnormal***
Normal	10(TN)	0(FP)
Abnormal	1(FN)	9(TP)

TABLE 5: CONFUSION MATRIX FOR SVM

Target class	**Predicted class**	
	Normal	***Abnormal***
Normal	6(TN)	4(FP)
Abnormal	3(FN)	7(TP)

TABLE 6: CONFUSION MATRIX FOR ANN

Target class	**Predicted class**	
	Normal	***Abnormal***
Normal	7(TN)	3(FP)
Abnormal	2 (FN)	8(TP)

TABLE 7: CLASSIFICATION RESULS

Methods	**Accuracy %**	**Sensitivity %**	**Specificity %**
1NN	85	100	70
3NN	95	100	90
SVM	65	70	60
ANN	75	80	70

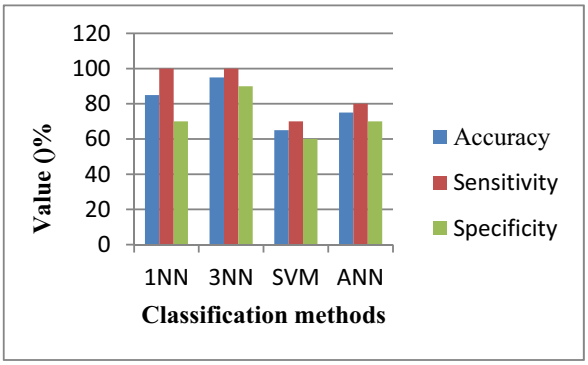


Fig. 4: Performance comparison of 1NN, 3NN, SVM, ANN classifiers

V. CONCLUSION AND FUTURE WORK

The CAD system is developed and presented here for the classification of mammogram into normal and abnormal breast tissues with the aim to support the radiologists in visual diagnosis. The experimental results show that 3NN gives the maximum accuracy rate for normal and abnormal classification (96%) compared to other classifiers.

The present work can be extended to the whole MIAS database with 322 mammogram images or other mammogram databases too. In future, more number of other statistical movement features may be considered with proper feature selection technique and accuracy may be improved further. A new mammogram database can be created by collecting mammogram images from different clinics and hospitals and this proposed method may be tested on that database.

REFERENCES

[1] American Cancer Society, "*Breast cancer facts & figures 2015-2016*," Atlanta, American Cancer Society, Inc. 2015.

[2] L. Tabar, P.Dean, *"Mammography and breast cancer: the new era,"* International Journal of Gynecology and Obstetrics, vol. 82, Issue 3, September 2003, pp. 319-326.

[3] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai, H. Du, "*Approaches for automated detection and classification of masses in mammograms,"* Pattern recognition, vol. 39, 2006, pp. 646-668.

[4] 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, October 2012, pp. 87-98.

[5] Maanasa N A S, V Gowri, "*Segmentation of mammogram using tumor-cut algorithm,*" International Journal of Engineering and Innovative Technology, vol. 2, Issue 10, April 2013, pp. 172-175.

[6] S. Deepa, V. Subbiah Bharathi, *"Textural feature extraction and classification of mammogram images using CCCM and PNN,"* IOSR Journal of Computer Engineering, vol. 10, Issue 6, 2013, pp. 07-13.

[7] Puneeth L., Krishna A.N. "*Classification of mammograms using texture features,*" International Journal of Innovative Research & Development, vol. 3, Issue 7, July 2014, pp. 373-377.

[8] K. Vaidehi, T. S. Subashini, *"An intelligent content based image retrieval system for mammogram image analysis,"* Journal of Engineering Science and Technology, vol. 10, 2015, pp.1453 – 1464.

[9] Subashini Sundaravinayagam, Bhavani Sankari. S, "*Detection and classification of masses in mammograms using a hybrid GA-PSO-KNN approach,*" International Journal of Advanced Research Trends in Engineering and Technology, vol. 2, May 2015.

[10] F. Shirazi, E. Rashedi, "*Detection of cancer tumors in mammography images using support vector machine and mixed gravitational search algorithm*," In proceedings of IEEE 1st Conference on Swarm

Intelligence and Evolutionary Computation, Bam, March 2016, pp. 98-101, doi:10.1109/CSIEC.2016.7482133.

[11] S. A. J Suckling, D Betal, N Cerneaz, D R Dance, S-L Kok, J Parker, I Ricketts, J Savage, E Stamatakis and P Taylor, "*The mammographic image analysis society digital mammogram databaseExerpta medica,*" International Congress Series 1069, 1994, pp. 375-378.

[12] R. N. Panda, M. A. Baig, B. K. Panigrahi, M. R. Patro "*Efficient CAD system based on GLCM & derived feature for diagnosing breast cancer*," International Journal of Computer Science and Information Technologies, vol. 6, 2015, pp. 3323-3327.

[13] Mister Khan, Sumit Kataria and Smriti, "*Mammography based breast cancer detection using different classifiers,*" National Conference on Emerging Trends in Electronics & Communication, vol. 1, No. 2, July 2015.

[14] K. Malathi, R. Nedunchelian, "*Comparision of various noises and filters for fundus Images using pre-processing techniques*," International Journal of Pharma and Bio Sciences, vol. 5, july 2014, pp. 499-508.

[15] K. Zuiderveld, "*Contrast Limited Adaptive Histogram Equalization*," Academic Press Inc, 1994.

[16] Pradeep N, Girish H, Karibasappa K, "*Segmentation and feature extraction of tumors from digital mammograms*," Computer Engineering and Intelligent Systems, vol 3, No.4, 2012.

[17] Biswajit Pathak, Debajyoti Barooah, "*Texture analysis based on the graylevel co-occurence matrix considering possible orientations,*" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, September 2013.

[18] Robert M. Haralick, K. Shanmugam, and ITS'HAK Dinstein, *"Textural features for image classification,"* IEEE Transaction on system, man and cybernetics, vol.SMC-3, No.6, November 1973, pp. 610-621.

[19] L.K. Soh, and C. Tsatsoulis, "*Texture analysis of sar sea ice imagery using grey level co-occurrence matrices,*" IEEE Transactions on Geoscience and Remote Sensing, vol.37, no. 2, March 1999, pp. 780-795.

[20] R. Nithya B. Santhi, "*Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast cancer*," International Journal of Computer Applications, vol. 28, August 2011, pp. 21-25.